Turkish Journal of Botany

http://journals.tubitak.gov.tr/botany/

Research Article

Turk J Bot
(2017) 41: 338-346
© TÜBİTAK
doi:10.3906/bot-1608-17

Molecular characterization of edible pea through EST-SSR markers

Mohammad NISAR[1], Asaf KHAN[2], Syed Fazal WADOOD[1], Aftab Ali SHAH[2], Fatih HANCI[3,*]
[1]Department of Botany, University of Malakand, Khyber Pakhtunkhwa, Pakistan
[2]Department of Biotechnology, University of Malakand, Khyber Pakhtunkhwa, Pakistan
[3]Atatürk Central Horticultural Research Institute, Yalova, Turkey

Received: 12.08.2016 • **Accepted/Published Online:** 09.03.2017 • **Final Version:** 18.07.2017

Abstract: The genetic diversity among 23 newly developed homosegregate pea lines (*Pisum sativum* L.) was assessed with a total of 13 expressed sequence tag (EST) based-simple sequence repeat (SSR) markers. The percentages of amplified and nonamplified primers were 92% and 8%, respectively, and 58.33% of the used primers gave the PCR product within the reported size range while 41.66% of primers gave a different product size. Polymorphism information content (PIC), major allele frequency, and variation in genetic diversity were calculated. The PIC ranged from 0.32 to 0.63 with an average of 0.50. Major allele frequency ranged from 0.48 to 0.78 with a mean value of 0.56. The variation in genetic diversity among these pea lines ranged from 0.36 to 0.68 with a mean value of 0.56. Cluster analysis based on a dendrogram divided the 23 pea lines into two main groups (L-1 and L-2), separated at 25% genetic distance. Seven subclusters were evident from these two main groups. L-1 grouped 51.2% (12 pea lines) while L-2 contained 47.8% (11 pea lines) of the total analyzed population. It was concluded that EST-SSR markers are useful for refinement of the pea linkage map.

Key words: Pea, expressed sequence tags, cluster analysis, genetic diversity, polymorphism information content

1. Introduction

Economically, legumes are the second best-known crop family with about 27% of the total world's crop production (Graham and Vance, 2003), and they are the 3rd topmost family of flowering plants with more than 650 genera and 18,000 species (Lewis et al., 2005). Among these, the pea (*Pisum sativum* L.) holds significant environmental benefits (Smykal et al., 2012) and has a rich history in genetic research. Pea is among the six major pulse crops cultivated throughout the world and is ranked the 2nd highest yielding legume globally after the common bean (Kumari et al., 2013). Mature kernels of the pea contain protein (23%–25%), slowly digestible starch (50%), soluble sugars (5%), vitamins, fiber, and minerals (Mahesh et al., 2014). Different stresses are constant threats to the field pea and the key to pea improvement programs is to have genetic variation for agronomically important traits within the parental lines that are used in crosses for maximizing the genetic gain (Taran et al., 2005). For this purpose, assessment of genetic diversity between germplasm/varieties to utilize their protection, conservation, and registration is also used in breeding programs to provide the bulk of allelic variation in breeding material (Jain et al., 2014).

Genetic diversity can be defined as variation and polymorphism at the DNA level. Allelic diversity enables certain species to increase and get healthier in new and difficult environments, ensuring long-term survival, and so is considered necessary for global food security (Able et al., 2007). Genetic diversity can be estimated through morphological, biochemical, and molecular (DNA) markers. Morphological markers are largely affected by environmental factors compared to biochemical markers. Therefore, molecular markers are the means to overcome the limitations associated with these markers (Rao, 2004) as estimation of genetic diversity through molecular markers is independent of environmental influences (Tatikonda et al., 2009). Various molecular markers such as intersimple sequence repeat (ISSR) (Zietkiewiez et al., 1994), sequence-related amplified polymorphism (SRAP) (Li and Quiros, 2001), single nucleotide polymorphism (SNP) (Jain et al., 2014), amplified fragment-length polymorphism (AFLP), random amplified polymorphic DNA (RAPD), and simple sequence repeats (SSR) markers are predominantly used for molecular characterization in plants (Belaj et al., 2003).

The effectiveness of these different DNA markers has been employed in several plant species (Jain et al., 2014).

* Correspondence: tanerfatih@gmail.com

Belaj et al. (2003) reported that SSR markers show the highest level of polymorphism compared to RAPD and AFLP in olive. The same results were reported in various other studies where SSR markers were compared to other molecular markers (Jain et al., 2014). Therefore, SSRs have increasingly become the marker set of choice for genetic studies as they are abundant, highly polymorphic, codominant in nature, and genome-specific (Cuevas and Prom, 2013; Izzah et al., 2013). Being so important, SSR markers were previously used for assessing genetic diversity in pea (Taran et al., 2005; Nasiri et al., 2009), but anonymous SSRs are more costly and time-consuming to develop from genomic libraries (Ramu et al., 2013). An alternate to SSR markers is to develop expressed sequence tag (EST)-based SSRs through available EST databases (Simko, 2009).

