Azarian Journal of Agriculture

Azarian Journals

Screening of Soybean (*Glycine max* L.) Genotypes through Multivariate Analysis

Kamrun Nahar Mili[1], Bir Jahangir Shirazy[2], Md. Mostofa Mahbub[3]

Article Info	ABSTRACT
	The physiological divergence was assessed in twenty-seven soybean genotypes by using principal component analysis, cluster mean analysis, principal coordinate analysis and canonical variate analysis to identify parental genotypes for the future breeding program in order to develop new high yielding varieties in a randomized complete block design with three replications. The genotypes under the experiment were grouped into five clusters. The highest number of genotypes found in cluster III. The highest intra-cluster distance was found in cluster II and while cluster V showed no intra-cluster distance values which revealed homogenous nature of the genotype within the cluster. The highest inter-cluster distance was found between cluster I and IV followed by I and V. Cluster II have early flowering genotypes whereas early maturity in cluster III and most of the desirable traits were found in cluster IV. Days to first flowering and pod length from cluster II, whereas pods per plant and yield per plant from cluster IV have the positive relative contribution to the entire divergence. According to principal component scores, LG-92P-1176 followed by KANH-33, AGS-79, MTD-452, GMOT-17, GC-82-332411, MTD-451 and BS-33 have the prominent influence towards varietal improvement. Selecting genotypes from distant clusters probably provide promising recombinants and better segregants for the future breeding platform.

INTRODUCTION

Soybean (*Glycine max* (L.) Merrill) is a leguminous crop which is one of the richest sources of oil as well as protein. Soybean kernel contains 36% protein, 35% carbohydrate, 19% oil, 5% minerals and some other components with vitamins (Abady et al. 2013). As a noble source of protein, unsaturated fatty acids, minerals like Ca and P together with vitamin A, B, C, and D, soybean can meet up different nutritional requirements (Mahbub et al. 2015). The current nutritional condition of the third world and some emerging countries like Bangladesh is a matter of great fear since the most of the publics are suffering from malnutrition (Mahbub et al. 2015

Shirazy et al. 2015; Hossain et al. 2016). Soybean can show a significant role in this case and can aid to meet up the nutritional shortage difficulty. Soya protein products can replace animal based foods (Henkel 2000). Furthermore soybean also contain frequent compounds that perform as antioxidant and are valuable to human fitness as they reduce the danger of cardiovascular syndromes, breast cancer, osteoporosis, diabetes and neurodegenerative sicknesses such as Alzheimer's and Parkinson's and decrease the menopausal signs (Kumar et al. 2014) and being a leguminous crop it recovers soil fertility and productivity by fixing the atmospheric nitrogen through *Rhizobium* bacteria that lives in root nodules (Jaiswal et al. 2012; Youseif et al. 2014). The statistical information regarding soybean acreage and production in Bangladesh is fragmentary and not available in systematic form, but it is thought to be of 10,000 acres with about 5,000 tons of production (Rahman 2003). As a result it is very suitable crop to fit into the cropping systems of Bangladesh.

Considering the potentiality of this crop, there is a need for improvement and to develop varieties suited to specific agro-ecological conditions and also for specific end use. Genetic diversity is a major factor that determines prospects of yield

[1] Department of Genetics and Plant Breeding, Sher-e-Bangla Agricultural University, Dhaka, Bangladesh
[2] Scientific Officer, Rice Farming Systems Division, Bangladesh Rice Research Institute (BRRI), Gazipur, Bangladesh
[3] Scientific Officer, Agronomy Division, BRRI, Gazipur, Bangladesh
*Email: bjshirazy@yahoo.com

