OPEN ACCESS

Research Article

Analysis of Genetic Variations in Soybean using Simple Sequence Repeat (SSR) Markers

Mashal Malik and Mudassar Nawaz Khan*

Genomics and Bioinformatics Division, Institute of Biotechnology and Genetic Engineering, The University of Agriculture Peshawar, Khyber Pakhtunkhwa, Pakistan.

Abstract | Soybean is nutritionally rich source of proteins and fats. It is grown in many countries of world including US, China, Brazil and India. Pakistan grows soybean on a very limited land owing to lack of farmers interest and government attention. Soybean lines; B24G14, B23G5, B20G16, B21G2, B21G9, B6G23, B23G16 and B29G11 were grown for one month and analyzed for morphological and molecular variations. In order to analyze genetic variations, Single sequence repeat (SSR) markers were used. The line B20G16 showed the highest morphological growth parameters of plant height, total number of leaves, leaf width, leaf length and leaf weight, while line B29G11 revealed the lowest. The maximum plant height (128.03 cm) was found for line B20G16. A maximum number of 48 leaves with leaf width of 2.34 cm, leaf length of 3.34 cm leaf and leaf weight of 156.11 mg were recorded in line B20G16. The line B29G11 grew with the least plant height (51.01 cm), number of leaves (24), leaf width (1.13 cm), leaf length (2.14 cm) and leaf weight (136.67 mg). Five SSR primers (Soy satt 001, Soy satt 005, Soy satt148, Soy satt160, Soy satt171) which were previously reported to be highly polymorphic and mapped on different linkage groups, were used for analyzing genetic variations in soybean. All the primers were amplified with two different-sized alleles against each primer. Dendrogram was constructed using software SAS by Neighbor-Joining method which clustered the soybean lines into two groups. Maximum genetic dissimilarity (7) was found between line B24G14 and line B23G16. From the results of the study, it was clear that B20G16 had maximum morphological growth, while B29G11 revealed minimum growth among the tested lines, although these two lines were clustered into the same cluster interestingly. It can be concluded that line B20G16 may be further analyzed and might be useful in breeding programs.

Introduction

Soybean (*Glycine max* (L.) Merr.) belongs to family Leguminoseae, sub-family Papilionaceae, with the genome size of 1.1 kb (Newell and Hymowitz, 1983). Soybean had originated in North and Central China and emerged as domestic crop in 11th century B.C. in the eastern half of Northern China; migrated to Southern China, Korea, Japan and South-East Asia probably between 11th to 13th centuries B.C (Hymowitz, 1970). In both South and North hemisphere geographical locations, soybean has wide range of latitudinal adaptation, thus reflecting subsequent breeding history and a very complex domestication origin (Carter *et al.,* 2004). The soybeans can germinate on well-drained soils

with maximum crop production on fertile loam soils. Soybean shows no sensitivity to acidic soils as many other legumes. For crop growth, the suitable soil pH is 6-7. There is enough availability of calcium and magnesium in this range of pH. Soybean is a high-quality food owing to its rich protein and oil content in comparison with other plants (Anderson and Bush, 2011). Soybean contains 40-42% protein and 18-22% oil varying with environmental and genetic factors (Krishnan, 2001). Soybean is blessed with 8 essential amino acids necessarily required for human proteins synthesis (Fukushima, 2001).

Soybean seeds were released commercially in Pakistan in the 1970s along with sunflower and safflower (Khurshid et al., 2017). Soybean cultivation is over a vast range of agro-ecological zones including rain-fed land. Lack of genetic breeding work on soybean has not popularized soybean among the farmers despite suitable soil and climatic conditions of Pakistan (Arshad et al., 2006). From the late 1990s to 2002, soybean cultivation area in the country increased to more than 6,000 hectares; however, the area under cultivation decreased in the recent years (Khurshid et al., 2017). Reports showed that Pakistan imported 2.2 million tons of oilseeds in year 2019 (USDA, 2020). Keeping in the mind the above mentioned facts, Pakistan should grow oilseeds crops at large area to cut its import bill.

The low genetic diversity due to less genetic pool in soybean is a challenge for breeders (Carter et al., 2004). Simple sequence repeats (SSRs), also called microsatellites, are polymerase chain reaction (PCR)-based markers. SSRs are the random repeats of nucleotide units dispersed through the genome of eukaryotic organisms (Schlotterer and Tautz, 1992). SSR is considered as one of the most efficient and powerful tool for analyzing genetic diversity due to its abilities of co-dominant inheritance, reproducibility, easy detection by PCR, affluence, far-reaching genome coverage and multiple alleles (Moe et al., 2012). The variability and large number of repeat sequences makes them an excellent tool for pedigree analysis, genotype differentiation, particular genotype identification, and genetic distance evaluation among organisms.