ESTs are obtained through partial sequencing of random cDNA clones and are fundamental molecular marker sources widely used nowadays. Markers from ESTs provide opportunities in the identification of unique genes and can increase the role of genetic markers by analyzing variation in the transcribed and known functional genes (Varshney et al., 2005). EST-SSRs have several fundamental advantages over traditional genomic SSR markers, such as being embedded in functional gene sequences, in direct association with transcribed genes, less costly, and highly transferable between related species (Varshney et al., 2005). EST-SSR markers have been used in several studies, such as evaluation of genetic diversity in grapevine, coffee, and sugar cane and in genetic map integration in soybean and genetic mapping of wheat, potato, and cotton (Missio et al., 2009). In pea, EST-SSR markers have been used to determine genetic diversity (Burstin et al., 2001). Loridon et al. (2005) used EST-SSR markers for molecular mapping in pea. Gong et al. (2010) also used EST-SSR markers in pea to determine its polymorphism and transferability.

In Pakistan, peas are grown on 10,000 ha with an overall production of 82,000 t. In Pakistan, the yield of peas per unit is less than the international average because of different factors such as poor weed control, poor cultural practices, and pest attacks. Among these, pests and diseases are the main reasons for the low yield (Khan et al., 2013). Therefore, molecular characterization and estimation of the genetic diversity of pea genotypes grown in this area are required to further improve the breeding programs. The present study was designed to estimate the level of genetic diversity in newly developed pea lines through EST-SSR markers.

2. Materials and methods

2.1. Plant material

During 2004 and 2005, a total of 177 pea genotypes were planted in the field of PGRI, NARC, Islamabad, Pakistan, for selection of powdery mildew resistant/susceptible plants. Out of 177 lines, the Fallon and 11760-3ER genotypes were highly resistant and susceptible, respectively. Through artificial inoculation, it was genetically confirmed that Fallon is a powdery mildew-resistant genotype from the 177 germplasm (Nisar et al., 2006, 2008). The genotypes Fallon (resistant/low yield) and 11760-3ER (susceptible/high yield) were crossed for this study (Nisar and Ghafoor, 2010). After crossing, the seeds of F_1 generations were grown for separation of morphological traits (Table 1). To get the homozygous population, the F_1 seeds were consecutively sown (2005–2015) until the F_{10} generation. A total of 20 different combinations based on morphological differences were developed and selected for the estimation of genetic diversity using EST-SSR markers (Table 2). In

Table 1. Nine contrasting traits between the selected parents of pea lines (Nisar et al., 2008).

S/No.	Contrasting traits	Parents	
		Fallon (female parent)	11760-3 (male parent)
1	Anthocyanin pigmentation	Absent	Present
2	Plant height	Dwarf	Tall
3	Plant color	Creamy	Purple
4	Pod color	Green	Purple
5	Tendril	Bushy	Normal
6	Numbers of leaflets	Leafless	Leaflets present
7	Seed color	Creamy	Brown
8	Seed texture	Smooth	Rough
9	Disease	Resistant	Susceptible

Table 2. Based on morphological traits separations, 20 different combinations developed from single parental cross of Fallon and 11760-3 genotypes.

PL	AP	NC	FLC	PC	TT	PH	PSH	SS	SC
PL-1	Present	Pink	Pink	Pink	Normal	Tall	Straight	Wrinkle	Brownish
PL-2	Present	Pink	Pink	Green	Bushy	Tall	Curve	Wrinkle	Brown
PL-3	Absent	White	White	Green	Bushy	Dwarf	Straight	Rough	Creamy
PL-4	Absent	White	White	Green	Normal	Tall	Straight	Rough	Creamy
PL-5	Present	Pink	Pink	Green	Normal	Dwarf	Straight	Wrinkle	Brownish
PL-6	Absent	White	White	Green	Normal	Dwarf	Straight	Wrinkle	Creamy
PL-7	Present	Pink	Pink	Green	Normal	Dwarf	Straight	Wrinkle	Brown
PL-8	Present	Pink	Pink	Green	Normal	Tall	Straight	Wrinkle	Brown
PL-9	Present	Pink	Pink	Pink	Normal	Tall	Straight	Wrinkle	Brown
PL-10	Absent	White	White	Pink	Normal	Tall	Straight	Wrinkle	Creamy
PL-11	Present	Pink	Pink	Pink	Normal	Dwarf	Straight	Wrinkle	Brownish
PL-12	Absent	White	White	Pink	Normal	Dwarf	Curve	Wrinkle	Creamy
PL-13	Absent	White	White	Pink	Normal	Tall	Curve	Wrinkle	Creamy
PL-14	Absent	White	White	Pink	Bushy	Tall	Curve	Rough	Creamy
PL-15	Present	Pink	Pink	Pink	Normal	Tall	Curve	Wrinkle	Brownish
PL-16	Present	Pink	Pink	Pink	Normal	Tall	Curve	Wrinkle	Brown
PL-17	Present	Pink	Pink	Pink	Normal	Dwarf	Straight	Wrinkle	Brownish
PL-18	Present	Pink	Pink	Pink	Normal	Tall	Straight	Wrinkle	Brown
PL-19	Present	Pink	Pink	Pink	Normal	Tall	Straight	Wrinkle	Brown
PL-20	Present	Pink	Pink	Pink	Normal	Tall	Straight	Wrinkle	Brown
PL-21	Absent	White	White	Green	Normal	Dwarf	Curve	Wrinkle	Creamy
PL-22	Absent	White	White	Green	Normal	Dwarf	Curve	Wrinkle	Creamy
PL-23	Absent	White	White	Green	Bushy	Dwarf	Curve	Rough	Creamy
PL-24	Present	Pink	Pink	Pink	Normal	Tall	Straight	Wrinkle	Brown

AP = Anthocyanin pigmentation, NC = node color, FLC = flower color, PC = pod color, TT = tendril type, PH = plant height, PSh = pod shape, SS = seed shape, SC = seed color.

the experiment, cultivar Climex was used as a control cultivar [pea line-22 (PL-22)].

