Xi'an Shiyou Daxue Xuebao (Ziran Kexue Ban)/
Journal of Xi'an Shiyou University, Natural Sciences Edition
ISSN: 1673-064X
E-Publication: Online Open Access
Vol: 67 Issue 08 | 2024
DOI: 10.5281/zenodo.13358030

MOLECULAR PROFILING OF SOYBEAN LINES USING SSR GENE-SPECIFIC MARKERS

SUMMRA SIDDIQUE *
Department of Plant Breeding and Genetics, University of Agriculture Faisalabad Punjab, Pakistan.
*Corresponding Author Email: samrasiddique33@yahoo.com

ZAHEER AHMED
Department of Plant Breeding and Genetics, University of Agriculture Faisalabad Punjab, Pakistan.

RIZWANA MAQBOOL
Department of Plant Breeding and Genetics, University of Agriculture Faisalabad Punjab, Pakistan.

FAISAL SAEED AWAN
Center of Agricultural Biochemistry and Biotechnology, University of Agriculture Faisalabad Punjab, Pakistan.

Abstract

Soybean is a high-nutritional crop that provides a sustainable source of dietary protein and edible oil for human consumption and animal diet. In this investigation, 9 potential soybean genotypes and four commercial checks on the basis of yield and quality characters were evaluated for molecular diversity through simple sequence repeat (SSR) markers. Out of 13 gene specific SSR-loci, all markers showed association with respective trait. The heterozygosity ranged from 0.41 to 0.83, lowest to highest was observed at locus Gm012 and Gm004 respectively. Maximum gene diversity was found at locus Gm011 (0.66), while minimal was observed for locus Gm012 (0.57) with the mean value of 0.64 per marker. The polymorphic information content (PIC) was recorded highest for Gm011 (0.59) followed by Gm003, Gm006, Gm009, Satt001, while the lowest value for Gm012 (0.51). Among the 13 SSR markers, Gm011 was found to be the most informative, while Gm012 was the least insightful for the genetic diversity of studied germplasm. The polymorphic information contents (PIC) indicating a moderate level of genetic diversity exhibited by the assessed germplasm. In addition, principal coordinate analysis delineated two groups, representing two diverse gene pools in soybean collection. The UPGMA inferred three clusters based on their genetic resemblance and maturity groups. Thus, the use of substantial genetic diversity confirmed in the present study will enable soybean breeders to select diverse parents to develop high-yielding varieties with improved quality traits.

Keywords: DNA, Gel Electrophoresis, Polymerase Chain Reaction, Primer Designing, Polymorphic Information Content**,** Soybean, SSR Markers.

1. INTRODUCTION

Soybean (*Glycine max* (L.) Merr.) Is one of the world's most important economic legume crop. Improved grain yield with increased oil and protein contents are key attributes for the selection of soybean genotypes [1]. Although Pakistan is an agricultural country, it cannot produce enough soybean due to limited genetic resources; hence, desired agro-morphological and nutritional characteristics are insufficient in the available germplasm. Thus, the development of adaptable soybean genotypes with higher grain yield and improved seed quality characters remains exceptionally requisite.

Xi'an Shiyou Daxue Xuebao (Ziran Kexue Ban)/
Journal of Xi'an Shiyou University, Natural Sciences Edition
ISSN: 1673-064X
E-Publication: Online Open Access
Vol: 67 Issue 08 | 2024
DOI: 10.5281/zenodo.13358030

The selection of divergent parents is a major challenge due to low diversity and unknown pedigrees. Therefore, the existence and extent of genetic diversity in crop improvement is a prerequisite for increased productivity [2]. Generally, the genetic diversity of soybean was estimated on the basis of agronomic traits such as grain yield and seed size [3], biochemical quality attributes such as oil and protein content [4] and various types of molecular markers including RAPD, AFLP and SSR polymorphism [5]. More importantly, the genetic markers have a high degree of precision and accuracy in screening of genetically diverse parents and marker assisted breeding for the improvement of trait of interest [6]. Among these molecular markers, SSRs have proven to be a more powerful tool for the determination of genetic diversity because of their distinctive features like high polymorphism, abundance in the genome, codominant inheritance, highly reproducible and stable to environmental fluctuations [7]. In soybean, microsatellite (SSR) markers are more convenient in defining genetically diverse breeding material and to broaden the genetic background of available germplasm through conventional breeding procedures [8], [9]. Integration of advanced molecular tools with traditional breeding techniques is the prerequisite for crop improvement [10]. Therefore, the assessment of genetic diversity in soybean is an efficient approach in combination with morphological and molecular markers for the selection of superior genotypes [11].

In Pakistan, very few studies have been concentrated on the genetic diversity of oilseed crops including soybean. Most of the soybean germplasm were characterized based on agro-morphological characters [12]. In order to minimize the overall import costs of soybean and edible oils, the present study aimed to determine the genetic diversity and identify the diverse genotypes that might be used as potential parents in future breeding programs for the development of higher grain yield and improved quality varieties.

Understanding the genetic diversity of soybean germplasm is essential to broaden the genetic base and to further utilize it in breeding program. Such an insight could be achieved through molecular characterization of soybean germplasm using DNA markers, which are more informative, stable and reliable, as compared to conventional methods like pedigree analysis and morphological diversity. Early studies have shown utilization of molecular markers for identification of genetically diverse genotypes to use in crosses in breeding programme [13], [14].

Among different types of DNA markers being utilized for molecular characterization and genetic diversity analysis in plants, simple sequence repeats (SSR) markers are considered as molecular marker of choice due to their abundance, high polymorphism rate and high reproducibility. SSR markers have been widely used in the genetic diversity studies of the soybean germplasm collections worldwide and high levels of polymorphism at SSR loci have been reported for both the number of alleles per locus and the gene diversity [15, [16], [17], [18], [19], [20], [21]. Therefore in the present study a set of nine soybean accessions used as parents and four check cultivars were selected based on agro-morphological and nutritional characterization from the whole germplasm collection for molecular characterization and genetic diversity analysis using SSR markers.

Xi'an Shiyou Daxue Xuebao (Ziran Kexue Ban)/
Journal of Xi'an Shiyou University, Natural Sciences Edition
ISSN: 1673-064X
E-Publication: Online Open Access
Vol: 67 Issue 08 | 2024
DOI: 10.5281/zenodo.13358030

2. MATERIALS AND METHODS

Growth of Seedlings

Selected nine parents and four check cultivars were sown in the glasshouse under controlled conditions. Trays filled with compost were used and two seeds were grown in each hole of the tray. After 10-12 days, all seedlings were grown and fresh leaf samples were collected in an ice bucket for DNA extraction and kept at -80^{o}C to store. These samples were then freeze dried so that they can be keep in room temperature too.

Figure 1: Seedlings of Soybean Accessions

DNA Extraction

Following QIAGEN instructions were followed for DNA extraction:

1. Freeze dried leaf samples were disrupted (≤100 mg wet weight or ≤20 mg lyophilized tissue) using the homogenizer.
2. 400 µl Buffer AP1 and 4 µl RNase A were added. Samples were vortex and incubated for 10 min at 65°C. Tubes were inverted 2–3 times during incubation.
3. 130 µl Buffer P3 was added and incubation was done for 5 min on ice.
4. Lysate was centrifuge for 5 min at 20,000 x g (14,000 rpm).
5. Pipet the lysate into a QIAshredder spin column placed in a 2 ml collection tube. Centrifuge for 2 min at 20,000 x g.