SSR markers have been used in rice (Temnykh et al., 2001), wheat (Roder et al., 1995), tomato (Broun and Tanksley, 1996), sorghum (Schloss et al., 2002) and soybean (Cregan et al., 1999; Tantasawat et al., 2011). Cultivated soybean has multiple origins from dissimilar wild soybean populations; this was revealed through chloroplast DNA's Haplotype analysis (Xu et al., 2002). Koutu et al. (2019) profiled 8 soybean varieties with 54 polymorphic SSR primers and detected 216 alleles with 4 alleles per locus. All analyzed loci were polymorphic. Tiwari et al. (2019) analyzed genetic diversity in 148 soybean lines using 26 SSR markers. The study detected a total of 71 alleles with an average of 2.8 alleles per locus. Seventy-five soybean genotypes of soybean were analyzed for genetic variability using 21 SSR markers (Hipparagi et al., 2017). A total of 60 alleles were amplified with an average of 2.85 alleles per locus. In a study conducted in India (Bisen et al., 2015), 38 soybean genotypes were analyzed for genetic variations using 16 SSR markers. Fifty-one alleles with an average of 2.22 alleles per locus were detected. In China as well as in Japan, for studying of genetic variations and flow of gene among wild soybean and cultivated soybean, the SSR markers were used (Kuroda et al., 2006; Jin et al., 2003). SSRs in soybean were revealed to be remarkably polymorphic (Diwan and Cregan, 1997; Akkaya et al., 1992). In U.S genetic pool, SSRs were used for the evaluation of the diversity of 43 commercial and ancestral cultivars (Akkaya et al., 1992). All these studies confirmed tremendous potential of SSR in unraveling genetic diversity in soybean germplasm.

Reduction in genetic base can occur by intensive selection which will have serious consequences for progress of future breeding. The local and international varieties grown need to be analyzed for their genetic variability. Knowledge regarding the genetic variations of the grown varieties/lines would be helpful to develop high-yield soybean varieties that can fit within the local climate, tolerate abiotic and biotic stresses, as well as give high yields. The soybean lines used in the study were B24G14, B23G5, B20G16, B21G2, B21G9, B6G23, B23G16 and B29G11. The current study was designed to analyze morphological and SSR-based genetic variations in the above-mentioned eight soybeans lines.

Materials and Methods

Plant materials and germination
Seeds of eight soybean lines were surface sterilized with ethanol, acetone and then rinsed with water three-times. The eight soybean lines were provided by the department of Plant Breeding and Genetics and

grown in green house at Institute of Biotechnology and Genetic Engineering, The University of Agriculture Peshawar. The studied soybean lines included B24G14 (Line 1), B23G5 (Line 2), B20G16 (Line 3), B21G2 (Line 4), B21G9 (Line 5), B6G23 (Line 6), B23G16 (Line 7) and B29G11 (Line 8). Soybean lines were grown in triplicates as each pot consisted of 3 seeds of single line, using complete randomized design. Standard agronomic practices such as irrigation and weeding were carried out for the plants during its growth.

Morphological measurements

Morphological plant growth parameters such as plant height, total number of leaves, leaf width, leaf length and leaf weight were measured and recorded.

Molecular characterization

DNA extraction

The DNA extraction was carried out from leaf samples by CTAB method (Doyle and Doyle, 1990). Approximately one gram of fresh leaves was collected from plants and were crushed in 700 µL of CTAB buffer that consisted of 100 mM Tris HCl, 1.4 M NaCl, 20 mM EDTA (pH 8). The extracts were incubated in water bath for 1 hour at 60 C°. After incubation, equal amount of Phenol: Chloroform: Isoamyl alcohol (24:25:1) was added to each sample tube. The samples were centrifuged at the speed of 12,000 rpm for 7 minutes. The aqueous phase was isolated in new tubes. Then equal amount of ice-cold isopropanol was added and the tubes were kept at -20 °C for overnight. The precipitated DNA was then centrifuged at 12,000 rpm for 7 minutes and the pellets were washed with 70% ethanol and centrifuged again. The pellet was kept for drying for about 1 hour and then the extracted DNA was suspended in 30µL of Tris-EDTA buffer (10 mM Tris-HCl, 1 mM EDTA, at pH 7.4).

DNA quantification and SSR primers optimization

In order to determine the quality and quantity of extracted DNA, Nano Drop (NanoDrop-200Oc, Thermo Scientific) spectrophotometer was used. The analysis performed at optical density of 260/280 showed us the purity of DNA. The recorded observations were compared with standard. Purity ranged from 1.7 to 2.0 at 260/280. The genomic DNA was diluted to a working concentration of 100 ng/µL and stored at 4 ºC. The five primers used

in the molecular analysis (Table 1) were selected based on the findings of Cregan et al. (1999). All the primers were tested to obtain an optimum annealing temperature for the primer (Table 1).

Polymerase chain reaction and data analysis

For the amplification of the SSRs, PCR was carried out. In a 10 µl reaction mixture, 5µl PCR mix, 3µl of double distilled water, 0.5µl of each SSR primer and 1µl of template were mixed. The PCR was carried out in Thermocycler (Applied Biosystems 2720) using 95 ºC for denaturation, followed by 95 ºC for 20 seconds, annealing temperature vary for each primer (Table 1), was set for 10 seconds and extension step was carried out at 72 ºC for 2 min, the cycle is repeated 40 times. PCR products obtained were resolved on a 1% agarose gel in 1X TBE buffer for approximately 1 hour at 100 volts. Ethidium bromide (1µg/mL) was used for staining. After running the gel, the results were visualized using gel documentation system (Alphalmager EC). The amplified PCR products was scored qualitatively with scoring as present (1) and absent (0) for each primer. Jaccard's coefficient was used to find out similarity coefficient among the verities. UPGMA clustering algorithm method was used further for the analysis of similarity matrix. Soybean lines were clustered into groups by constructing dendrogram using software SAS by Neighbor-Joining method.