improvement in future (Rahman et al. 2012). Knowledge of genetic diversity within a crop and correlation among the yield contributing characters is essential for the long-term success of a breeding program and maximizes the exploration of germplasm resources (Rahman et al. 2014). The quantification of genetic diversity through biometrical procedures (Rao 1952 and Anderson 1957) has made it possible to choose genetically diverse parents having high genetic divergence (Upadhaya and Mehta 2010). Moreover, evaluation of genetic diversity is important to know the source of genes for a particular trait within the available germplasm (Tomooka 1991). Multivariate analysis acts as a useful tool to quantify the degree of divergence between the biological populations at genotypic level and to assess the relative contribution of different components to the total divergence both inter and intra-cluster levels (Mahbub et al. 2016). The utility of multivariate analysis for measuring the degree of divergence and assessing the relative contribution of different character to the total divergence in self-pollinated crops has been established by several workers (Golakia and Makne 1992). In the present study, genetic divergence of soybean has been assessed in the light of eleven important morphological and physiological traits to identify suitable genotypes for the breeding program.

MATERIALS AND METHODS

The study was conducted to assess the morphogenetic divergence among twenty-seven soybean genotypes. The experiment was carried out during December 2013 to April 2014 at the field laboratory of Sher-e-Bangla Agricultural University, Dhaka, Bangladesh. Soybean genotypes were sown in a randomized complete block design with three replications; each plot consisted of a single row of 4m long with a row to row distance of 1m maintaining 10 plants per meter. Sowing was done with the help of hand drill. Ten random plants were used to take the data from field. All intercultural operations were carried out following standard procedures as described by Mondal et al. 2001. All intercultural operations were carried out

following standard procedures as described by Mondal et al. 2001.

Statistical analysis

The data were analyzed by Mahalanobis D^2 statistics, principal component analysis, cluster mean analysis, principal coordinate analysis (PCoA) and canonical variate analysis (CVA). Twenty-seven soybean genotypes were grouped into the cluster by using Genstat v 5.5 software.

RESULTS AND DISCUSSION

Cluster analysis

The distribution pattern in table 1 indicated that the maximal number of genotypes (11) was comprised of cluster III followed by cluster I (6), cluster II (6), cluster IV (3) and cluster V (1). Mikel et al. 2010, Wang et al. 2012 and Mahbub et al. 2016 also grouped the soybean genotypes into different clusters and found a different number of genotypes in the different cluster. The constructed dendrogram (Figure 1) revealed twenty seven soybean genotypes into five major clusters.

There was much variation among the clusters compared to intra-cluster variations (Table 2 and Figure 2). The highest intra-cluster distance was observed in cluster II (1.61) and the lowest in cluster V (0.000). The inter-cluster distance was maximum between cluster I and IV (15.10) followed by cluster I and V (12.89); cluster IV and V (12.24); cluster II and IV (11.07); and cluster I and II (10.07) suggesting that the genotypic materials belonging to these clusters may be used as parents for hybridization program to develop desirable variety because crosses between genetically divergent lines will generate heterotic segregants (Khan et al. 2016). From Table 3 it was observed that G8 and G10 have highest inter-genotypic distance whereas G3 and G6 have the lowest distance. G8 found in cluster IV and G10 in Cluster I, which have the highest inter-cluster distance (Table 1). Moreover, we know that heterosis can be best exploited and chances of getting transgressive segregants are maximum when generating diverse lines are crossed. Because

Table 1. Distribution of genotypes of soybean in different clusters

Cluster no.	No. of Genotypes	Number of populations	Name of genotypes
I	G1, G4, G10, G16, G17, G19	6	LG-92P-1176, AGS-79, SHOHAG, AGS-95, BARI SOYBEAN-6, MTD-451
II	G2, G21, G23, G24, G26, G27	6	P1-4174-75, 86017-66-6, LG-92P-12-18, BS-33, BS-13, CHINA-1
III	G3, G5, G6, G7, G9, G12, G13, G18, G20, G22, G25	11	KANH-33, MTD-452, GMOT-17, JOYAWAZA, YESOY-4, GC-82-332411, PK-327, NS-1, GC-830059, MTD-16, ASSET-95
IV	G8, G14, G15	3	F-85-11347, ASSET-93-19-13, PK-329
V	G11	1	AUSTRALIA

2

the genotypes from diverge cluster may be advised for inclusion in hybridization program as they are expected to develop excellent segregants (Mahbub and Shirazy, 2016).