Xi'an Shiyou Daxue Xuebao (Ziran Kexue Ban)/
Journal of Xi'an Shiyou University, Natural Sciences Edition
ISSN: 1673-064X
E-Publication: Online Open Access
Vol: 67 Issue 08 | 2024
DOI: 10.5281/zenodo.13358030

6. Transfer the flow-through into a new tube without disturbing the pellet if present. Add 1.5 volumes of Buffer AW1, and mix by pipetting.
7. Transfer 650 µl of the mixture into a DNase Mini spin column placed in a 2 ml collection tube. Centrifuge for 1 min at ≥6000 x g (≥8000 rpm). Discard the flow through. Repeat this step with the remaining sample.
8. Place the spin column into a new 2 ml collection tube. Add 500 µl Buffer AW2, and centrifuge for 1 min at ≥6000 x g. Discard the flow-through.
9. Add another 500 µl Buffer AW2. Centrifuge for 2 min at 20,000 x g. Note: Remove the spin column from the collection tube carefully so that the column does not come into contact with the flow-through. 10. Transfer the spin column to a new 1.5 ml or 2 ml micro centrifuge tube.
11. Add 100 µl Buffer AE for elution. Incubate for 5 min at room temperature (15–25°C). Centrifuge for 1 min at ≥6000 x g.
12. Repeat step 11

DNA Quantification

The concentration of extracted DNA was quantified through Nanodrop spectrophotometer (Model: ND1000) using a wavelength of A260/A280 nm with an optimum range of 1.8 to 2.0 for extracted DNA purity. First blank (d_3H_2O) was run and equipment was adjusted to auto-zero for blank. The samples from each parent were loaded on the spectrophotometer to measure their DNA concentration.

DNA Quality

Good quality DNA of known concentration was needed for subsequent use in PCR. To check the quality of the extracted DNA, 5 µL DNA of each parent was run on 1% agarose gel stained with ethidium bromide and visualized under UV light. The gel picture was analyzed in gel documentation system for measurement of DNA concentrations. The brighter bands having higher densities possessed high DNA concentrations and vice versa.

Primer Designing

Different yield and quality controlling genes of soybean were studied from previous literature and their sequence was retrieved from Soybase database (https://www.soybase.org) 12 SSR primers were designed using Primer 3 software. One reported primer (soysat001) was used in the study.

Xi'an Shiyou Daxue Xuebao (Ziran Kexue Ban)/
Journal of Xi'an Shiyou University, Natural Sciences Edition
ISSN: 1673-064X
E-Publication: Online Open Access
Vol: 67 Issue 08 | 2024
DOI: 10.5281/zenodo.13358030

Table 1: List of SSR Primers, their sequence and annealing temperature used in this study

Sr. No	Gene ID	Forward primer	Reverse primer	Annealing temprature
1.	Glyma.04G102900	CTGATGCGCAGTGCTACATT	TGGGTGATTGGTCAACAAGA	56
2.	Glyma.11G229600	AAGTTGAACTGCCGTTTTGG	AACAGAAACCCAATGGTTCG	56
3.	Glyma.04G116500	GAAAATTCTGGGGCAGATGA	CCAGCCTCAAGGAGAACTTG	58.5
4.	Glyma.14G194300	GCAGAAACTGGATGCAATGA	TAGAGGGCATGAGGAAATGG	57
5.	Glyma.16G003500	TTGGCATCACACTCTTGCTC	TCCAAGACAAAGCAGCAATG	56
6.	Glyma.04G077600	AACCATGGCGAACTTGAAAC	GGTTTGCGGTGAATTGAGAT	53
7.	Glyma.05G004300	TGCGTTACATGGTTGCAAAT	TAAAATCCGGCAAAACTTGG	56
8.	Glyma.05G004300	CCACAAAACTTCCTCCCTGA	CGCTGTCATCTTTCACTCCA	55
9.	Glyma.05G013800	GGAGATCCAGTGTGCCATTT	TCTTTGGATGGCAAACTTCC	55
10.	Glyma.10G275800	AAGCTTGGTGCTCTTTGCAT	TGGGTTGTTGTTCATCTCCA	57
11.	Glyma.14G121400	CCACAAAACTTCCTCCCTGA	CGCTGTCATCTTTCACTCCA	57
12.	Glyma.02G005200	GGAGATCCAGTGTGCCATTT	TCTTTGGATGGCAAACTTCC	52
13.	Soysat001	TGTGCAATGATAGTACATAGA	GTGCTGATTGAACTATTTGTA	57

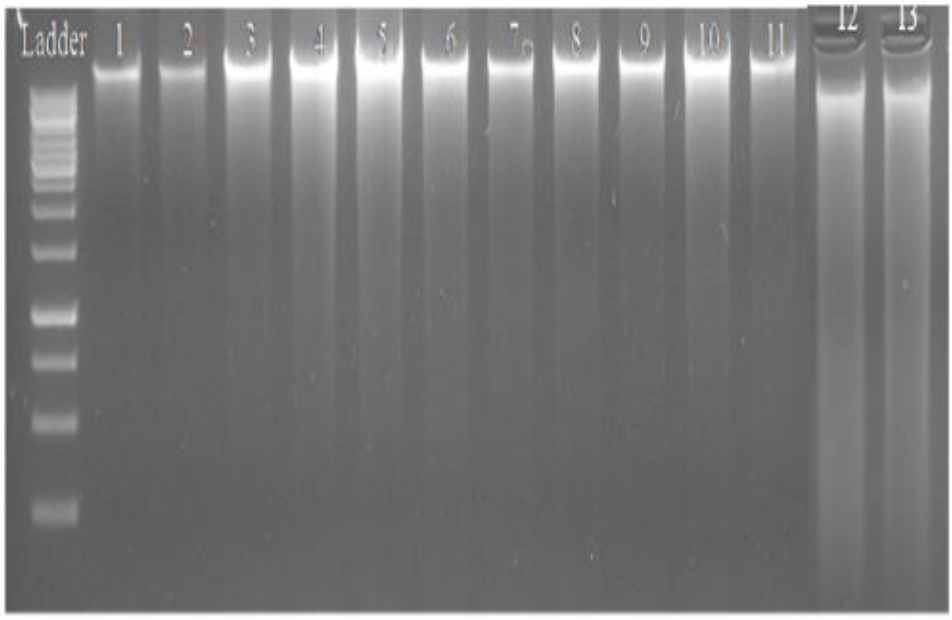


Figure 2: Smears of extracted DNA soybean samples

SSR Primers Stock Solution (First Dilution)

- The exact amount of each oligo was supplied by the manufacturer.
- The “nmol” amount was multiplied by 10 and SSR pellet was dissolved in that much volume of deionized water.
- Example: 26.10nmols x 10 = 261µl
- The oligo pellet was dissolved in 261µl to get 100 µM stock solution

Xi'an Shiyou Daxue Xuebao (Ziran Kexue Ban)/
Journal of Xi'an Shiyou University, Natural Sciences Edition
ISSN: 1673-064X
E-Publication: Online Open Access
Vol: 67 Issue 08 | 2024
DOI: 10.5281/zenodo.13358030

Second Dilution

- 10µl of first solution + 190µl deionized water to yield 50µM solution.

Working Dilution

10µl of second dilution + 90ul deionized water to yield 5µM solution

Polymerase Chain Reaction

Once high quality DNA of known concentration was in hand, a series of different PCRs were performed for testing of specific genes in the samples. PCR was performed with selected primers at their optimized annealing temperature to explore genetic diversity among nine parents and four check cultivars. Working solution of samples containing 15ng of DNA in one µL of d_3H_2O was prepared.

For PCR the procedure was:

- 6 µl DreamTaq Green PCR master mix was added in 1.5 ml eppendorf tubes.
- 0.5 µl of diluted forward and reverse primers were added.
- 2-3 µl DNA samples were added in each eppendorf tube.
- 10-11 µl deionized water was added in each tube to make 20 µl total volume.