2.2. DNA isolation

Total genomic DNA was isolated according to Dellaporta et al. (1983) with some modifications. Ten seeds of each line were ground into a fine powder. About 400 µL of DNA extraction buffer and phenol chloroform/isoamyl alcohol (24:24:1) were added to 0.5 g of powder and mixed well in 1.5-mL Eppendorf tubes using a vortex. The extraction buffer (200 mL, pH 8.5) contained 0.5% SDS, 200 mM Tris-HCl (pH 8.0), 25 mM EDTA, and 200 mM NaCl. The crude homogenate was then centrifuged at 13,000 rpm for 13 min at 25 °C using a tabletop centrifuge (Dynamica, UK). The supernatant was transferred to a fresh tube, to which 400 µL of isopropanol and 50 µL of 4% sodium acetate were added. The tubes were centrifuged at 8000 rpm for 8 min and a white nucleic acid pellet was formed at the bottom of tubes. The pellet was washed with 70% ethanol and subsequently dissolved in TE buffer after getting completely dried. The samples were then kept in the refrigerator and the quality of the DNA was confirmed using agarose gel electrophoresis.

2.3. Detection of genomic DNA

The genomic DNA was run on 1% agarose gel. The gel was prepared by taking 1 g of agarose powder and dissolving it in 100 mL of TBE. The mixture was boiled in a microwave oven to dissolve the agarose completely. After slightly cooling, about 3 µL ethidium bromide was added to the medium. The gel was cast in a gel tray with an inserted comb. After solidification, the comb was removed and the gel was shifted to electrophoresis containing 1X TBE. Genomic DNA (1 µL) was loaded in each well and electrophoresis was run at a constant voltage of 100 V with 120 mA for 15 min. The DNA was visualized under UV light using a gel documentation system (Vilber Lourmat, USA).

2.4. EST-SSR assay

This assay was carried out using thirteen EST-SSR primers (Table 3). Polymerase chain reaction (PCR) was performed using thermal cycler model SC 300 (Australia). Each reaction tube (20 µL) contained 23% of PCR Master Mix [PCR water = 14.8 µL, 10X buffer = 2.5 µL (Thermo Scientific), $MgCl_2$ = 25 mM (1.5 µL), dNTPs = 10 mM (0.5 µL), primers F = 10 pmol (0.5 µL), primers R = 10 pmol (0.5 µL), Taq DNA polymerase = 0.2 µL (Thermo Scientific)] with 25 ng/µL template DNA. The reaction conditions were an initial denaturation at 95 °C for 5 min, followed by 94 °C for 30 s, annealing temperature of 54 °C for 30 s, and extension at 72 °C for 1 min, with a final extension of 5 min at 72 °C. The PCR product was then separated on 20% polyacrylamide gel electrophoresis (PAGE). The product of PCR was visualized using ethidium bromide staining solution and was photographed using a gel documentation system (Vilber Lourmat).

2.5. Data analysis

For data analysis, each amplified band was defined as a single character. The allelic data for all genotypes were scored in the form of binary matrix "zero one table" where "0" represented the absence and "1" represented the presence of bands, and they were subjected to analysis with PC-ORD software (McCune and Grace, 2005). A simplified representation of genetic distances was based on a dendrogram obtained from the unweighted pair group method based on arithmetic averages (UPGMA). Power Marker program version 3.25 (Liu and Muse, 2005) was used to calculate allele frequencies at each locus and gene diversity H value according to Nei (1973). The polymorphic information content (PIC) value for each marker was determined using the following equation:

$$PIC = 1 - \sum_{i=1}^{n} p_i^2 - 2\sum_{i=1}^{n-1}\sum_{j=i+1}^{n} p_i^2 p_j^2$$

where p_i is the frequency of the ith allele, and n is the number of alleles (Botstein et al., 1980).

3.Results

3.1. Marker efficiency and allelic richness

A total of 13 EST-SSR markers were used in the study as shown in Table 3. Among these 13 primers, 12 primers were successfully amplified, while primer PEA121 did not amplify any marker. Variation based on alleles was clearly detected among the primers. The number of polymorphic alleles ranged from two to five (Table 4). In PSAT8487 the number of polymorphic alleles was five, while only two polymorphic alleles were observed in PEA063 and PEA069. The number of polymorphic alleles was four in both PSAT7598 and PEA128, while three polymorphic alleles were detected for the remaining primers. The percentage of three polymorphic alleles was 58%, followed by 16% for both two and four polymorphic alleles. The mean value of alleles per locus was 4.69. The frequency of major alleles ranged from 0.48 to 0.78. Major allele frequency was recorded the same, i.e. 0.48 and 0.61, in three out of 12 amplified primers. In PEA120 the highest major allele frequency of 0.78 was observed. The mean value of major allele frequency obtained was 0.56.