Cluster means for the characters

Intra-cluster mean for 12 physiological characters are present in Table 4. It was perceived that cluster IV showed the maximum value for plant height, the number of branches per plant, number of pod per plant, stover yield, single pod weight per plant and yield per plant. A number of seeds per pod have the higher value in cluster V; pod length and 100 seed weight in cluster II. Cluster II suited for early flowering and cluster III for early maturity which is also very much desirable. Among these clusters, the cluster II and

Table 2. Intra (bold) and inter cluster distances (D^2) for 27 genotypes of soybean

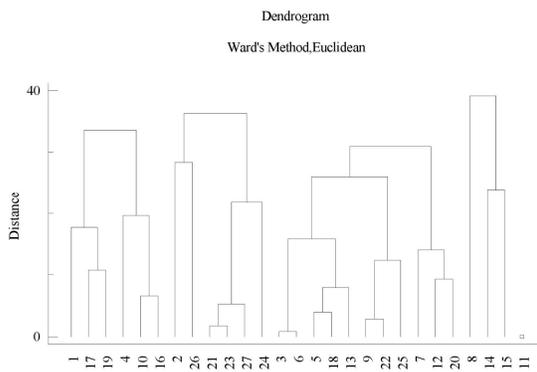| Cluster	I	II	III	IV	V
I	**1.55**	10.07	5.32	15.10	12.89
II		**1.64**	6.34	11.07	2.93
III			**1.40**	10.01	9.24
IV				**1.41**	12.24
V					**0.00**

Figure 1. Dendrogram showing genetic relationships among the soybean genotypes

IV showed the maximum desirable traits which could be reflected in the assortment of parents for the future breeding program. Mahbub et al. (2016) also study about cluster mean values for different quantitative characters with soybean genotypes.
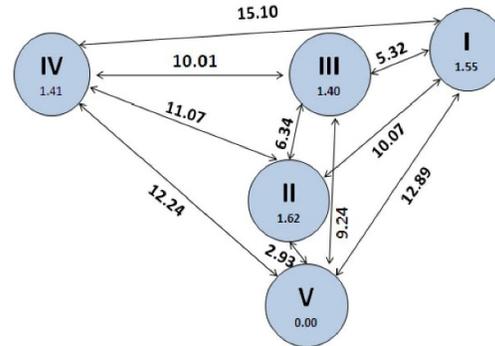

Figure 2. Intra and inter cluster distances (D2) of 27 genotypes in soybean

The genotypic materials from these clusters could be used for breeding purpose in crop improvement project.
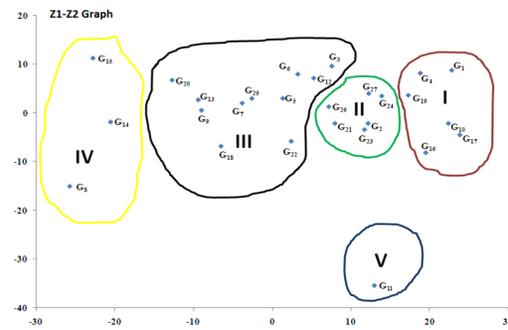

Figure 3. Scatter diagram of 27 genotypes of soybean based on their principle component scores super imposed with clustering

Contribution of traits towards diversity

The characters having the greater contribution to the total divergence were given highest emphasis for grouping on the cluster for hybridization. The contribution of characters towards divergence obtain from CVA is presented in Table 5, where the values of vector-1 and 2 transpired that contribution of pod length and pods per plant were found prominent to the total divergence. In vector-1, the other important traits responsible for differentiation were days to first flowering, number of pod per plant, pod length and yield per plant

Table 3. Ten highest and ten lowest inter genotypic distance among 27 genotypes of soybean