Results and Discussion

Morphological parameters

Plant height

Average plant height (cm) from triplicate grown eight soybean lines was recorded and standard deviation was calculated (Figure 1). The maximum plant height (128.03 cm) was recorded in line B20G16 while minimum plant height (51.01 cm) was recorded in line B29G11. The variations in plant height among the lines were statistically significant ($p < 0.05$) when determined by One-way ANOVA.

Number of leaves

Average number of leaves from triplicate grown 8 soybean lines was recorded (Figure 2). The maximum number of leaves (48) was found in line B20G16 and minimum numbers (24) were reported in line B29G11. The variations among the lines were statistically significant ($p < 0.05$) when determined by One-way ANOVA.

Table 1: *Primer sequences used in the study.*

| Primer name	Forward primer	Reverse primer	Annealing temperature
Soy satt 001	TGTGCAATGATAGTACATAGATAT	GTGCTGATTGAACTATTTGTAGT	50 °C
Soy satt 005	TATATCCTAGAGAAGAACTAAAAAA	GTCGATTAGGCTTGAAATAATAC	50 °C
Soy satt 148	TTAAGGATTAATTGAGACAAAATCA	CTAAAGCATCACAAAACAGAGC	48 °C
Soy satt 160	ACATCAAAAGTTTATAACGTGTAG	CTCCCACACAGTTTTCATATAAT	48 °C
Soy satt 171	TTGAGGGCTCCCACACAGTT	CAAAAGTTTATAACGTGTAGATTAA	51 °C
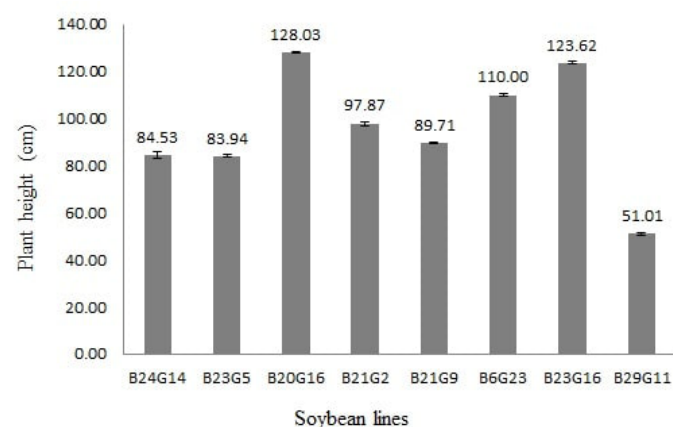
Figure 1: *Plant height variations in soybean lines grown for one-month.*
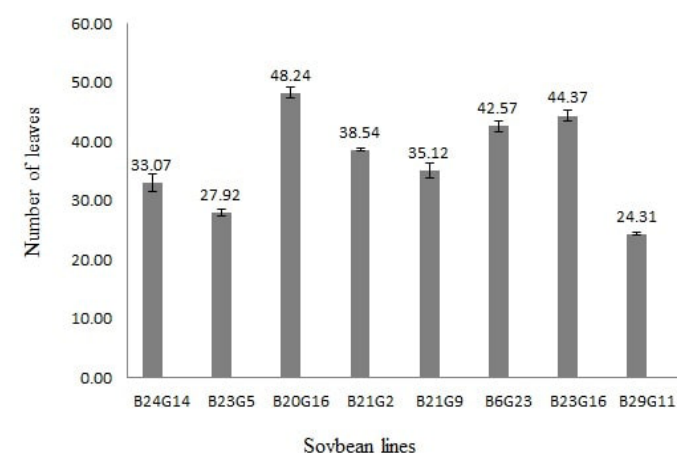

Figure 2: *Total number of leaves in soybean lines grown for one-month.*

Leaf width

Average leaf width (cm) from triplicate grown 8 soybean lines was recorded and standard deviation was calculated (Figure 3). The maximum leaf width (2.34 cm) was recorded in line B20G16 and minimum (1.13 cm) in B29G11. The variations in leaf width among the lines were statistically significant (p<0.05) when determined by One-way ANOVA.

Leaf length

Average leaf length from triplicate grown 8 soybean lines was recorded and standard deviation was calculated (Figure 4). The maximum leaf length (3.34

cm) was recorded in line B20G16 and minimum (2.14 cm) in B29G11. The variations were statistically significant (p<0.05) when determined by One-way ANOVA.
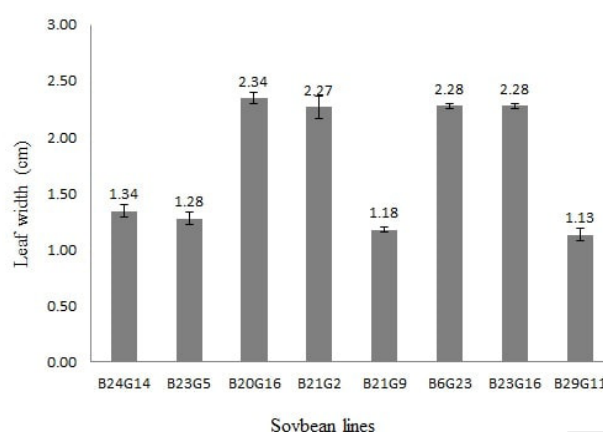