3.2. PCR product variability

PCR products were compared with a 50-bp DNA ladder (Figure 1). Seven out of 12 primers were found to have allele size well within the reported range. It was observed that four markers showed different PCR products from the reported range. These included PEA143, PEA063, PEA251, PEA128, and PEA132. Their reported ranges are 269–293 bp, 300–312 bp, 242 bp, 202–217 bp, and 262–282 bp, respectively, while they were observed in ranges of 200–210 bp, 181–197 bp, 128–134 bp, 88–102 bp, and 137–147 bp, respectively. The percentage of primers within their reported size ranges was 58.34%, while 41.66% of primers were found to have a size range different from that reported.

3.3. Gene diversity and polymorphism information content

Gene diversity ranged from 0.36 to 0.68 (Table 4). High gene diversity of 0.68 was recorded in PSAT8487, followed by 0.67 in PSAT7598. Primer PEA120 was found to be the primer with the lowest recorded gene diversity. The average gene diversity calculated was 0.56. The PIC value ranged from 0.32 to 0.63. A high PIC value of 0.63 was recorded in PSAT8487, followed by PSAT7598 with a PIC value of 0.62. The lowest PIC value of 0.32 was recorded in PEA120. The mean PIC value observed was 0.50.

3.4. Cluster analysis

To find out the genetic diversity based on DNA, the analysis of 23 genotypes of pea was subjected to PC-ORD software. The dendrogram obtained from the analysis of 23 pea lines is divided into two linkages (L-1 and L-2) at

Table 3. List of EST-SSR primers used for estimation of genetic diversity in pea.

Primer	Accession no.	Sequence	Ta (°C)	Putative function	References
P251	32543524-1	F-ATCCAGAACTCACAACAT R-TAGAATCAAAACACGACC	55	P54 protein	Gong et al., 2010
P1109	32545076-1	F-CTCCATCTCAAGAAATCC R-CACATAACTAAAAAACCC	55	Histone H1 SUT1 subtype 7	Gong et al., 2010
PEA032	FG534835.1	F-ACCGTCTGATTTGATTAC R-CTCTGCCAACTATGTCCT	55	GRAS family transcription	Xu et al., 2012
PEA063	FG535978.1	F-TGCTGGGACTGCGATTCTA R-ATCCTCATCACCGTCAACC	56	COL domain class	Xu et al., 2012
PEA069	FG529639.1	F-CACCCACTCATTGAGATTA R-ACATACAGCAGCATTACACT	56	Nuclear acid binding protein	Xu et al., 2012
PEA090	FG533337.1	F-TGATGGAAGATGGGAAGA R-ATGGCATAGCAACAAGGA	56	Hypothetical protein MtrDRAFT_ AC148289g12v2	Xu et al., 2012
PEA120	FG533943.1	F-TCGTCACCGATTCAGTTC R-ACGGAGGAGCGATAGGAT	55	Catalytic/oxidoreductase, acting on NADH or NADPH	Xu et al., 2012
PEA121	CA902458.1	F-TGGATGTTAATTTGAGGGTG R-AAGGTCACTTGCCTGTTG	55	MADS transcription factor	Xu et al., 2012
PEA128	EX570575.1	F-GAGGTGCTTAGGCTTTC R-TGGCTCCAATTCATTCATA	55	Cytochrome P450 monooxygenase CYP78A29	Xu et al., 2012
PEA132	FG538740.1	F-GACACTGCTCCTCCACGAA R-CCCTGCCGCATGTACCTTA	56	110 kDa 4SNcTudor domain protein	Xu et al., 2012
PEA143	GH720573.1	F-ATCTTACTGCCATCTCC R-TAGTCATTCATGCCACA	56	Sucrose transport protein SUT1	Xu et al., 2012
Psat8487	JR344281	F-TGTTTCCAGAAGGTTATGGCCC R-AGATTCTTCGTTGCCTTTGCTTTGA	54	HXXXD-type acyltransferase	Zhuang et al., 2013
Psat7598	JR344275	F-ACTACAGGAGTTGAATTTGCGGA R-CAACATCAACAAAACAAGAACACG	54	Basic helix-loop-helix protein	Zhuang et al., 2013

T_a = Annealing temperature.

Table 4. Summary of each primer pair showing genetic diversity and polymorphism information contents in analyzed pea lines.

Marker	SS	NO	AN	AR (bp)	RS (bp)	A	MAF	GD	PIC
P251	23	23	3	128–134	242	1	0.65	0.5	0.43
P1109	23	23	3	385–405	383	1	0.48	0.64	0.56
PEA032	23	23	3	170–190	169–190	1	0.61	0.54	0.48
PEA063	23	23	2	181–197	300–312	1	0.61	0.48	0.36
PEA069	23	23	2	163	141–176	1	0.65	0.45	0.35
PEA090	23	23	3	150–160	161–179	1	0.57	0.55	0.47
PEA120	23	23	3	150–160	157–178	1	0.78	0.36	0.32
PEA128	23	23	4	88–102	202–217	1	0.61	0.55	0.49
PEA132	23	23	3	137–147	262–282	1	0.61	0.51	0.42
PEA143	23	23	3	200–210	269–293	1	0.7	0.46	0.41
Psat7598	23	23	4	190–220	209	1	0.48	0.67	0.62
Psat8487	23	23	5	150–170	155	1	0.48	0.68	0.63
Mean	23	23	4.69			1	0.56	0.56	0.5

SS = Sample size, NO = number of observations, AN = allele number, AR = allele range, RS = reported size, A = availability (maximum value), MAF = major allele frequency, GD = gene diversity, PIC = polymorphism information content.