	Highest distance				Lowest distance		
Sl No.	Genotype		Distance	Sl No.	Genotype		Distance
1	G8	G10	10.51	1	G3	G6	1.74
2	G10	G8	10.51	2	G6	G3	1.74
3	G4	G8	10.17	3	G21	G23	1.81
4	G24	G8	9.47	4	G23	G21	1.81
5	G17	G8	9.06	5	G27	G23	1.93
6	G16	G8	9.01	6	G5	G18	2.13
7	G11	G8	8.83	7	G22	G21	2.14
8	G12	G8	8.76	8	G9	G22	2.19
9	G19	G8	8.69	9	G18	G5	2.23
10	G24	G18	8.59	10	G13	G9	2.26

Table 4. Cluster mean values of 12 different characters of 27 genotypes of soybean

Characters	I	II	III	IV	V
Plant height (cm)	48.78	54.36	57.38	69.9	34.57
Days to first flowering	81.44	81.34	82.03	86.67	77.67
Days to 50% flowering	84.16	84	84.45	90	81
Days to 80% maturity	124.28	128.67	123.88	133.44	157
Number of branches per plant	5.35	5.61	6.42	9.05	4.47
Number of pod per plant	38.7	43.1	48.28	92.2	38.27
Number of seeds per pod	2.54	3	2.96	3	3.33
Pod length (cm)	3.19	3.99	3.62	3.79	3.95
100 seed weight (g)	8.13	13.02	7.14	8.93	12.37
stover yield (g)	4.74	6.75	8.31	14.36	8.21
Single pod weight per plant (g)	2.96	5.49	3.78	6.99	4.72
Yield per plant (g)	12	14.21	19.4	35.73	22.59

Table 5. Relative contributions of the 12 characters of 27 varieties to the total divergence

Characters	Vector -1	Vector- 2
Plant height (cm)	-0.1883	0.0054
Days to first flowering	2.195	0.7711
Days to 50% flowering	-2.0313	-0.6766
Days to 80% maturity	-0.1178	0.0178
Number branches per plant	-1.1267	-0.1455
Number of pod per plant	0.0708	0.137
Number of seed per pod	-6.7284	-5.0663
Pod length (cm)	0.499	0.9861
Hundred seed weight (g)	-0.4394	-0.8838
Stover yield (g)	-0.5726	-0.2168
Single pod weight per plant (g)	-0.6258	0.2357
Yield per plant (g)	0.112	0.047

while in vector-2, plant height, days to first

flowering, days to 80% maturity, number of pod per plant, pod length, single pod weight per plant and yield per plant was important. A number of seed per pod and days to 50% flowering has highly negative relation with vector-1 and number of seed per pod with vector-2. Interesting plant height, days to 80% maturity and single pod weight per plant have the negative relation with one vector but the positive relation with opposite vector. So during selection program, these characters should be under considering for the breeding program.

In table 6, where the values of PCA 1 and PCA 2 exposed that contribution of LG-92P-1176 followed by KANH-33, AGS-79, MTD-452,

GMOT-17, GC-82-332411, MTD-451 and BS-33 were found noticeable to the total divergence. In PCA 1, the other important genotypes responsible for differentiation were P1-4174-75, SHOHAG, AUSTRALIA, AGS-95, BARI SOYBEAN-6, 86017-66-6, MTD-16, LG-92P-12-18, BS-13 and CHINA-1; while in PCA 2, the JOYAWAZA, YESOY-4, PK-327, PK-329, GC-830059, ASSET-95, BS-13 and CHINA-1 were important. Emphasis should be given on SHOHAG, AUSTRALIA, PK-327, AGS-95, BARI SOYBEAN-6 and GC-830059 also for their negative characters because positively and

Table 6. Z1-Z2 score of 27 genotypes of soybean

Genotypes	PCA 1	PCA 2
LG-92P-1176	22.85	8.81
P1-4174-75	12.22	-2.09
KANH-33	7.61	9.67
AGS-79	18.86	8.22
MTD-452	1.38	3.04
GMOT-17	3.3	8.02
JOYAWAZA	-3.78	2.06
F-85-11347	-25.71	-15.02
YESOY-4	-8.96	0.59
SHOHAG	22.44	-2.1
AUSTRALIA	13.02	-35.4
GC-82-332411	5.31	7.2
PK-327	-9.37	2.75
ASSET-93-19-13	-20.52	-1.81
PK-329	-22.75	11.3
AGS-95	19.55	-8.11
BARI SOYBEAN-6	23.88	-4.45
NS-1	-6.48	-6.79
MTD-451	17.33	3.67
GC-830059	-12.69	6.79
86017-66-6	8.01	-2.1
MTD-16	2.45	-5.77
LG-92P-12-18	11.76	-3.37
BS-33	13.96	3.53
ASSET-95	-2.56	3.01
BS-13	7.23	1.33
CHINA-1	12.3	4.03