Figure 3: *Plant leaf width variations in soybean lines grown for one month.*

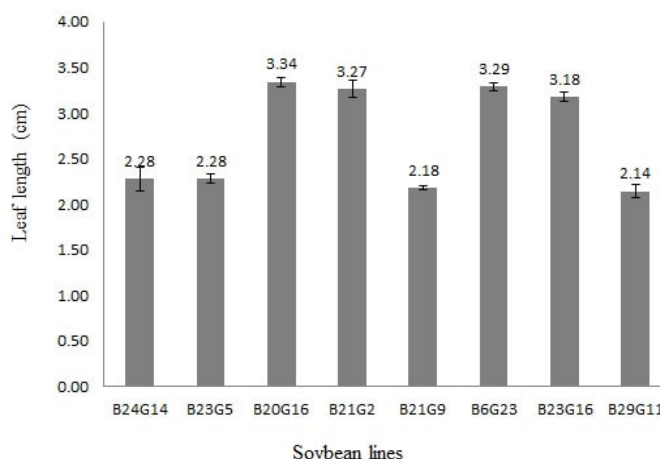**Figure 4:** *Leaf length variations in soybean lines grown for one-month.*

Leaf weight

Average leaf weight from triplicate grown 8 soybean lines was recorded and standard deviation was calculated (Figure 5). The maximum leaf weight (156 mg) was recorded in line B20G16 and minimum (136 mg) in B29G11. The variations in weight were statistically significant (p<0.05) when determined by One-way ANOVA.
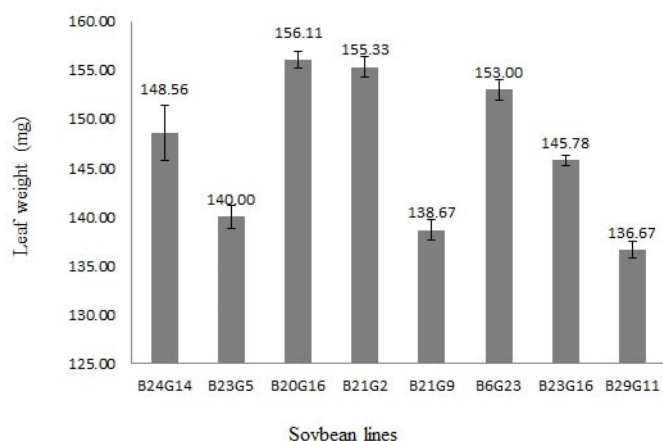
Figure 5: *Leaf weight variations in soybean lines grown for one month.*

PCR amplification with selected primers

For PCR amplification of SSR markers, primers Soy satt 001, Soy satt 005, Soy satt 148, Soy satt 160, Soy satt 171 were used to detect genetic variability in the selected soybean lines. A total of 8 lines were amplified with different sized 2 alleles against each primer. Allele size ranged from 191 to 200 for PCR amplification product of primers Soy satt 001 (Figure 6A). Allele size ranged from 120 to 129 for PCR amplification product of primers Soy satt 005 (Figure 6B). Allele size ranged from 120 to 126 for PCR amplification product of primers Soy satt 148 (Figure 6C). Allele size ranged from 123 to 129 for PCR amplification product of primers Soy satt 160 (Figure 6D). For primer Soy satt 171, allele size of PCR amplification product ranged from 151 to 158 (Figure 6E). The frequency of each allele among these soybean lines varied from as low as 0.125 to as high as 0.875. The details of allelic frequency for each amplification are given in Table 2.

Table 2: *Allelic sizes and frequency of PCR amplification product using selected primers.*