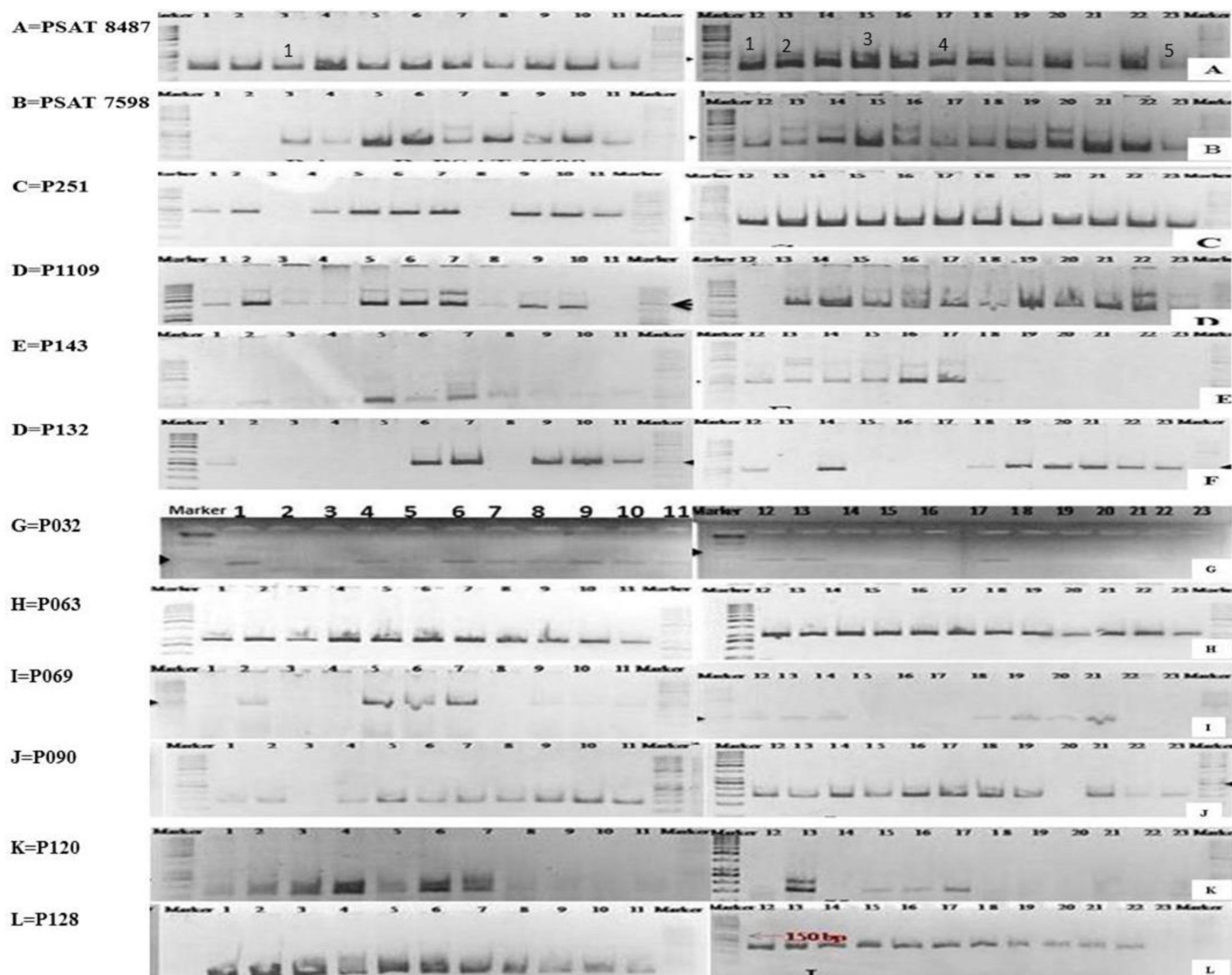

Figure 1. Amplification pattern of 23 pea lines using EST-SSR markers.
A = Primer PSAT-8487, B = primer PSAT-7598, C = primer P251, D = primer P1109, E = primer P143, F = primer P132, G = primer P032, H = primer P065, I = primer P069, J = primer P090, K = primer P120, L = primer P128.