PCR tubes were placed in the thermal cycler and PCR was performed with the PCR profile given in table 2.

Table 2: Thermocycling Conditions of PCR Reaction used for Estimation of genetic Diversity

Steps	Temperature	Time	Cycles
Initial denaturation	94°C	1 min	1
Denaturation	94°C	30 sec	35-40
Annealing	55-65°C	30-45 sec	
Extension	72°C	1 min	
Final extension	72°C	10 min	
Holding	4°C	∞	until turned off

Agarose Gel Electrophoresis and Gel Documentation System

Gel electrophoresis is a tool used for the separation and purification of macromolecules especially nucleic acids (DNA, RNA, and Protein) on the basis of their charge as well as size. The larger molecules covers less distance than the smaller molecules because small molecules can move rapidly through the pores of the gel and hence go farther than the larger molecule. Nucleic acid molecules have negative charge on them and hence move from negative electrode to the positive one.

2% agarose gel was prepared dissolving 2g agarose in 100 mL of 1X TAE buffer and heated in microwave oven for 2-3 minutes. After heating, the liquid gel was cooled and add 2.0 µl of ethidium bromide (EtBr) solution. In the meantime, gel casting tray was

Xi'an Shiyou Daxue Xuebao (Ziran Kexue Ban)/
Journal of Xi'an Shiyou University, Natural Sciences Edition
ISSN: 1673-064X
E-Publication: Online Open Access
Vol: 67 Issue 08 | 2024
DOI: 10.5281/zenodo.13358030

prepared by placing stoppers at both ends. The gel solution was gently poured in gel tray to avoid bubble formation, comb was inserted at required places and kept at room temperature for gel solidification. After about 30 minutes, the gel tray along with gel was put into the horizontal gel container of the agarose gel apparatus containing 1XTAE buffer. The wells in gel formed by the comb were kept toward cathode of the apparatus. To separate the DNA fragments according to their size, 10 µl of PCR product was loaded in the wells. The gel was run supplying 60 volts of current for 50 minutes. After running the gel, it was viewed in UV gel documentation system.

SSR allele Scoring and Data Analysis

The presence or absence of SSR fragment in each accession was recorded for all the gene specific SSR markers. The SSR bands appearing without ambiguity were scored as 1 (present) and 0 (absent) for each primer. The polymorphism information content (PIC), a measure of the allelic diversity at a locus, was determined by using the formula:

$$PIC = 1-\Sigma Pi^2$$

Where, Pi is the frequency of the ith allele in the set of genotypes analyzed, calculated for each SSR locus. The genetic similarity among genotypes was estimated based on Jaccard's similarity coefficient. The resulting similarity matrix was further analyzed using the unweighted pair-group method arithmetic average (UPGMA) clustering algorithm for construction of dendrogram; the computations were carried out using powermarker version 3.25 and PAST 3 software.

3. RESULTS

Molecular Characterization of Soybean Accessions

Improvement of yield, oil and protein content in soybean is attributed to increased use of genetically diverse parents in breeding programme. Knowledge on the genetic diversity of germplasm helps in selection of parental genotypes for development of segregation population and to develop varieties. Morphological traits based genetic diversity is prone to environmental variations and availability of limited number of morphological markers has limited their use in genetic diversity studies. On the other hand, molecular markers based genetic diversity is not influenced by environmental factors therefore highly reproducible and also widely distributed throughout the genome. To efficiently broaden the genetic base of modern soybean cultivars, an insight into molecular diversity is necessary. In the present study, an attempt has been made to study the molecular diversity of selected parents for breeding program and four local cultivars using gene specific 13 SSR markers. More than 0.5 value of polymorphic information content (PIC) observed in SSR markers among the soybean accessions in the present study demonstrated the effectiveness of SSR markers in determining genetic variation.

A total of thirteen microsatellite (SSR) markers were utilized to characterize and assess genetic diversity for yield and quality related traits among selected parents and check cultivars. These accessions were selected from Principal component analysis based on

Xi'an Shiyou Daxue Xuebao (Ziran Kexue Ban)/
Journal of Xi'an Shiyou University, Natural Sciences Edition
ISSN: 1673-064X
E-Publication: Online Open Access
Vol: 67 Issue 08 | 2024
DOI: 10.5281/zenodo.13358030

yield and quality contributing traits for both spring and autumn season. All SSR markers showed amplification except one SSR marker which showed no amplified products. These were used in the analysis of genetic diversity. Major allele frequency, genetic diversity, heterozygosity and polymorphic information content (PIC) has been given in Table 5.3. Major allelic frequencies were calculated for each locus, and major allele frequency ranged from 0.33 to 0.58 with an average of 0.42 per locus. The heterozygosity ranged from 0.41 to 0.83, lowest to highest was observed at locus Gm012 and Gm004 respectively. Maximum gene diversity was found at locus Gm011 (0.66), while minimal was observed for locus Gm012 (0.57) with the mean value of 0.64 per marker. The polymorphic information content (PIC) was recorded highest for Gm011 (0.59) followed by Gm003, Gm006, Gm009, Satt001, while the lowest value for Gm012 (0.51). Among the 13 SSR markers, Gm011 was found to be the most informative, while Gm012 was the least insightful for the genetic diversity of studied germplasm.

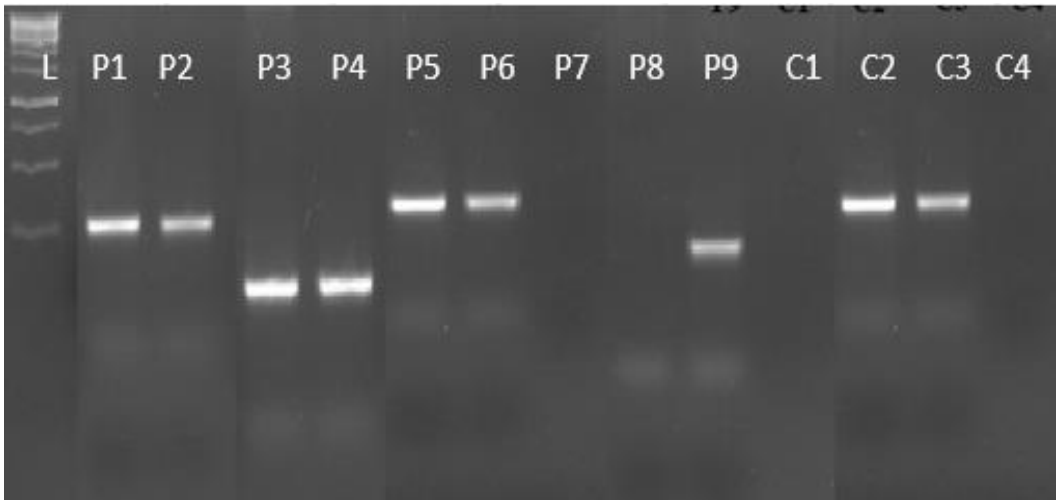
These SSR markers are linked with different agronomic and quality traits of soybean. Amplified product of potential parents and check cultivars of some targeting genes is shown in Figs 3-8. Gm11 had high PIC value (0.59) which indicates that this SSR marker had high diversity in soybean accessions. The lower PIC values indicates low allelic diversity in present set of soybean accessions.

To assess the genetic resemblances among the selected parents and local soybean accessions, Jaccard's similarity coefficients were calculated. The pairwise genetic similarity among selected soybean accessions used as parents and local cultivars varied from 0.28 to 0.90. The similarity coefficients matrix was used for UPGMA cluster analysis. The dendrogram constructed based on genetic similarities between genotypes showed that these genotypes formed three major clusters (Fig 2). The genetic structure of the 13 soybean accessions based on the dendrogram does not correlate with country of acquisition or morphological differentiation based on maturity duration.

