EVALUATION OF GENETIC DIVERSITY AMONG SOYBEAN [*GLYCINE MAX* (L.)] GENOTYPES USING MULTIVARIATE ANALYSIS

Nidhi Dubey, Harshal A. Avinashe* and A. N. Shrivastava[1]

Department of Genetics and Plant Breeding, Lovely Professional University, Phagwara - 144 411 (Punjab), India.
[1]Department of Plant Breeding and Genetics, JNKVV, Jabalpur - 482 003 (M.P), India.

Abstract

Divergence analysis among 50 genotypes was carried out using Mahalonobis D^2 statistic. The genotypes were grouped into ten clusters. Maximum differences among the genotypes within the same cluster (intracluster) were shown by cluster II (141.18) followed by cluster V (125.80), cluster VII (119.12) and cluster I (95.74). Clusters III, IV, VI, VIII, IX and X showed zero intracluster distances. Diversity among the clusters varied from 53.05 to 3181.79 intercluster distances. Cluster II and X showed maximum inter cluster distance (3181.79) followed by that between cluster II and IX (2270.24). The lower intercluster distance was noticed between cluster IV and VI (53.05) followed by that between cluster VI and VII (71.78). Genotypes belonging to clusters with maximum intra-cluster distance are genetically more divergent and hybridization between divergent clusters is likely to produce wide variability with desirable segregants. The most important trait causing maximum genetic divergence was biological yield per plant (78.53%) followed by days to maturity (8.90%), protein content (5.55%), seed yield per plant (3.10%), days to 50% flowering (2.29%) and oil (1.14%). Hence, it is advisable to select divergent parents based on these three characters and attempt crossing between them so as to achieve a broad spectrum of favourable genetic variability for yield improvement in soybean.

Key words : Soybean, genetic divergence, D^2 analysis, cluster analysis, inter and intra-cluster distance.

Introduction

Soybean (*Glycine max* L. Merrill) is an important source of high quality protein and oil. It is however, characterized by low yield only because of low yield varieties, lodging and pod shattering, which are major production constrains. Soybean being the richest, cheapest and easiest source of best quality proteins and fats and having a vast multiplicity of uses as food and industrial products is sometimes called a wonder crop (Gopinath *et al.*, 2015). The assessment of available genetic variability is of utmost importance in all the crop improvement programmes. This is important for several reasons: the ability to distinguish reliably different genotypes is important for designing the breeding programmes, population-genetic analysis, genetic engineering and an estimation of the amount of variation within genotypes and between genotypes is useful for predicting potential genetic gains in a breeding programme and in setting up appropriate cross-breeding strategies (Bhakuni *et al.*,

2017). Genetic variability is the basic requirement for crop improvement as this provides wider scope for selection. Knowledge of diversity patterns will allow breeders to better understand the evolutionary relationships among accessions, to sample germplasm in a more systematic fashion and to develop strategies to incorporate useful diversity in their breeding programs (Naik *et al.*, 2016). Genetic diversity is the most important tool in the hands of the plant breeder in choosing the right type of parents for hybridization programme. To analyze genetic diversity in any population methods have relied on pedigree data, morphological data, agronomic performance data, biochemical data, and more recently molecular (DNA-based) data. For reasonably accurate and unbiased estimates of genetic diversity, adequate attention has to be devoted to utilization of various data sets using clustering procedures and other multivariate methods in analyses of data. The divergence can be studied by technique using D^2 statistics developed by Mahalanobis (1936). It is based on multivariate analysis and grouped into various cluster. This is considered as

---

the most effective method to assess the genetic diversity present among the genotypes included in the study. The present investigation aimed to estimate the magnitude of genetic divergence present in the genotypes and to identify the diverse genotypes for future.

Materials and Methods

The present study was conducted under All India Coordinated Research Project on Soybean at Seed Breeding Farm, Department of Plant Breeding and Genetics, College of Agriculture, Jawaharlal Nehru Krishi Vishwa Vidhyalaya Jabalpur (M.P.), India to analyze the genetic diversity among soybean genotypes during *Rabi*, 2013-14. Source and pedigree of the material are given in table 1.