25% genetic distance (Figure 2). L-1 grouped 51% (12 pea lines) while L-2 contained 47.82% (11 pea lines) of the total population. The obtained dendrogram was further classified into seven clusters. Linkage-1 has three clusters, i.e. cluster-1, cluster-2, and cluster-3. Linkage-2 consists of four clusters that are cluster-4, cluster-5, cluster-6, and cluster-7. Cluster-1 of linkage-1 consists of four pea lines: PL-1, PL-2, PL-3, and PL-4. Cluster-2 of the linkage grouped PL-5 and PL-6, while cluster-3 of the same linkage clustered genotypes PL-7, PL-8, PL-10, PL-9, PL-11, and PL-12. Among the four clusters of linkage-2 the genotypes PL-13 and PL-14 were clustered in linkage-2. Similarly, PL-15 and PL-16 were grouped together in cluster-5. The remaining two clusters, i.e. cluster-6 and cluster-7, gathered genotypes PL-19, PL-20, PL-21, PL-22, PL-23, PL-17, and PL-18, respectively. The genotypes in one cluster are most likely identical at the DNA level and have less genetic diversity.

4. Discussion

The application of modern molecular markers in pea includes marker-assisted selection, identification of regions affecting quantitative trait loci (Taran et al., 2005), and estimation of diversity (Baranger et al., 2004). Comprehensive analysis of genetic diversity could be useful for genetic and genomic analysis and the utilization of genetic variation in pea breeding. The present study was designed to investigate changes in the genetic diversity in pea by the analysis of microsatellites (EST-SSRs). A total of 23 new pea lines were selected. The results of 20 EST-SST primer pairs showed clear bands and were applied to investigate the genetic diversity among selected pea lines.

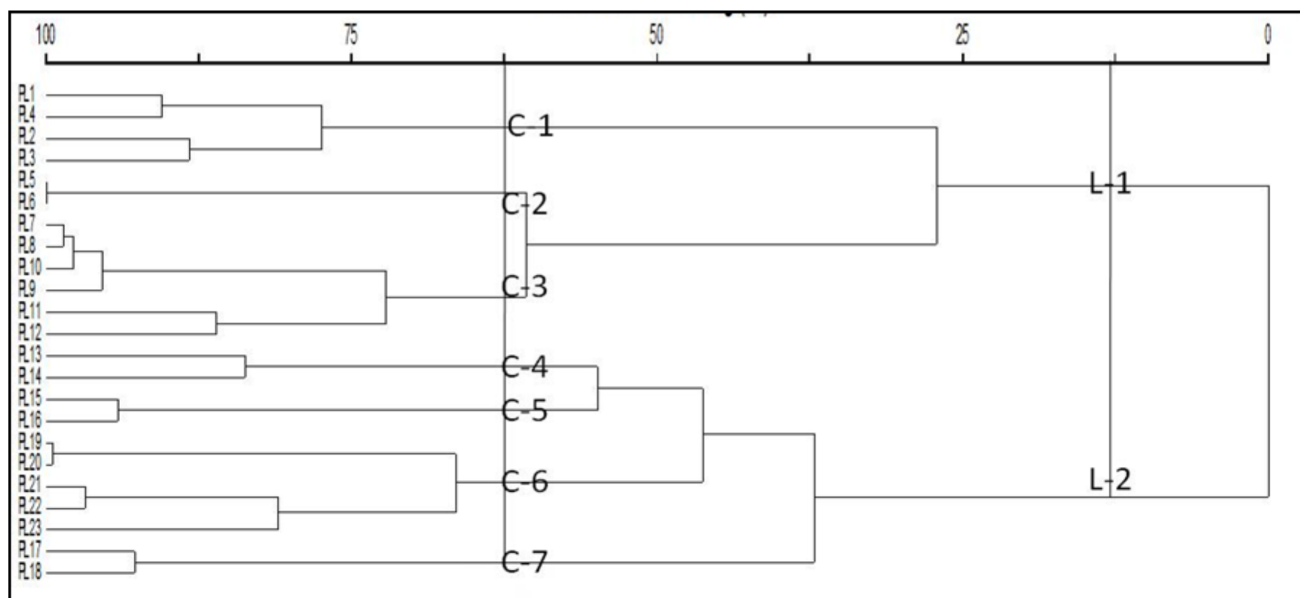

Figure 2. Dendrogram illustrating coefficient similarities among 23 pea genotypes based on data obtained from EST-SSR markers. L = Linkage, C = cluster.

It was observed that five markers showed a different PCR product size range compared to the already reported size, which is in accordance with a previous observation (Gong et al., 2010). The PIC value is a measure of polymorphism for a marker locus used in linkage analysis (Shete et al., 2000) and determines the level of allelic variation. In our study, the PIC value ranged from 0.32 to 0.63 with a mean of 0.50, compared to 0.18–0.58 with mean of 0.41 as previously reported (Gong et al., 2010). This value varied from 0.055 to 0.660 with a mean of 0.460 in the work of Ahmad et al. (2012). Usually genetic variation is assessed empirically from statistics based on population gene frequencies, but alternative statistics based on allelic diversity can be useful in obtaining relevant information (Caballero and García-Dorado, 2013).