Table 3: Major allele frequency, Gene diversity, Heterozygosity and Polymorphic information content (PIC) of 13 SSR loci used in the study

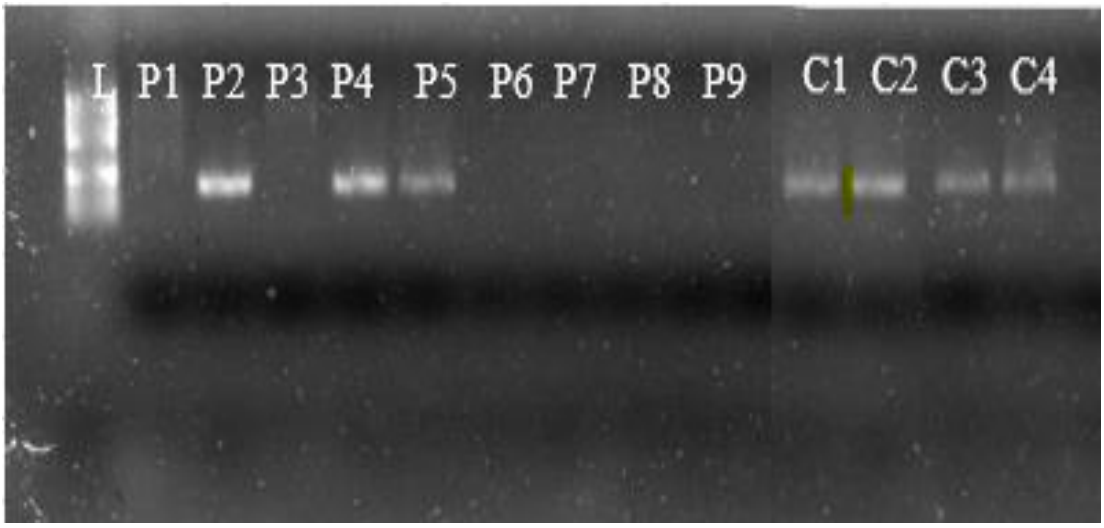
SSR markers	Major Allele Frequency	Gene Diversity	Heterozygosity	PIC
Gm001	0.50	0.6250	0.5000	0.5547
Gm002	0.42	0.6563	0.5833	0.5827
Gm003	0.38	0.6563	0.7500	0.5815
Gm004	0.42	0.6250	0.8333	0.5454
Gm005	0.41	0.6563	0.5833	0.5827
Gm006	0.37	0.6563	0.7500	0.5815
Gm007	0.50	0.6250	0.5000	0.5547
Gm008	0.41	0.6563	0.5833	0.5827
Gm009	0.37	0.6563	0.7500	0.5815
Gm010	0.41	0.6563	0.5833	0.5827
Gm011	0.33	0.6667	0.6667	0.5926
Gm012	0.58	0.5729	0.4167	0.5101
Satt001	0.37	0.6563	0.7500	0.5815
Mean	0.42	0.6434	0.6346	0.5703

Xi'an Shiyou Daxue Xuebao (Ziran Kexue Ban)/
Journal of Xi'an Shiyou University, Natural Sciences Edition
ISSN: 1673-064X
E-Publication: Online Open Access
Vol: 67 Issue 08 | 2024
DOI: 10.5281/zenodo.13358030



L= 50bp Ladder, **P1=**LN86-1595, **P2=**L2-3, **P3=**CX210**, P4=**AP 1776**, P5=**Vernal**, P6=** L-4303**, P7=**GUTWEIN-180, **P8=** 9162, **P9=**Toano, **C1=** Malakand, **C2=**Ajmeri, **C3=**Rawal, **C4=** NARC-2

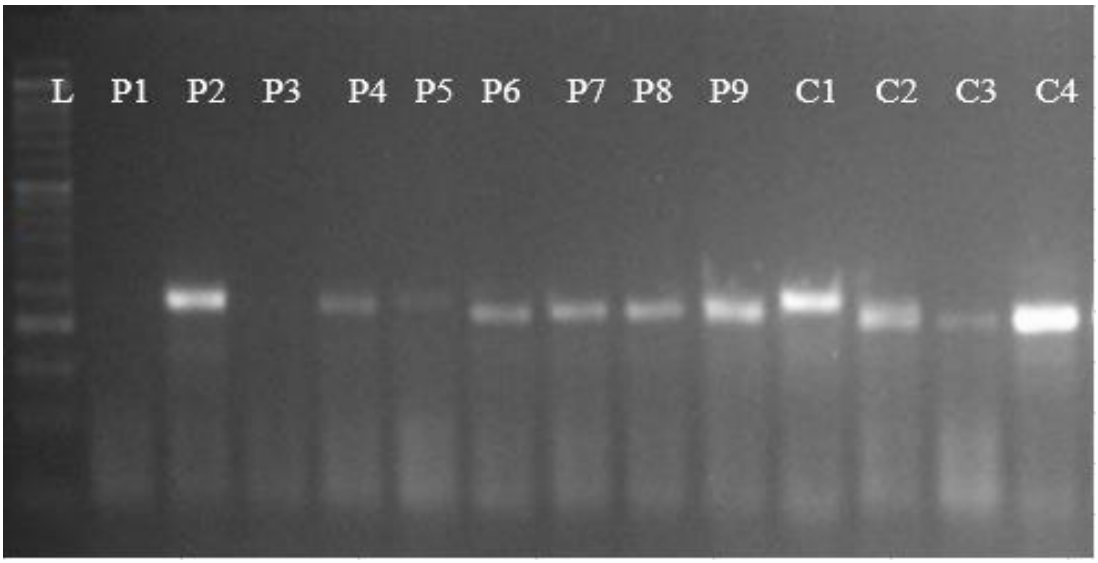
Figure 3: PCR analysis of potential parents and check cultivars by targeting Glyma.10G275800 gene



L= 50bp Ladder, **P1=**LN86-1595, **P2=**L2-3, **P3=**CX210**, P4=**AP 1776**, P5=**Vernal**, P6=** L-4303**, P7=**GUTWEIN-180, **P8=** 9162, **P9=**Toano, **C1=** Malakand, **C2=**Ajmeri, **C3=**Rawal, **C4=** NARC-2

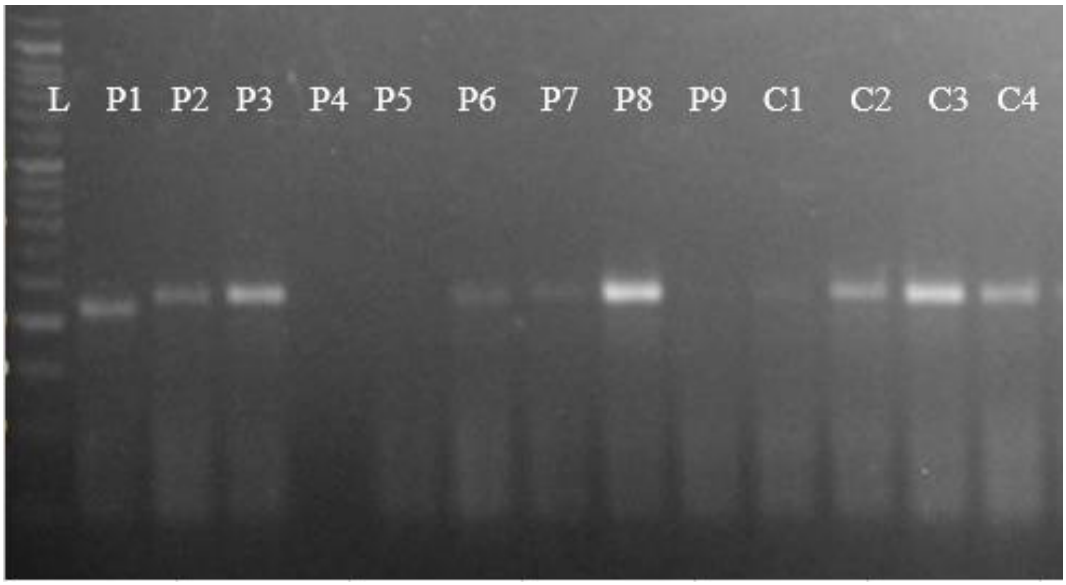
Figure 4: PCR analysis of potential parents and check cultivars by targeting Glyma.05G004300 gene