SSR primer	No. of alleles	Allele size	Allelic frequency
Soy satt 001	2	191	0.75
		200	0.25
Soy satt 005	2	120	0.25
		129	0.75
Soy satt 148	2	120	0.75
		126	0.25
Soy satt 160	2	123	0.125
		129	0.875
Soy satt 171	2	151	0.250
		158	0.750
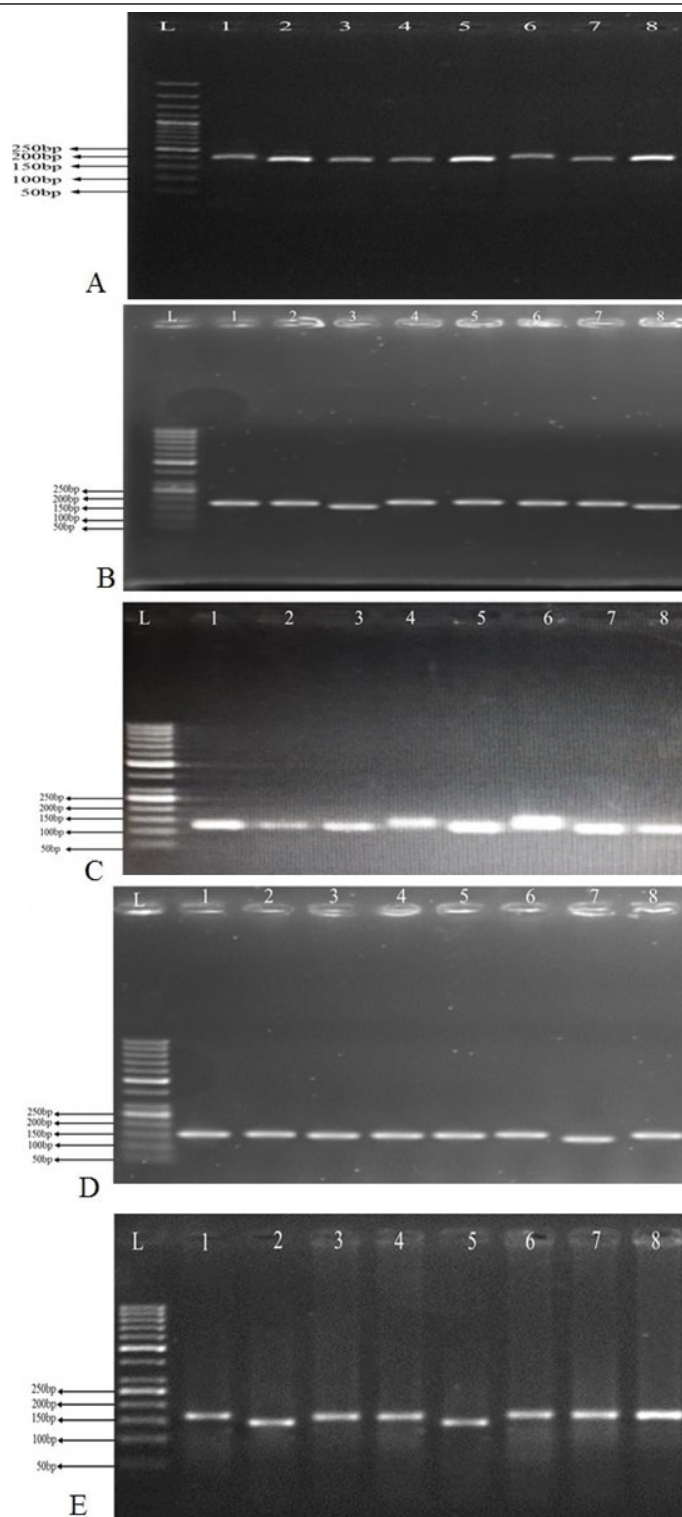
Figure 6: *PCR amplification of SSR markers in 8 soybean lines using primer Soy satt 001 (A), Soy satt 005 (B), Soy satt 148 (C), Soy satt 160 (D), Soy satt 171 (E). L represent ladder. Numbers above bands presents soybean lines.*

Genetic relationship among the selected soybean lines

Dendrogram showed two main clusters or groups for eight soybean lines. Each of them was further divided into sub-groups (Figure 7). The sub-group I consisted of line B24G14, line B6G23 and line B21G2. The sub-group II consisted of Lines B23G5, B21G9, B20G16, B29G11 and B23G16.

```
                            +------------------- line B24G14
               +-----------3
+--------------4           +------------------line B6G23
!              !           !
!              !           +------------------------line B21G2
!              !
--7                        +line B23G5
!              +-----------------------------1
!              !           +line B21G9
!           +--5
!           ! !           +line B20G16
+-----------6 +-----------------------------2
!                         +line B29G11
!
+-----------------------------------------line B23G16
```
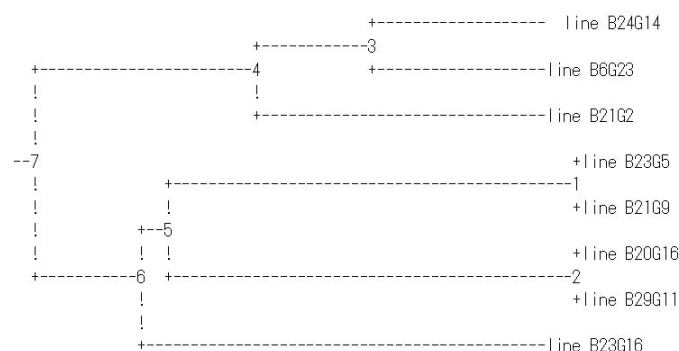
Figure 7: *Dendrogram of soybean lines showing genetic divergence.*

Soybean is leguminous crop rich in protein and oil content (Anderson and Bush, 2011). Keeping in mind the agricultural importance of soybean, current research was designed to analyze the morphological and genetic variations in soybean lines that previously showed better yields in the fields. Eight lines were grown in green house and morphological parameters such as plant height, total number of leaves, leaf width, leaf length and leaf weight were measured. The soybean lines were analyzed at molecular level for genetic variations by using SSR markers mapped at different linkage groups (Cregan et al., 1999).

Conventionally, genetic variation in soybean has been based on the changes in agro-morphological traits and pedigree information. This evaluation is significant for breeding programs. Soybean varieties are improved to specific agro-ecological areas and the phenotypes are greatly influenced by environmental factors (Li and Nelson, 2001). In the current study, soybean lines showed noticeable variations regarding morphological parameters. The line B20G16 showed highest whereas, line B29G11 lowest morphological parameters. The maximum plant height (128.03 cm), total number of leaves (48), leaf width (2.34 cm), leaf length (3.34 cm) and leaf weight (156.11 mg) was recorded in line B20G16. The least growth parameters of plant height (51), total number of leaves (24), leaf width (1.13 cm), leaf length (2.14 cm) and leaf weight (136.67 mg) were recorded in line B29G11. Results of the current study were almost in agreement with the findings of Malik et al. (2009), who reported high variability in plant height in the range of 47.67 cm to 90.27 cm. Pankaj (2013) and Awal (2014) reported plant heights lesser than current study (30 cm and 31 cm, respectively) in various soybean genotypes. The variations in plant height might account for genetic make-up as well environmental factors.