The advancement of molecular techniques has made it possible to observe genetic erosion at the allelic level. Richness with alleles is important both evolutionary as well as for breeders' aspects as a source for the continuous development and adaptation of the crop (Wouw et al., 2010). It has been reported that Turkish pea revealed 6.1 average alleles per SSR locus (Sarıkamış et al., 2010). Similarly, 2 to 4 alleles per locus were reported in Spanish pea landraces (Martin-Sanz et al., 2011). Using SSR markers, Hagenblad et al. (2014) reported 5 to 10 alleles in the Swedish garden pea. The number of alleles per locus in the present study ranged from 2 to 5 with the mean value of 4.69 compared to the work of Zhuang et al. (2013), where the allele number ranged from 2 to 4. Another study also showed a variable number of alleles ranging from 1 to 7 using EST-SSR markers (Xu et al., 2012). Similarly, allele number per locus averaged 3.1 in the work of Teshome et al. (2015). Observed heterozygosity ranged from 0 to 0.43 in the work of Zhuang et al. (2013). In our study the range of gene diversity was observed from 0.36 to 0.68, which differs from 0 to 0.8889 as observed previously (Xu et al., 2012). The average gene diversity was 0.346 for each of the primer pairs, as shown previously (Cieslarová et al., 2012). Earlier studies in pea using EST-SSRs and genomic SSRs revealed similar results with most polymorphic loci having 7 alleles (Burstin et al., 2001; Loridon et al., 2005).

There are two main ways of analyzing the resulting distance (or similarity) matrix, namely principal coordinate analysis (PCA) and dendrograms (or clustering, tree diagrams). PCA is used to produce a two- or three-dimensional scatter plot of the samples such that the distances among the samples in the plot reflect the genetic distances between them with minimum distortion. Another approach is to produce a dendrogram (or tree diagram), a grouping of samples together in clusters that are more genetically similar to each other than to samples in other clusters (Govindaraj et al., 2015). Our study divided all 23 pea lines into two main linkages that were further subdivided into seven subclusters, comparable to previous studies where cluster analysis grouped 35 pea accessions into two major clusters and eight subclusters (Gong et al., 2010; Xu et al., 2012; Zhang et al., 2013). Similarly, the study of Gixhari et al. (2014) delineated 28 pea lines into 27 groups. The information revealed in cluster analysis may be useful in designing a breeding program (Ahmad et al., 2012). In our study, the two parental lines pl-21 and pl-23 (from which the crosses of other lines were obtained)

were grouped in the same cluster. In our study, the results obtained from the constructed phylogenetic tree were the same as those of the scattered plot.

The present study was initiated to investigate the genetic diversity in 23 segregating pea lines using EST-SSR markers. It was concluded that conventional plant breeding increases genetic diversity and produces new combinations of important traits, which ultimately has a remarkable impact on agricultural yield.

Acknowledgment

The authors acknowledge the Higher Education Commission of Pakistan for provision of funds to carry out this research work under the approved project entitled "Marker Assisted Selection (MAS) Breeding in *Pisum sativum* Using Molecular Markers (RAPD and SSR Markers)". Project No: 1473. PI of the project: Dr. Mohammad Nisar, Associate Professor, University of Malakand, Pakistan.

References

Able JA, Langridge P, Milligan AS (2007). Capturing diversity in the cereals: many options but little promiscuity. Trends Plant Sci 12: 71-79.

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.

Baranger A, Aubert G, Arnau G, Lainé A, 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: 1309-1321.

Belaj A, Satovic Z, Cipriani G, Baldoni L, Testolin R, Rallo L (2003). Comparative study of the discriminating capacity of RAPD, AFLP and SSR markers and of their effectiveness in establishing genetic relationships in olive. Theor Appl Genet 107: 736-744.

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

Burstin J, Potier GDJ, Weinachter C, Aubert G, Baranger A (2001). Microsatellite polymorphism in *Pisum sativum*. Plant Breeding 120: 311-317.

Caballero A, García-Dorado A (2013). Allelic diversity and its implications for the rate of adaptation. Genetics 195: 1373-1384.

Cieslarová J, Hýbl M, Griga M, Smýkal P (2012). Molecular analysis of temporal genetic structuring in pea (*Pisum sativum* L.) cultivars bred in the Czech Republic and in former Czechoslovakia since the mid-20th century. Czech J Genet Plant 48: 61-73.

Cuevas HE, Prom LK (2013). Assessment of molecular diversity and population structure of the Ethiopian sorghum [*Sorghum bicolor* (L.) Moench] germplasm collection maintained by the USDA-ARS national plant germplasm system using SSR markers. Genet Resour Crop Ev 60: 1817-1830.

Dellaporta SL, Wood J, Hicks JB (1983). A plant DNA minipreparation: Version II. Plant Mol Biol Rep 1: 19-21.

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.

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-Sc B 11: 702-707.

Govindaraj M, Vetriventhan M, Srinivasan M (2015). Importance of genetic diversity assessment in crop plants and its recent advances: an overview of its analytical perspectives. Genetics Research International 2015: 431487.

Graham PH, Vance CP (2003). Legumes importance and constraints to greater use. Plant Physiol 131: 872-877.

Hagenblad J, Bostrom E, Nygards L, Leino M (2014). Genetic diversity in local cultivars of garden pea (*Pisum sativum* L.) conserved on farm and in historical collections. Genet Resour Crop Ev 61: 413-422.

Izzah NK, Lee J, Perumal S, Park JY, Ahn K, Fu D (2013). Microsatellite-based analysis of genetic diversity in 91 commercial *Brassica oleracea* L. cultivars belonging to six varietal groups. Genet Resour Crop Ev 60: 1967-1986.

Jain M, Moharana KC, Shankar R, Kumari R, Garg R (2014). Genome wide discovery of DNA polymorphisms in rice cultivars with contrasting drought and salinity stress response and their functional relevance. Plant Biotechnol J 12: 253-264.