Xi'an Shiyou Daxue Xuebao (Ziran Kexue Ban)/
Journal of Xi'an Shiyou University, Natural Sciences Edition
ISSN: 1673-064X
E-Publication: Online Open Access
Vol: 67 Issue 08 | 2024
DOI: 10.5281/zenodo.13358030



L= 50bp Ladder, **P1=**LN86-1595, **P2=**L2-3, **P3=**CX210**, P4=**AP 1776**, P5=**Vernal**, P6=** L-4303**, P7=**GUTWEIN-180, **P8=** 9162, **P9=**Toano, **C1=** Malakand, **C2=**Ajmeri, **C3=**Rawal, **C4=** NARC-2

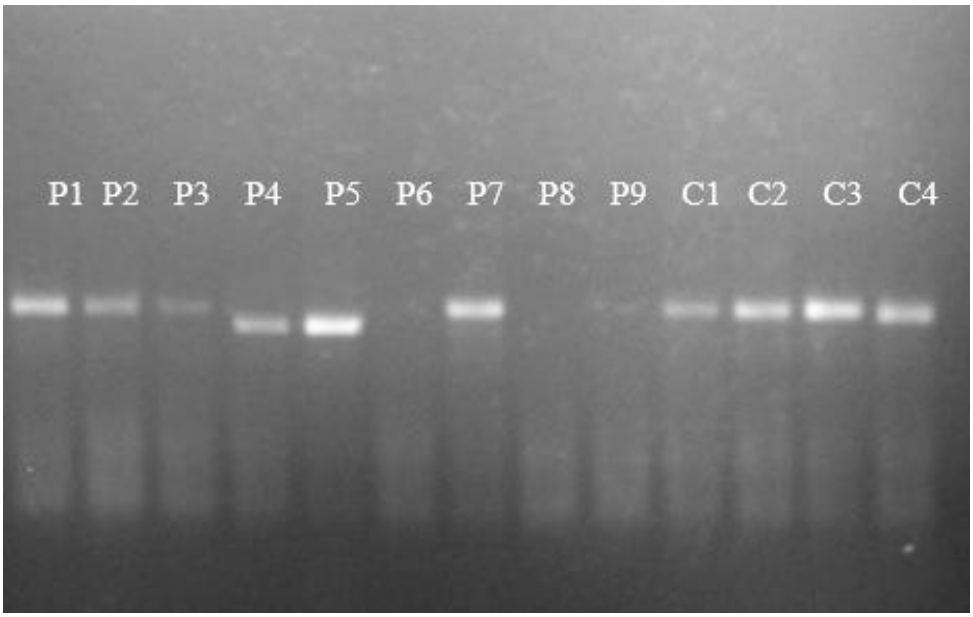
Figure 5: PCR analysis of potential parents and check cultivars through Soy sat001 marker



L= 50bp Ladder, **P1=**LN86-1595, **P2=**L2-3, **P3=**CX210**, P4=**AP 1776**, P5=**Vernal**, P6=** L-4303**, P7=**GUTWEIN-180, **P8=** 9162, **P9=**Toano, **C1=** Malakand, **C2=**Ajmeri, **C3=**Rawal, **C4=** NARC-2

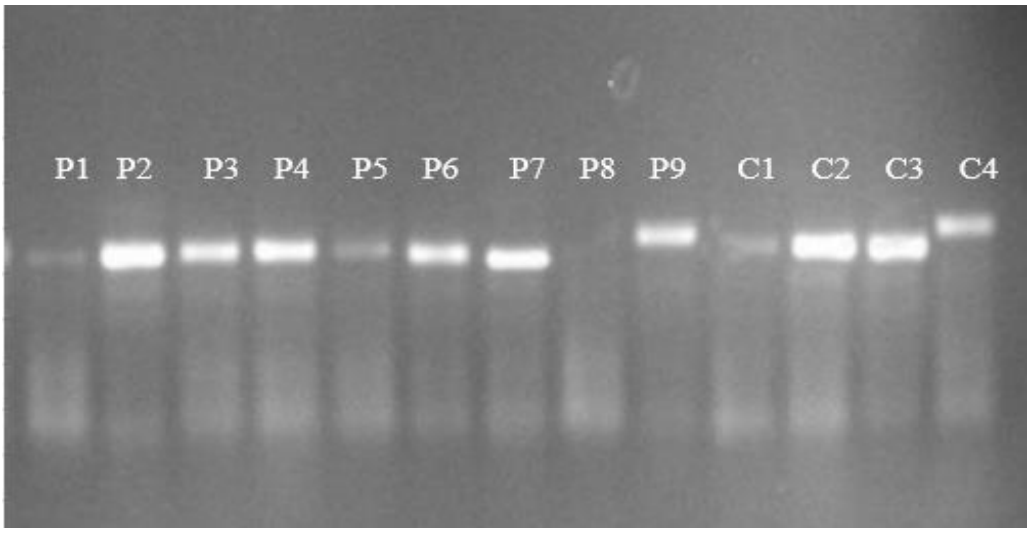
Figure 6: PCR analysis of potential parents and check cultivars by targeting GLYMA.04G102900 gene

Xi'an Shiyou Daxue Xuebao (Ziran Kexue Ban)/
Journal of Xi'an Shiyou University, Natural Sciences Edition
ISSN: 1673-064X
E-Publication: Online Open Access
Vol: 67 Issue 08 | 2024
DOI: 10.5281/zenodo.13358030



P1=LN86-1595, **P2=**L2-3, **P3=**CX210**, P4=**AP 1776**, P5=**Vernal**, P6=** L-4303**, P7=**GUTWEIN-180, **P8=** 9162, **P9=**Toano, **C1=** Malakand, **C2=**Ajmeri, **C3=**Rawal, **C4=** NARC-2

Figure 7: PCR analysis of potential parents and check cultivars by targeting GLYMA.14G194300 gene



P1=LN86-1595, **P2=**L2-3, **P3=**CX210**, P4=**AP 1776**, P5=**Vernal**, P6=** L-4303**, P7=**GUTWEIN-180, **P8=** 9162, **P9=**Toano, **C1=** Malakand, **C2=**Ajmeri, **C3=**Rawal, **C4=** NARC-2

Figure 8: PCR analysis of potential parents and check cultivars by targeting Glyma.02G005200 gene

Xi'an Shiyou Daxue Xuebao (Ziran Kexue Ban)/
Journal of Xi'an Shiyou University, Natural Sciences Edition
ISSN: 1673-064X
E-Publication: Online Open Access
Vol: 67 Issue 08 | 2024
DOI: 10.5281/zenodo.13358030