negatively distant genotypes could provide some fascinating outcomes.

Based on principal component scores -1 and 2 obtained from the principal component analysis, a two-dimensional scatter diagram using component score 1 as X axis and component score 2 as Y axis was constructed, which has been presented in Figure 3. The positions of the genotypes in the diagram were apparently distributed into scattered, which point out that considerable diversity existed among the genotypes.

Selection of the parents

Identification and utilization of diverse germplasm are the central issues in plant breeding. Three factors (choice of the particular cluster, selection of specific variety from a cluster and relative contribution of the character to the total

divergence) should be reflected for deciding on parents for a breeding program (Chaudhary et al. 1977). More accurate and complete description of genotypes and patterns of genetic diversity could help determinate future breeding strategies and facilitate introgression of diverse germplasm into the current commercial soybean genetic base. Genetic diversity analysis reveals genetic backgrounds and interactions of germplasm and manages crop primary pools (Brown-Guedira et al. 2000). Principal component analysis is useful as it gives information about the groups where certain traits are more important allowing the breeders to conduct specific breeding program (Salimi et al. 2012). Considering all the above criteria genotypes from cluster II and IV could be selected as parents for the future breeding program.

CONCLUSION

Based on the results of this experiment, it may be concluded that breeding program should be undertaken by picking as parental genotypes from diverse distant clusters along with considering the cluster mean values of different traits for find the desirable characters which have active relative contribution to the total divergence aimed at developing anticipated varieties by selection of superior genotypes through the successive positive principal coordinate and canonical variants involvement generations.

REFERENCES

Abady S. Merkeb F. Dilnesaw Z. (2013) Heritability and path-coefficient analysis in soybean (*Glycine Max* L. Merrill) genotypes at Pawe, North Western Ethiopia. Journal of Environmental Science and Water Resource, 2(8): 270 – 276.

Anderson E. (1957) A semigraphical method for the analysis of complex problems. Proceedings of the National Academy of Sciences, 43(10): 923-927.

Kumar A. Kumar A. Lal S.K. Jolly M. Sachdev A. (2014) Influence of gamma rays and ethyl methane sulphonate (EMS) on the levels of phytic acid, raffinose family oligosaccharides and antioxidants in soybean seeds of different genotypes. Journal of Plant Biochemistry and Biotechnology, Pp: 1-6.

Mahbub M.M. Rahman M.M. Hossain M.S. Mahmud F. Kabir M.M.M. (2015) Genetic variability, correlation and path analysis for yield and yield components in soybean. American-Eurasian Journal of Agricultural & Environmental Sciences, 15(2): 231-236.

Mahbub M.M. Rahman M.M. Hossain M.S. Nahar L. Shirazy B.J. (2016) Morphophysiological Variation in Soybean (*Glycine max* (L.) Merrill). American-Eurasian Journal of Agricultural and Environmental Sciences, 16(2):234-238.

Mahbub M.M. Shirazy B.J. (2016) Evaluation of Genetic Diversity in Different Genotypes of Soybean (*Glycine max* (L.) Merrill). American Journal of Plant Biology, 1(1):24-29.

Chaudhary B.D. Kakar S.N. Singh R.K. (1977) Estimation of genetic parameters in barley (*Hordeum vulgare* L.). Theoretical and Applied Genetics, 49(4):153-156.

Brown-Guedira G.L. Thompson J.A. Nelson R.L. Warburton M.L. (2000) Evaluation of genetic diversity of soybean introductions and North American ancestors using RAPD and SSR markers. Crop Science, 40(3):815-823.