The field experiment was laid out in randomized block design (RBD) with three replications. Fifty genotypes were planted with a spacing of 30 cm row to row and 10 cm plant to plant distance. All the recommended agronomical practices and plant protection measures were adopted to raise the healthy crop. The data was recorded on plant height at harvest (cm), number of pods per plant, number of pods per node, number of seeds per

plant, number of seeds per pod, 100 seed weight (g), biological yield per plant (g), harvest index (%), seed yield per plant (g), protein content (%) and oil content (%) on a sample of 5 plants per replication in each genotypes whereas for days to flowering and days to maturity data were taken on whole plot basis.

The data on morphological traits was subjected to analysis of variance on the basis of model described by Panse and Sukhatme (1985) for individual characters. The replicated data were subjected to genetic divergence analysis using Mahalanobis's D^2 – statistic (Mahalanobis, 1936).

Results and Discussion

In the present investigation, 13 important yield and yield contributing traits have been studied to evaluate the pattern and extent of genetic variability and relation among 50 genotypes of soybean. The ANOVA for different characters revealed that mean sum of squares due to genotypes were highly significant for all the characters indicating the presence of significant genetic variability among genotypes, which provides scope for selection and further use of these genotypes in crop improvement. It

Table 1 : Pedigree and source of 50 genotypes of *Glycine max* L. Merrill.

S. no.	Variety	Pedigree	S. no.	Variety	Pedigree
01.	JS 335	JS 78-77 X JS 71-5	26.	JS 20-98	JS 97-52 X SL 710
02.	JS 93-05	2nd selection from PS 73-22	27.	JS 20-100	JS 20-30 X JS 93-05
03.	JS 95-60	2nd selection from PS 73-22	28.	JS 20-101	JS 97-52X[JS 97-51XSL 96]
04.	JS 97-52	PK 327 X L129	29.	JS 20-102	JS 97-52 X JSM 281
05.	JS 20-29	JS 97-52 X JS 95-56	30.	JS 20-103	JS 97-52 X JSM 299
06.	JS 20-34	JS 98-63 X PK 768	31.	JS 20-104	JS 92-22 X SL 517
07.	JS 20-30	JS 97-52 X SL 710	32.	JS 20-105	JS 97-52X[JS 97-51XSL 96]
08.	JS 20-35	JS 99-81 X JS 99-83	33.	JS 20-106	JS 99-76 X JSM 275
09.	JS 20-41	JS 97-52 X JS 20-02	34.	JS 20-107	JS 97-52 X JSM 281
10.	JS 20-53	JS 97-52 X JS 20-02	35.	JS 20-108	JS 97-52 X JSM 286
11.	JS 20-65	JS 97-52XJS(IS)90-5-12-1	36.	JS 20-109	JS 97-52 X JSM 299
12.	JS 20-68	JSM 240 X JSM 189	37.	JS 20-110	JS 97-52 X JS 93-05
13.	JS 20-69	JS 97-52 X SL 710	38.	JS 20-111	JS 97-52 X JS 95-60
14.	JS 20-71	JS 97-52XJS(IS)90-5-12-1	39.	JS 20-112	JS 97-52 X SL 96
15.	JS 20-72	JS 99-81 X JS 99-83	40.	JS 20-113	JS 92-22 X SL 517
16.	JS 20-76	JS 97-52 X SL 710	41.	JS 20-114	JS 99-76 X JSM 275
17.	JS 20-79	JS 97-52XJS(IS) 90-5-12-1	42.	JS 20-115	JS 97-52 X JSM 288
18.	JS 20-87	JS 97-52XJS(IS) 90-5-12-1	43.	JS 20-116	JS 97-52 X JSM 120 A
19.	JS 20-89	JS 97-52 X JSM 286	44.	JS 20-117	JS 99-76 X JSM 275
20.	JS 20-90	JS 97-52 X JS 95-56	45.	JS 20-118	JS 97-52 X JS 20-02
21.	JS 20-92	JS 97-52 X JSM 52	46.	JS 20-119	JS 97-52 X JS(IS) 90-5-12-1
22.	JS 20-94	JS 97-52 X JS 20-02	47.	JS 20-120	JS 97-52 X SL 710
23.	JS 20-95	JS 97-52 X JS 95-56	48.	JS 20-121	JS 97-52 X JSM 52
24.	JS 20-96	JS 97-52 X JSM 286	49.	JS 20-122	JS 97-52 X JSM 120
25.	JS 20-97	JSM 259 X PK 768	50.	JS 20-123	JS 97-52 X JS 95-56

Table 2 : ANOVA and Per cent contribution of characters towards divergence in 50 soybean genotypes.