Cluster Analysis

To assess the genetic resemblances among the genotypes, Jaccard's similarity coefficients were calculated for all SSR markers among 13 soybean accessions. The pairwise genetic similarity among 13 soybean accessions varied from 0.40 to 0.99. The similarity coefficients matrix was used for UPGMA cluster analysis. The dendrogram constructed based on genetic similarities between genotypes showed that the 13 genotypes formed three major clusters (Fig. 2). The genetic structure of the 13 soybean accessions based on the dendrogram does not correlate with origin or morphological differentiation based on maturity groups. The Cluster I contains three genotypes two from USA and one Pakistani variety, namely Gutwein-180 belongs to MG-I, Malakand (Pakistani variety) and LN86-1595 belongs to MG-II. The Cluster II contains eight genotypes six from USA and two Pakistani varieties, namely L2-3 belongs to MG-II, CX210 belongs to MG-II, Vernal belongs to MG-V, NARC 2 (Pakistani variety), AP 1776 belongs to MG-I, Toano belongs to MG-V, Ajmeri (Pakistani variety) and 9162 belongs to MG-I. The cluster III contain only two genotypes namely L4303 from USA and belongs to MG-II and Rawal (Pakistani variety). Hence, closed genetic similarity was seen between the USA and Pakistani accessions. In cluster I, Gutwein-180 and Malakand has 67.4% genetic similarity, while in cluster II, L2-3 and CX210 has 82.5% genetic similarity. L4303 and Rawal in cluster III has almost 68.5% genetic similarity.

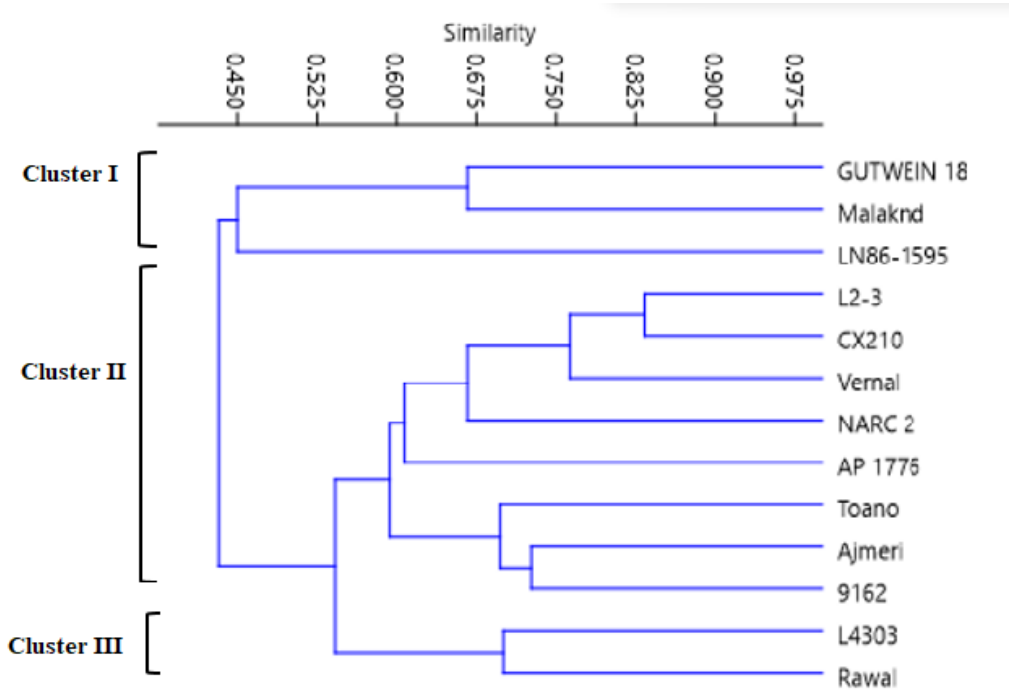


Figure 9: Dendrogram showing genetic relationships among 13 soybean accessions based on UPGMA clustering of Jaccard's similarity coefficients

Xi'an Shiyou Daxue Xuebao (Ziran Kexue Ban)/
Journal of Xi'an Shiyou University, Natural Sciences Edition
ISSN: 1673-064X
E-Publication: Online Open Access
Vol: 67 Issue 08 | 2024
DOI: 10.5281/zenodo.13358030

Principal Coordinate analysis (PCoA)

The association of the nine selected parents and four check cultivars was examined using genetic distance matrix data of 13 SSR markers. The Principle Coordinate Analysis (PCoA) was employed to gain insight into the germplasm and genetic relationship among genotypes. The PCoA separated the germplasm into two groups with a clear distinction, which identified in the population structure. The first two principal coordinates contributed 48.48 % of the cumulative variability, with an individual contribution of 27.49% and 20.99%, respectively (Fig 5.2). LN86-1595, Gutwein-180, AP 1776, NARC 2, Malakand and Ajmeri showed closed relationship and categorized in one group while 9162, Rawal, Toano, L4303, CX210, Vernal and L2-3 showed more similar genetic makeup and placed in another group.

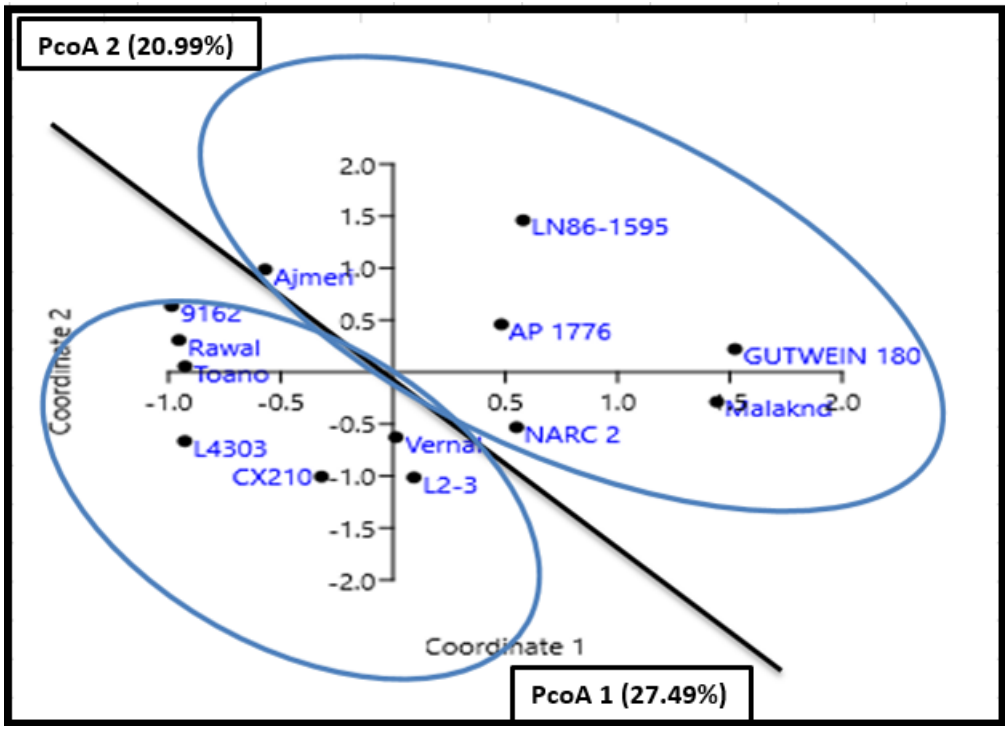


Figure 10: Principal Coordinate Analysis (PCoA) presenting the genetic diversity of Soybean Genotypes at Different Coordinates

4. DISCUSSION

The characterization of soybean germplasm is a prerequisite for the conservation, management, and utilization of available genetic resources efficiently for current and future breeding programs [22] (Kumar *et al*., 2015). The knowledge of genetic diversity would help the plant breeders to broaden the genetic base of locally adapted cultivars. Therefore, genetic diversity estimation has become a vital approach for identifying genetically divergent parents with desirable traits [23] (Oliveira *et al.*, 2017).