Leaf is essential for plants because it is the site for photosynthesis. The numbers of leaves per plant and leaf surface have essential characteristics for germplasm classification. In the present study, number of leaves ranged from 24 to 48. Similar results were published by Mudibu et al. (2011) who reported 16 to 44 numbers of leaves per plant. Puech et al. (1974) reported 14 maximum numbers of leaves in his study. The variations in number of leaves per plant are indicating noticeable differences in plant photosynthesis capabilities. Present study showed leaf length ranged from 2.14 to 3.34 cm and leaf width ranged from 1.13 to 2.34 cm. Mudibu et al. (2011) reported mean values for leaf length vary from 5.7 mm to 11.9 mm. Significant variations were also recorded between leaf width, which ranged from 4.6 to 6.9 mm. The results of the current study revealed that lines B20G16, B21G2 and B23G16 revealed comparable higher growth rate as compared to lines B24G14, B23G5, B21G9, B6G23 and B29G11.

Genetic variation using morphological characters and molecular markers have been reported in many species, particularly in Apium (Castellini et al., 2006), chamomile (Solouki et al., 2008) and coriander (Lopez et al., 2008). Genetic variation using molecular markers have been reported in the *Changium smyrnioides* through RAPD (Fu et al., 2003), in Carum L. by Papini et al. (2007), in carrot through AFLP by Santos and Simon (2002) and in soybean through SSR (Cregan et al., 1999; Tantasawat et al., 2011). In the current study, all the primers were amplified in each soybean genotype. A size variation of amplification product against each primer was noted as it was evident in genotypes. From the results of dendrogram, it was seen that lines B20G16 and B29G11 were grouped in the same sub-cluster that indicated their closer genetic resemblance. However, there was huge variation in morphological characters among them. Comparable results were found by Solouki et al. (2008), where out of 25 soybean lines, only 11 revealed resemblances in both morphological and molecular results. These findings specified a slight association among the two dendrogram created with markers and phenotype. Hamza et al. (2004) and Liu et al. (2007) revealed a perfect correlation among the morphological characters and genetic variations as discovered by RAPD analysis. Conversely, Schut et al. (1997) and Eshraghi et al. (2006) conveyed a few correlations among the morphological and molecular characters. Similar results for other species

have been reported by Ali *et al.* (2007) and Lopez *et al.* (2008). Assumed this, the vast variation b/w the lines B20G16 and B29G11 regarding morphological growth parameters and their genetic clustering into the same sub-group in dendrogram point out that these molecular markers might not be linked to the studied growth parameters.

Conclusions and Recommendations

Analysis of eight soybean lines using SSR markers revealed that there was genetic diversity among these lines and also in-terms of growth parameters of plant height, total number of leaves, leaf width, leaf length and leaf weight. The soybean lines B20G16 showed the highest growth parameters among the tested line; whereas, line B29G11 revealed the lowest growth parameters among the tested lines. Furthermore, two different-sized alleles were detected for each primer amplification. It is interesting to note that lines B20G16 and B29G11 revealed extremely contrasting growth rates; however, they were seen clustered in the same sub-group in the dendrogram. This close linkage in dendrogram and contrast in growth parameters could be because these SSR markers might not be linked to the analyzed morphological growth parameters. Based on the findings, it is recommended that these genotypes along with other available germplasm, may be studied on a larger scale to select better genotypes for breeding programs.

Acknowledgements

Novelty Statement

The selected genotypes were analyzed for first-time for their growth parameters and genetic linkages. Results of the current study may provide a base-line for selection in breeding programs with extensive germplasm of soybean.

Author's Contribution

Mudassar Nawaz Khan: Designed, supervised the research and also wrote the manuscript.
Mashal Malik: Performed all the experiments of the study.

Conflict of interest
The authors have declared no conflict of interest.

References

Akkaya, M.S., A.A. Bhagwat and P.B. Cregan. 1992. Length polymorphisms of simple sequence repeat DNA in soybean. Genetics, 132(4): 1131-1139. https://doi.org/10.1093/genetics/132.4.1131

Ali, G.M., S. Yasumoto and M. Seki-Katsuta. 2007. Assessment of genetic diversity in sesame (*Sesamum indicum* L.) detected by amplified fragment length polymorphism markers. Elec. J. Biotech., 10(1): 12-23. https://doi.org/10.2225/vol10-issue1-fulltext-16

Anderson, J.W. and H.M. Bush. 2011. Soy protein effects on serum lipoproteins: A quality assessment and meta-analysis of randomized, controlled studies. J. Am. Coll. Nutr., 30(2): 79-91. https://doi.org/10.1080/07315724.2011.10719947

Arshad, M., N. Ali and A. Ghafoor. 2006. Character correlation and path coefficient in soybean *Glycine max* (L.) Merrill. Pak. J. Bot., 38(1): 121–130.