S. no.	Characters	Source of variation			CV	Percent contribution	Rank
		Replication	Treatments	Error			
	degree of freedom	2	49	98			
1.	Days to 50% flowering	0.06	103.852**	1.086	1.58	2.29%	V
2.	Days to maturity	0.093	246.487**	1.52	0.97	8.90%	II
3.	Plant height	18.391	150.232**	12.046	9.13	0.00%	
4.	No. of pods plant^{-1}	143.836	754.511**	60.775	12.38	0.08%	
5.	No. of pods node^{-1}	0.888	3.642**	0.761	13.99	0.00%	
6.	No. of seeds plant^{-1}	122.775	1942.256**	184.714	16.19	0.00%	
7.	No. of seeds pod^{-1}	0.089	0.304**	0.058	17.38	0.00%	
8.	100 seed weight	0.01	4.666**	1.019	8.11	0.08%	
9.	Biological yield plant^{-1}	0.196	501.771**	0.32	1.51	78.53%	I
10.	Harvest index	6.808	429.396**	26.028	16.85	0.08%	
11.	Seed yield plant^{-1}	1.067	30.307**	2.936	16.44	3.10%	IV
12.	Protein content (%)	1.06	20.611**	0.368	1.52	5.55%	III
13.	Oil content (%)	1.076	5.512**	0.368	3.41	1.14%	VI

*, ** Significant at 5% and 1% level.
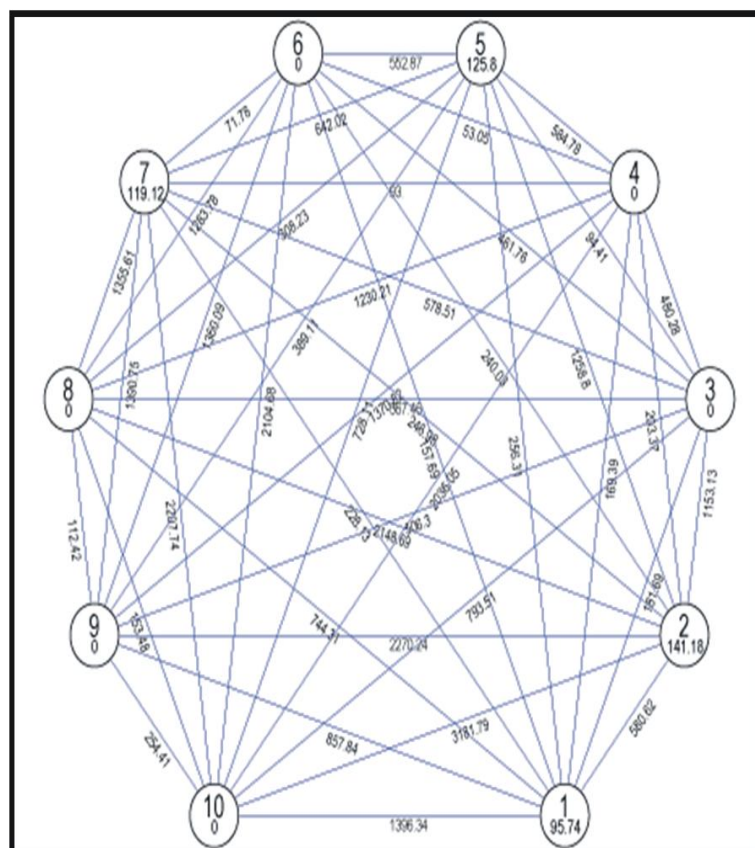
Fig. 1 : Intra and inter-cluster distance for 50 genotypes of soybean.

reported by Reni and Rao (2013).

A method suggested by Tocher (Rao, 1952) was used to group the genotypes into different clusters based on the D^2 values. 50 genotypes were grouped into ten clusters. Among ten clusters, cluster I was the biggest with 16 genotypes followed by cluster II and V with 12 genotypes and cluster VII with 4 genotypes. Cluster III, IV, VI, VIII, IX and X were solitary. The clustering pattern and the distribution of genotypes into different clusters are presented in table 3.