Xi'an Shiyou Daxue Xuebao (Ziran Kexue Ban)/
Journal of Xi'an Shiyou University, Natural Sciences Edition
ISSN: 1673-064X
E-Publication: Online Open Access
Vol: 67 Issue 08 | 2024
DOI: 10.5281/zenodo.13358030

The selection of potential parents based on morphological traits was supported by molecular data in the determination of genetic diversity and genetic structure of assessed soybean germplasm would facilitate the long-term breeding and selection programs. Thus, the combination of morphological and molecular markers is a useful strategy for the selection of diverse progenitors [24] (Tripathi and Khare, 2016).

Hyten *et al*. (2006) [25] concluded that, during soybean domestication, 50% of the genetic diversity and 81% of the rare alleles were lost, and that there were changes in 60% of allelic frequencies. Moreover, the introduction of few accessions in producing countries might have caused losses of approximately 79% of the rare alleles previously found in domestic populations of soybean. A large significant number of rare alleles may contribute to soybean breeding, since they are absent from elite cultivars.

Genetic diversity has played a key role in the collection, conservation, and exploitation of important genetic resources for the genetic improvement of soybean [26],[27](Rao and Hodgkin, 2002; Nawaz *et al*., 2020). In soybean crossbreeding, the diverse parents is more appropriate in order to minimize the genetic instability of segregating populations [28](Mihaljevic *et al.,* 2020). Many researchers have recently assessed the genetic diversity in various legume crops, including chickpea [29](Samyuktha *et al.,* 2018), cluster bean [30](Tribhuvan *et al*., 2019) and soybean [9](Hwang *et al.,* 2020). The molecular analysis revealed a high polymorphism (94%), which is maybe likely due to the polymorphic nature of SSRs was previously confirmed by Dong *et al.*, (2014)[31]; Hipparagi *et al*., (2017) [8]and Tasma *et al.*, (2018)[32]. The current results are also in agreement with Kachare *et al.,* (2020) [33], who assessed the genetic diversity among the soybean genotypes and found high polymorphism for tested SSR markers. We observed considerably high allelic variation for studied germplasm compared to the earlier study of Bisen *et al*. (2015) [34], who estimated the genetic diversity among the active soybean collection. In contrast, Gupta and Manjaya (2017) [35] reported much higher allelic richness for Indian cultivars than our study. Thus, the molecular markers system is a reliable approach in the identification of diverse soybean populations [32](Tasma *et al.,* 2018).

The polymorphic information content (PIC) predicts the allelic frequencies and signifies the informativeness of markers (Dong *et al.* 2014) [31]. It is noteworthy that PIC value is not always constant due to the variations in genetic architecture of evaluated germplasm (Hwang *et al*., 2020) [9]. Therefore, marker discrimination power primarily depends on the average PIC value [36] (Fu *et al.,* 2016). In the current study, all SSR markers exhibited high PIC values (>0.5), which may be attributed to the higher allelic variation retained by the germplasm.

The UPGMA clustering categorized the 59 soybean genotype into five distinct clusters, indicating the genotypes are more similar in the same cluster and are different from the other clusters (Kachare *et al.*, 2020)[33]. Appiah-Kubi *et al*. (2014) [37] observed closed relationship between genotypes and their geographical origin while assessing genetic diversity among the soybean genotypes by using 20 SSR markers.

Xi'an Shiyou Daxue Xuebao (Ziran Kexue Ban)/
Journal of Xi'an Shiyou University, Natural Sciences Edition
ISSN: 1673-064X
E-Publication: Online Open Access
Vol: 67 Issue 08 | 2024
DOI: 10.5281/zenodo.13358030

5. CONCLUSION

The genetic diversity of the investigated accessions is high, distributed over three groups and several subgroups, and exhibits a moderate level of association between genetic divergence and different maturity groups. Genetic diversity of soybean is effectively investigated using microsatellites markers, which allow a more complete coverage of the existent genetic variation.

The present study highlighted the genetic behavior of soybean genotypes to improve seed quality without offsetting grain yield. Significant parallelism was observed between phenotypic diversity and molecular diversity; hence, the essential genotypes, namely LN86-1595, Gutwein-180, AP 1776, L2-3, Vernal, CX210, Toano, L4303 and 9162 were identified based on mean performance and multivariate analysis. Among the 13 SSR markers, Gm011 was found to be the most informative, while Gm012 was the least insightful for the genetic diversity of studied germplasm.The polymorphic information contents (PIC) indicating a moderate level of genetic diversity exhibited by the assessed germplasm. UPGMA clustering divided the germplasm into three distinct groups, signifies that the genetic architecture of USA genotypes and Pakistani varieties has variability. Potential genotypes with higher grain yield, enhanced quality and broad genetic diversity will play a significant role in hybrid-oriented breeding programs to develop high yielding varieties with improved quality attributes. Moreover, the molecular diversity information will help plant breeders and researchers to postulate the best strategy and identify the desirable parents for future breeding program.

Acknowledgements

This work was supported by Higher Education Commission (HEC) Islamabad, Pakistan and Pakistan Agriculture Research Council under project PARB-961.

References

1) Malik, M.F., M. Ashraf, A.S. Qureshi and A. Ghafoor. 2007. Assessment of genetic variability, correlation and path analyses for yield and its components in soybean. Pak. J. Bot. 39:405-413.

2) Aditya, J.P., P. Bhartiya, and A. Bhartiya. 2011. Genetic variability, heritability and character association for yield and component characters in soybean (*G. max* (L.) Merrill). J. Cent. Eur. Agric. 12:27–34.

3) Marconato, M. B., E.M. Pereira, F.M. Silva, E.H. Bizari, J.B. Pinheiro, A.O. Mauro, and S.H. Uneda-Trevisoli. 2016. Genetic divergence in a soybean (*Glycine max*) diversity panel based on agro-morphological traits. Genet. Mol. Res. 15:1–10.

4) Li, S., H.Xu, J. Yang and T. Zhao. 2019. Dissecting the genetic architecture of seed protein and oil content in soybean from the Yangtze and Huaihe river valleys using multi-locus genome-wide association studies. Int. J. Mol. Sci. 20:1-18.

5) Idrees, M., and M. Irshad. 2014. Molecular markers in plants for analysis of genetic diversity: a review. Eur. Acad. Res. 2:1513–1540.

ISSN: 1673-064X
E-Publication: Online Open Access
Vol: 67 Issue 08 | 2024
DOI: 10.5281/zenodo.13358030

6) Nadeem, M.A., M.A. Nawaz, M.Q. Shahid, Y. Dogan, G. Comertpay, M. Yıldız, R. Hatipoglu, F. Ahmad, A. Alsaleh, N. Labhane, O.H. zkan. 2018. DNA molecular markers in plant breeding: current status and recent advancements in genomic selection and genome editing. Biotechnol. Biotechnol. Equip. 32:261–285.

7) Varshney, R.K., A. Graner and M.E. Sorrells. 2005. Genic microsatellite markers in plants: features and applications. Trends Biotechnol. 23:48–55.

8) Hipparagi, Y., R. Singh, D.R. Choudhury and V. Gupta. 2017. Genetic diversity and population structure analysis of Kala bhat (*Glycine max* (L.) Merrill) genotypes using SSR markers. Hereditas. 154:1–11.

9) Hwang, T.Y., B.S. Gwak, J. Sung and H.S. Kim. 2020. Genetic diversity patterns and discrimination of 172 Korean soybean (*Glycine max* (L.) merrill) varieties based on SSR analysis. Agriculture. 10:1–15.

10) Bezerra, A.R., T. Sediyama T, F.L. da Silva, A. Borem, A.F. da Silva and F.C. S. Silva. 2017. Agronomical aspects of the development of cultivars. In Soybean breeding. Springer. Cham. 395-411.