Awal, A.M., 2014. Grain yield and seed quality of early maturing dwarf soybean genotypes as influenced by planting time and spacing. Ph.D. thesis. Dept. Agronomy, Bangabandhu Sheikh Mujibur Rahman Agricultural University, Gazipur, Bangladesh.

Bisen, A., D. Khare, P. Nair and N. Tripathi. 2015. SSR analysis of 38 genotypes of soybean (*Glycine max* (L.) Merr.) genetic diversity in India. Physiol. Mol. Biol. Plants, 21(1): 109-115. https://doi.org/10.1007/s12298-014-0269-8

Broun, P. and S.D. Tanksley. 1996. Characterization and genetic mapping of simple repeat sequences in the tomato genome. Mol. Gen. Genet., 250(1): 39-49. https://doi.org/10.1007/BF02191823

Carter, T.E., R.L. Nelson, C.H. Sneller and Z. Cui. 2004. Genetic diversity in soybean. In: Shibles, R.M., Harper, J.E., Wilson, R.F. and Shoemaker, R.C. (Ed), Soybeans: Improvement, production and uses. American Society of Agronomy, Inc. Crop Science Society of America, USA. pp. 303-416. https://doi.org/10.2134/agronmonogr16.3ed.c8

Castellini, G., R. Torricelli, E. Albertini and M. Falcinelli. 2006. Morphological and molecular characterization of celery landrace from central Italy, apium graveolens l. Var. Dulce (miller) pers. In: Proceedings of the 50th Italian society of agricultural genetics annual congress, 10-14 September, Ischia, Italy.

Cregan, P.B., T. Jarvik, T.A.L. Bush, R.C. Shoemaker, K.G. Lark, A.L. Kahler, N. Kaya, T.T. VanToai, D.G. Lohnes, J. Chung and J.E. Specht. 1999. An integrated genetic linkage map of the soybean genome. Crop Sci., 39(5): 1464-1490. https://doi.org/10.2135/cropsci1999.3951464x

Diwan, N. and P. Cregan. 1997. Automated sizing of fluorescent-labeled simple sequence repeat (SSR) markers to assay genetic variation in soybean. Theor. Appl. Genet., 95(5-6): 723-733. https://doi.org/10.1007/s001220050618

Doyle, J.J. and J.L. Doyle. 1990. Isolation of plant DNA from fresh tissue. Focus, 12: 13-15.

Eshraghi, P., R. Zarghami and H. Ofoghi. 2006. RAPD analysis of micropropagated plantlets in date palm. Pak. J. Biol. Sci., 9(1): 111-114. https://doi.org/10.3923/pjbs.2006.111.114

Fu, C., Y. Qiu and H. Kong. 2003. RAPD analysis for genetic diversity in *Changium smyrnioides* (Apiaceae), an endangered plant. Bot. Bull. Acad. Sin., 44: 13-18.

Fukushima, D., 2001. Recent progress in research and technology on soybeans. Food Sci. Tech. Res., 7(1): 8-1. https://doi.org/10.3136/fstr.7.8

Hamza, S., W.B. Hamida, A. Rebai and M. Harrabi. 2004. SSR-based genetic diversity assessment among Tunisian winter barley and relationship with morphological traits. Euphytica, 135(1): 107-118. https://doi.org/10.1023/B:EUPH.0000009547.65808.bf

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): 9. https://doi.org/10.1186/s41065-017-0030-8

Hymowitz, T., 1970. On the domestication of the soybean. Econ. Bot., 24(4): 408-421. https://doi.org/10.1007/BF02860745

Jin, Y., T. He and B. Lu. 2003. Fine scale genetic structure in a wild soybean (*Glycine soja*) population and the implications for conservation. New Phytol., 159(2): 513-519. https://doi.org/10.1046/j.1469-8137.2003.00824.x

Khurshid, H., D. Baig, S.A. Jan, M. Arshad and M.A. Khan. 2017. Miracle crop: the present and future of soybean production in Pakistan. MOJ Biol. Med., 2(1): 189–191. https://doi.org/10.15406/mojbm.2017.02.00042

Koutu, G.K., A. Shrivastava, Y. Singh and S. Tiwari. 2019. Molecular characterization and genetic diversity assessment of soybean varieties using SSR markers. Int. J. Curr. Microbiol. App. Sci., 8(4): 173-182. https://doi.org/10.20546/ijcmas.2019.804.018

Krishnan, H.B., 2001. Biochemistry and molecular biology of soybean seed storage proteins. J. New Seeds 2(3): 1-25. https://doi.org/10.1300/J153v02n03_01

Kuroda, Y., A. Kaga, N. Tomooka and D.A. Vaughan. 2006. Population genetic structure of Japanese wild soybean (*Glycine soja*) based on microsatellite variation. Mol. Ecol., 15(4): 959-974. https://doi.org/10.1111/j.1365-294X.2006.02854.x

Li, Z. and R.L. Nelson. 2001. Genetic diversity among soybean accessions from three countries measured by RAPDs. Crop Sci., 41(4): 1337-1347. https://doi.org/10.2135/cropsci2001.4141337x

Liu, P., Y.S. Yang, C.Y. Hao and W.D. Guo. 2007. Ecological risk assessment using RAPD and distribution pattern of a rare and endangered species. Chemosphere, 68(8):1497-1505. https://doi.org/10.1016/j.chemosphere.2007.03.018

Lopez, P.A., M.P. Widrlechner, P.W. Simon, S.R. Boylston, T.A. Isbell, T.B. Bailey, C.A. Gardner and L.A. Wilson. 2008. Assessing phenotypic, biochemical, and molecular diversity in coriander (*Coriandrum sativum* L.) germplasm. Genet. Resour. Crop Evol., 55(2): 247-275. https://doi.org/10.1007/s10722-007-9232-7

Malik, M.F.A., A.S. Qureshi, M. Ashraf, M.R. Khan and A. Javed. 2009. Evaluation of genetic diversity in soybean (*Glycine max*) lines using seed protein electrophoresis. Aust. J. Crop Sci., 3(2): 107-112.