Khan TN, Ramzan A, Jillani G, Mehmood T (2013). Morphological performance of peas (*Pisum sativum*) genotypes under rainfed conditions of Potowar region. J Agric Res 51: 51-60.

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

Lewis GP, Schrire B, Mackinder B, Lock M, editors (2005). Legumes of the World. Kew, UK: Royal Botanic Gardens.

Li G, 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: 45-46.

Liu K, Muse SV (2005). Power-Marker: integrated analysis environment for genetic marker data. Bioinformatics 21: 2128-2129.

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

Mahesh C, Tarte G, Ashok K, Nilesh C (2014). Studies on the effect of pH and temperature on the activity of esterase enzyme isolated from *Pisum sativum*. Asian Journal of Biomedical and Pharmaceutical Science 3: 2231-2560.

Martin-Sanz A, Caminero C, Jing R, Flavell AJ, Vega MP (2011). Genetic diversity among Spanish pea (*Pisum sativum* L.) landraces, pea cultivars and the World *Pisum* sp. core collection assessed by retrotransposon-based insertion polymorphisms (RBIPs). Span J Agric Res 9: 166-178.

McCune B, Grace JB (2005). Multivariate Analysis of Ecological Data (PC-ORD ver. 5.10 MjM Software). Gleneden Beach, OR, USA: MjM.

Missio RF, Caixeta ET, Zambolim EM, Pena GF, Ribeiro AP, Zambolim L, Pereira AA, Sakiyama NS (2009). Assessment of EST-SSR markers for genetic analysis on coffee. Bragantia 68: 573-581.

Nasiri J, Haghnazari A, Saba J (2009). SSR diversity on field peas. Afr J Biotechnol 8: 3405-3417.

Nei M (1973). Analysis of gene diversity in subdivided populations. P Natl Acad Sci USA 70: 3321-3323.

Nisar M, Ghafoor A (2010). Inheritance studies of *Pisum sativum* f1, f2 and f3 generation based on morphological traits and selection of high yielding powdery mildew resistant lines. Legume Genomics and Genetics 1: 18-23.

Nisar M, Ghafoor A, Ahmad H, Khan MR, Qureshi AS, Ali H, Aslam M (2008). Evaluation of genetic diversity of pea germplasm through phenotypic trait analysis. Pak J Bot 40: 2081-2086.

Nisar M, Ghafoor A, Khan MR, Qureshi AS (2006). Screening of *Pisum sativum* (L.) germplasm against *Erysiphe pisi* (Syd). in relation with vegetative traits and SDS-PAGE profile. Acta Biol Cracov Bot 48: 33-37.

Ramu P, Billot C, Rami JF, Senthilvel S, Upadhyaya HD, Ananda L, Hash CT (2013). Assessment of genetic diversity in the *Sorghum* reference set using EST-SSR markers. Theor Appl Genet 126: 2051-2064.

Rao NK (2004). Plant genetic resources: advancing conservation and use through biotechnology. Afr J Biotechnol 3: 136-145.

Sarıkamış G, Yanmaz R, Ermiş S, Bakır M, Yüksel C (2010). Genetic characterization of pea (*Pisum sativum*) germplasm from Turkey using morphological and SSR markers. Genet Mol Res 9: 591-600.

Shete S, Tiwari H, Elston RC (2000). On estimating the heterozygosity and polymorphism information content value. Theor Popul Biol 57: 265-271.

Simko I (2009). Development of EST-SSR markers for the study of population structure in lettuce (*Lactuca sativa* L.). J Hered 100: 256-262.

Smykal P, Aubert G, Burstin J, Coyne CJ, Ellis NT, Flavell AJ, Ford R, Hýbl M, Macas J, Neumann P (2012). Pea (*Pisum sativum* L.) in the genomic era. Agronomy 2: 74-115.

Taran B, Zhang C, Warkentin T, Tullu A, Vandenberg A (2005). Genetic diversity among varieties and wild species accessions of pea (*Pisum sativum* L.) based on molecular markers, and morphological and physiological characters. Genome 48: 257-272.

Tatikonda L, Wani SP, Kannan S, Beerelli N, Sreedevi TK, Hoisington DA, Devi D, Varshney RK (2009). AFLP-based molecular characterization of an elite germplasm collection of *Jatropha curcas* L., a biofuel plant. Plant Sci 176: 505-513.

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

Varshney RK, Graner A, Sorrells ME (2005). Genic microsatellite markers in plants: features and applications. Trends Biotechnol 23: 48-55.

Wouw MVD, Hintum TV, Kik C, Treuren RV, Visser B (2010). Genetic diversity trends in twentieth century crop cultivars: a meta-analysis. Theor Appl Genet 120: 1241-1252.

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.

Zhuang X, McPhee KE, Coram TE, Peever TL, Chilvers MI (2013). Development and characterization of 37 novel EST-SSR markers in *Pisum sativum* (Fabaceae). Applications in Plant Sciences 1: 1-4.

Zietkiewiez E, Rafalski A, Labuda D (1994). Genome fingerprinting by simple sequence repeat (SSR)-anchored polymerase chain reaction amplification. Genomics 20: 176-183.