The average D^2 value of intra and inter cluster distances are given in table 4 (fig. 1). Maximum differences among the genotypes within the same cluster (intracluster) were shown by cluster II (141.18) followed by cluster V (125.80), cluster VII (119.12) and cluster I (95.74). Clusters III, IV, VI, VIII, IX and X showed zero intracluster distances. Diversity among the clusters varied from 53.05 to 3181.79 intercluster distances. Cluster II and X showed maximum inter cluster distance (3181.79) followed by that between cluster II and IX (2270.24). The lower intercluster distance was noticed between cluster IV and VI (53.05) followed by that between cluster VI and VII (71.78). The perusal of mean in table 4 revealed that inter-cluster distances were greater than intra-cluster distances revealing considerable amount of genetic diversity among the genotypes studied. Genotypes

was observed that biological yield per plant (78.53%) was the highest contributor towards divergence followed by days to maturity (8.90%), protein content (5.55%), seed yield per plant (3.10%), days to 50% flowering (2.29%) and oil (1.14%) (table 2). Similar findings were

Table 3 : Distribution of 50 genotypes into different clusters.

S. no.	Cluster	No. of genotypes	Name of genotypes
1	I	16	JS 20-115, JS 20-116, JS 20-107, JS 20-108, JS 20-35, JS 20-94, JS 20-76, JS 20-98, JS 20-72, JS 20-122, JS 97-52, JS 20-69, JS 20-79, JS 20-92, JS 20-105, JS 20-119
2	II	12	JS 20-95, JS 20-123, JS 20-113, JS 20-104, JS 20-117, JS 20-87, JS 20-90, JS 20-29, JS 20-65, JS 20-120, JS 20-121, JS 20-101
3	III	1	JS 20-53
4	IV	1	JS 20-110
5	V	12	JS 20-96, JS 20-114, JS 20-41, JS 20-103, JS 20-30, JS 20-89, JS 20-106, JS 20-100, JS 93-05, JS 20-118, JS 20-68, JS 335
6	VI	1	JS 20-71
7	VII	4	JS 20-109, JS 20-111, JS 20-102, JS 20-112
8	VIII	1	JS 95-60
9	IX	1	JS 20-34
10	X	1	JS 20-97

Table 4 : Average intra and intercluster D^2 values of soybean genotypes.

Cluster	I	II	III	IV	V	VI	VII	VIII	IX	X
I	95.74	580.62	181.69	169.39	256.31	157.69	228.13	744.31	857.84	1396.34
II		141.18	1153.13	203.37	1258.80	240.03	246.98	2148.69	2270.24	3181.79
III			0.00	480.28	94.41	461.76	578.51	367.46	506.30	793.51
IV				0.00	584.78	53.05	93.00	1230.21	1370.82	2036.05
V					125.80	552.87	642.02	308.23	389.11	728.11
VI						0.00	71.78	1283.78	1360.09	2104.68
VII							119.12	1355.61	1390.75	2207.74
VIII								0.00	112.42	153.48
IX									0.00	254.41
X										0.00

* Diagonal values – Intracluster distance

Table 5 : Mean values of thirteen characters in ten clusters in soybean genotypes.

S. no.	Cluster	No. of genotypes	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10	X11	X12	X13
1	I	16	67.65	129.96	37.60	64.26	6.45	84.91	1.33	12.31	35.75	29.43	10.38	39.62	17.37
2	II	12	70.14	133.39	44.04	72.22	6.66	99.76	1.40	12.52	54.94	22.60	12.41	38.72	17.39
3	III	1	64.00	131.67	33.80	52.93	6.26	55.13	1.05	11.87	25.34	25.60	6.47	41.19	15.55
4	IV	1	68.00	129.67	42.13	74.87	7.34	80.31	1.08	11.90	45.21	21.28	9.60	41.51	16.87
5	V	12	63.75	123.61	33.83	52.67	5.67	74.01	1.42	12.80	36.97	25.34	9.31	39.23	17.80
6	VI	1	72.00	134.00	42.67	65.73	5.70	46.01	0.70	12.47	43.40	13.21	5.73	35.79	15.45
7	VII	4	69.33	128.50	40.77	72.70	6.83	100.13	1.38	12.79	45.51	28.00	12.83	36.22	17.96
8	VIII	1	53.67	109.00	25.73	33.40	4.98	73.41	2.19	14.90	17.85	61.26	10.93	43.18	17.46
9	IX	1	51.00	101.00	24.80	34.53	4.48	69.11	2.00	11.53	17.43	45.07	7.87	35.77	17.70
10	X	1	51.00	105.00	21.87	23.80	3.92	42.67	1.80	13.50	9.43	60.94	5.73	42.30	15.48
	Mean	5	63.05	122.58	34.72	54.71	5.83	72.55	1.44	12.66	32.02	34.44	9.13	39.35	16.90