Moe, K.T., W.J. Hong, S.W. Kwon and Y.J. Park. 2012. Development of cDNA-derived SSR markers and their efficiency in diversity assessment of cymbidium accessions. Elec. J. Biotech., 15(2): 4-4. https://doi.org/10.2225/vol15-issue2-fulltext-4

Mudibu, J., K.K.C. Nkongolo and A. Kalonji-Mbuyi. 2011. Morphovariability and agronomic

characteristics of soybean accessions from the Democratic Republic of Congo (DR-Congo) gene pool. J. Plant Breed. Crop Sci., 3(11): 260-268.

Newell, C.A. and T. Hymowitz. 1983. Hybridization in the genus Glycine Subgenus Glycine Wild. (Leguminosae, Papilionoideae). Am. J. Bot., 70(3): 334–348. https://doi.org/10.1002/j.1537-2197.1983.tb06400.x

Pankaj, B., 2013. Plant characterization and seed quality of some dwarf soybean genotypes. M.S. thesis. Dept. Agronomy, Bangabandhu Sheikh Mujibur Rahman Agricultural University, Gazipur, Bangladesh.

Papini, A., F. Banci and E. Nardi. 2007. Molecular evidence of polyphyletism in the plant genus *Carum* L. (Apiaceae). Genet. Mol. Biol. 30(2): 475-482. https://doi.org/10.1590/S1415-47572007000300029

Puech, J., P. Lencrerotamd and M. Hernandez. 1974. Role of environmental factors in soybean quantitative and qualitative production I. Growth, development and yield under irrigated and non-irrigated conditions (in French). Ann. Agron., 25(5): 659-679.

Roder, M.S., J. Plaske, S.U. Konig, A. Bolner, M.E. Sorrels, S.D. Tanksley and M.W. Ganal. 1995. Abundance, variability and chromosomal location of microsatellites in wheat. Mol. Gen. Genet., 246(3): 327-333. https://doi.org/10.1007/BF00288605

Santos, C.A. and P.W. Simon. 2002. Some AFLP amplicons are highly conserved DNA sequences mapping to the same linkage groups in two F2 populations of carrot. Genet. Mol. Biol., 25(2): 195-201. https://doi.org/10.1590/S1415-47572002000200013

Schloss, S.J., S.E. Mitchell, G.M. White, R. Kukatla, J.E. Browers, A.H. Paterson and S. Kresovich. 2002. Characterization of RFLP clone sequences for gene discovery and SSR development in *Sorghum bicolor* (L) Moench. Theor. Appl. Genet., 105(6-7): 912-920. https://doi.org/10.1007/s00122-002-0991-4

Schlotterer, C. and D. Tautz. 1992. Slippage synthesis of simple sequence DNA. Nucl. Acids Res., 20(2): 211-215. https://doi.org/10.1093/nar/20.2.211

Schut, J.W., X. Qi and P. Stam. 1997. Association between relationship measures based on AFLP markers, pedigree data and morphological traits in barley. Theor. Appl. Genet. 95(7): 1161-1168. https://doi.org/10.1007/s001220050677

Solouki, M., H. Mehdikhani, H. Zeinali and A.A. Emamjomeh. 2008. Study of genetic diversity in Chamomile (*Matricaria chamomilla*) based on morphological traits and molecular markers. Sci. Hort., 117(3): 281-287. https://doi.org/10.1016/j.scienta.2008.03.029

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(3): 283-290.

Temnykh, S., G. DeClerck, A. Lukashova, L. Lipovich, S. Cartinhour and S. McCouch. 2001. Computational and experimental analysis of microsatellites in rice (*Oryza sativa* L.): frequency, length variation, transposon associations, and genetic marker potential. Genome Res., 11(8): 1441-1452. https://doi.org/10.1101/gr.184001

Tiwari, S., N. Tripathi, K. Tsuji, K. Tantwai. 2019. Genetic diversity and population structure of Indian soybean (*Glycine max* (L.) Merr.) as revealed by microsatellite markers. Physiol. Mol. Bio. Plants 25(4): 953-964. https://doi.org/10.1007/s12298-019-00682-4

USDA, 2020. https://www.usda.gov/ accessed on October 21, 2020.

Xu, D.X., J.A. Abe, J.G. Gai and Y.S. Shimamoto. 2002. Diversity of chloroplast DNA SSRs in wild and cultivated soybeans: evidence for multiple origins of cultivated soybean. Theor. Appl. Genet., 105(5): 645-653. https://doi.org/10.1007/s00122-002-0972-7