Salimi S. Lahiji H.S. Abadi G.M. Salimi S. Moradi S. (2012) Genetic diversity in soybean genotypes under drought stress condition using factor analysis and cluster analysis. World Applied Sciences Journal, 16(4):474-478.

Golakia P.R. Makne V.G. (1992) D^2 analysis in Virginia runner groundnut genotypes. Indian Journal of Genetics, 55 (3): 252-256.

Henkel J. (2000) Soybean: Health claims for soybean protein, questions about other components. FDA Consumer. www.cfsan.fda.gov [Accessed: 19 January 2017].

Khan M.R.A. Mahbub M.M. Reza M.A. Shirazy B.J. Mahmud F. (2016) Selection of Field Pea (*Pisum sativum* L.) Genotypes through Multivariate Analysis. ScientiaAgriculturae, 16 (3): 98-103.

Hossain M.M. Mahbub M.M. Shirazy B.J. (2016) Growth and Yield Performance of Mugbean Varieties in Summer Cultivation. Scientia Agriculturae, 16 (3): 79-82.

Shirazy B.J. Islam M.M. Haque M.A. Mahbub M.M. Somee T.A. (2015) Influence of Combined Effect of Nitrogen and Micronutrients on Yield and Yield Contributing Characters of Seasame (*Seasamum indicum* L.). Botany Research International, 8(4): 73-76.

Jaiswal S.K. Anand A. Dhar B. Vaishampayan A. (2012) Genotypic characterization of phage-typed indigenous soybean bradyrhizobia and their host range symbiotic effectiveness. Microbial ecology, 63(1): 116-126.

Youseif S.H. El-Megeed F.H.A. Ageez A. Mohamed Z.K. Shamseldin A.Saleh S.A. (2014) Phenotypic characteristics and genetic diversity of rhizobia nodulating soybean in Egyptian soils. European Journal of Soil Biology, 60: 34-43.

Rahman M.M. Rasaul M.G. Hossain M.A. Iftekharuddaula K.M. Hasegawa H. (2012) Molecular Characterization and Genetic Diversity Analysis of Rice (*Oryza sativa* L.)

Using SSR Markers. Journal of Crop Improvement, 26(2): 244-257.

Rahman M.M. Syed M. Akter A. Alam M.M. Ahsan M.M. (2014) Genetic Variability, Correlation and Path Coefficient Analysis of Morphological Traits in Transplanted Aman Rice (*Oryza sativa* L.). American-Eurasian Journal of Agricultural and Environmental Sciences, 14(5): 387-391.

Rahamn L. (2003) Studies on the development of varieties, production technology, food and fish feed uses of soybean in Bangladesh (BAU-USDA soybean project BG-ARS 107), p: 6.

Rao C.R. (1952) Advanced statistical methods in Biometrical Research. John willy and sons. New York. Pp:390.

Tomooka N. (1991) Genetic diversity and landrace differentiation of mungbean, *Vigna radiata* (l.) Wilczek, and evaluation of its wild relatives (The subgenus ceratotropis) as breeding materials, Techinical Bulletin of Tropical Agriculture Research Center. Ministry of Agriculture, Forestry and Fisheries, Japan. 28. pp.10-17.

Mondal M.R.I. Wahhab M.A. (2001) Production technology of oil crops. Oil seed research centre. Bangladesh Agricultural Research Institute. 57-67.

Mikel M.A. Diers B.W. Nelson R.L. Smith H.H. (2010) Genetic diversity and agronomic improvement of North American soybean germplasm. Crop Science, 50(4): 1219-1229.

Upadhayay D. Mehta N. (2010) Biometrical Studies in Dolches Bean (*Dolchos lablab* L.) for Chhattisgarh plains. Research Journal of Agricultural Sciences, 1(4): 441-447.

Wang K.J. Li X.H. (2012) Genetic diversity and geographical peculiarity of Tibetan wild soybean (*Glycine soja*). Genetic Resources and Crop Evolution, 59(4): 479-490.