11) Sulistyo, A., F.C. Indriani, M.J. Mejaya, A.N. Sugiharto, and J. Agranoff. 2019. Genetic diversity of Indonesian soybean (*Glycine max* L. Merrill) germplasm based on morphological and microsatellite markers. InIOP Conference Series: Environ Earth Sci. 293:1–11.

12) Iqbal, Z., M. Arshad, M. Ashraf, R. Naeem, M.F. Malik and A. Waheed. 2010. Genetic divergence and correlation studies of soybean [*Glycine max* (L.) Merrill.] Genotypes. Pak. J. Bot. 42:971–976.

13) Maughan, P.J., M.A.S. Maroof, G.R. Buss and G.M. Huestis. 1996. Amplified fragment length polymorphism (AFLP) in soybean: species diversity, inheritance and near isogenic line analysis. Theor. Appl. Genet. 93:392–401.

14) Thompson, J.A., and R.L. Nelson. 1998. Utilization of diverse germplasm for soybean yield improvement. Crop Sci. 38:1362–1368.

15) Maughan, P.J., M.A. S. Maroof, and G.R. Buss. 1995. Microsatellite and amplified sequence length polymorphism in cultivated and wild soybean. Genome. 38:715–725.

16) Abe, J., D.H. Xu, Y. Suzuki, A. Kanazawa and Y. Shimamoto. 2003. Soybean germplasm pools in Asia revealed by nuclear SSRs. Theor. Appl. Genet. 106:445–453.

17) Wang, L., R. Guan, L. Zhangxiong, R. Chang and L. Qiu. 2006. Genetic diversity of Chinese cultivated soybean revealed by SSR markers. Crop Sci. 46:1032–1038.

18) Fu, Y.B., G.W. Peterson and M.J. Morrison. 2007. Genetic diversity of Canadian soybean cultivars and exotic germplasm revealed by simple sequence repeat markers. Crop Sci. 47:19-47.

19) Li, W., Y. Han, D. Zhang, M. Yang, W. Teng, Z. Jiang, L. Qiu and G. Sun. 2008. Genetic diversity in soybean genotypes from north-eastern China and identification of candidate markers associated with maturity rating. Plant Breed. 127:494–500.

20) Singh, R.K., V.S. Bhatia, K.V. Bhat, T. Mohapatra, N.K. Singh, K.C. Bansal and K.R. Koundal. 2010. SSR and AFLP based genetic diversity of soybean germplasm differing in photoperiod sensitivity. Genet. Mol. Bio. 33:319–324.

21) Tantasawat, P., J. Trongchuen, T. Prajongjai, S. Jenweerawat, and W. Chaowiset. 2011. SSR analysis of soybean (*Glycine max* (L.) Merr.) Genetic relationship and variety identification in Thailand. Aust. J. Crop Sci. 5:283–290.

22) Kumar, A., A. Pandey, C. Aochen, and A. Pattanayak. 2015. Evaluation of genetic diversity and interrelationships of agromorphological characters in soybean (*Glycine max*) genotypes. Proc. Natl. Acad. Sci. India B.

ISSN: 1673-064X
E-Publication: Online Open Access
Vol: 67 Issue 08 | 2024
DOI: 10.5281/zenodo.13358030

23) Oliveira, M.M., L.B. Sousa, M.C. Reis, E.S. Junior, D.B. Cardoso, O.T. Hamawaki and A.P. Nogueira. 2017. Evaluation of genetic diversity among soybean (*Glycine max*) genotypes using univariate and multivariate analysis. Genet. Mol. Res. 16:1–10

24) Tripathi, N., and D. Khare. 2016. Molecular approaches for genetic improvement of seed quality and characterization of genetic diversity in soybean: a critical review. Biotechnol. Lett. 38:1645–1654.

25) Hyten, D.L., Q. Song, Y. Zhu, I.Y. Choi, R.L. Nelson, J.M. Costa, J.E. Specht, R.C. Shoemaker, P.B. Cregan. 2006. Impacts of genetic bottlenecks on soybean genome diversity. Proceedings of the National Academy of Science of the United States of America. 103: 16666-16671.

26) Rao, V.R., and T. Hodgkin. 2002. Genetic diversity and conservation and utilization of plant genetic resources. Plant Cell. Tissue Organ. Cult. 68:1–19.

27) Nawaz, M.A., X. Lin, T.F. Chan, J. Ham, T.S. Shin, S. Ercisli, K.S. Golokhvast, H.M. Lam and G. Chung. 2020. Korean wild soybeans (*Glycine soja* Sieb & Zucc.): geographic distribution and germplasm conservation. Agronomy. 10:1–20

28) Mihaljevic, Zulj., M., H. Sarcevic, A. Lovric, Z. Andrijanic, A. Sudaric, G. Jukic and I. Pejic. 2020. Genetic diversity of European commercial soybean [*Glycine max* (L.) Merr.] Germplasm revealed by SSR markers. Genet. Resour. Crop Evol. 18:1–4.

29) Samyuktha, S.M., J.R. Kannan Bapu and S. Geethanjali. 2018. Molecular genetic diversity and population structure analysis in chickpea (*Cicer arietinum* L.) germplasm using SSR markers. Int. J. Curr. Microbiol. App. Sci. 7:639–651.

30) Tribhuvan, K.U., P. Sharma, A. Das, K. Kumar, A. Tyagi, A.U. Solanke, R. Sharma, P.V. Jadhav, M. Raveendran and B. Fakrudin. 2019. Identification of genomic SSRs in cluster bean (*Cyamopsis tetragonoloba*) and demonstration of their utility in genetic diversity analysis. Ind. crop Prod. 133:221–231.

31) Dong, D., X. Fu, F. Yuan, P. Chen, S. Zhu, B. Li, Q. Yang, X. Yu, and D. Zhu. 2014. Genetic diversity and population structure of vegetable soybean (*Glycine max* (L.) Merr.) In China as revealed by SSR markers. Genet. Resour. Crop Evol. 61:173–183.

32) Tasma, I.M., N.M. Yani, R. Purwaningdyah, D. Satyawan, K. Nugroho, P. Lestari, K.R. Trijatmiko and M. Mastur. 2018. Genetic diversity analysis and F2 population development for breeding of long juvenile trait in soybean. Agro. Biogen. 14:11–22.

33) Kachare, S., S. Tiwari, N. Tripathi and V.V. Thakur. 2020. Assessment of genetic diversity of soybean (*Glycine max*) genotypes using qualitative traits and microsatellite markers. Agric. Res. 9:23–34.

34) Bisen, A., D. Khare, P. Nair and N. Tripathi N. 2015. SSR analysis of 38 genotypes of soybean (*Glycine max* (L.) Merr.) Genetic diversity in India. Physiol. Mol. Biol. Plants. 21:109–115.

35) Gupta, S.K., and J.G. Manjaya. 2017. Genetic diversity and population structure of Indian soybean [*Glycine max* (L.) Merr.] Revealed by simple sequence repeat markers. J. Crop Sci. Biotechnol. 20:221–231.

36) Fu, K., Z. Guo, X. Zhang, Y. Fan, W. Wu, D. Li, Y. Peng, L. Huang, M. Sun, S. Bai and X. Ma. 2016. Insight into the genetic variability analysis and cultivar identification of tall fescue by using SSR markers. Hereditas. 153:1–9.

37) Appiah-Kubi, D., J.Y. Asibuo, M.D. Quain, A. Oppong, and R. Akromah. 2014. Diversity studies on soybean accessions from three countries. Biocatal. Agric. Biotechnol. 3:198–206.