X1 - Days to 50% flowering, **X2** - days to maturity, **X3** - plant height, **X4** - Number of pods plant^{-1}, **X5** - Number of pods node^{-1}, **X6** - Number of seeds plant^{-1}, **X7** - Number of seeds pod^{-1}, **X8** - 100 seed weight, **X9** - biological yield per plant, **X10** - harvest index, **X11** - Seed yield plant^{-1}, **X12** – Protein content, **X13** – Oil content

belonging to clusters with maximum intra-cluster distance are genetically more divergent and hybridization between divergent clusters is likely to produce wide variability with desirable segregants. The results are in close proximity with the findings of Jain *et al.* (2017) and Parameshwar *et al.* (2011) and Patil *et al.* (2011).

The cluster means and general mean values for 13 characters of 50 genotypes have been represented in table 5. The data revealed that differences in cluster means had existed. Cluster I comprised of 16 genotypes which were characterized as having above average values for all the characters except number of seeds per pod, 100 seed weight and harvest index. Cluster II had 12 genotypes that indicated above average values for all the characters except number of seeds per pod, 100 seed weight, harvest index and protein content. Cluster III comprised of only one genotype which was characterized as having above average values for days to 50% flowering, days to maturity, number of pods per node and protein content. Cluster IV comprised of one genotype which was characterized as having above average values for all the characters except number of seeds per pod, 100 seed weight, harvest index and oil content. Cluster V consisting of 12 genotypes showed above average values for days to 50% flowering, days to maturity, number of seeds per plant, 100 seed weight, harvest index, seed yield per plant and oil content. Cluster VI had one genotype showed above average values for days to 50% flowering, days to maturity, plant height, number of pods per plant and biological yield per plant. Cluster VII comprising of 4 genotypes showed above average values for all the characters except number of seeds per pod, harvest index and protein content. Cluster VIII involving one genotype was characterized by above average values for number of seeds per plant, number of seeds per pod, 100 seed weight, harvest index, seed yield per plant, protein content and oil content. Cluster IX had one genotype showed above average values for number of seeds per pod, harvest index and oil content. Cluster X involving one genotype was characterized by above average values for number of seeds per pod, 100 seed weight, harvest index and protein content. Similar findings were reported by Jain *et al.* (2017) and Mahesh *et al.* (2017).

Conclusion

The present study indicated that the distribution of genotypes into different clusters was at random and sufficient D^2 values among different cluster suggests that the genetic constitution of the promising lines in one cluster is in close proximity with the promising lines in other clusters of the pair may lead to desirable segregants

having broad genetic base through hybridization between genotypes of two distant clusters. This finding will be helpful in planning future hybridization programme should involving diverse genotypes for crop improvement.

Acknowledgement

Competing interests

The authors have declared that no competing interest exists. This manuscript is neither published nor submitted for publication in some other journal or book.

References

Bhakuni, V., P. S Shukla, K. Singh and V. K. Singh (2017). Morphological Characterization and Assessment of Genetic Variability in Soybean Varieties. *International Journal of Current Microbiology and Applied Sciences*, **6(3)** : 361-369.

Gopinath, P. and P. Pavadai (2015). Morphology and Yield parameters and Biochemical analysis of Soybean [*Glycine max* (L.) Mrr.] Using Gamma rays, EMS and DES treatment. *International Letters of Natural Sciences*, **35** : 50-58.

Jain, R. K., A. Joshi and D. Jain (2017). Morphological Characterization and Assessment of Genetic Diversity in Soybean [*Glycine max* (L.) Merrill] Genotypes. *Int. J. Curr. Microbiol. App. Sci.*, **6(7)** : 2371-2381.

Mahalanobis, P. C. (1936). On the generalized distances in statistics. *Proceedings of National Institute of Sciences of India*, **2** : 49-55.

Mahesh, M., S. Jain and V. S. Bhadoria (2017). Genetic Divergence Studies in Soybean (*Glycine max* L.). *International Journal of Agriculture Science*, **9(19)** : 4196-4197.

Naik, S. M., K. Madhusudan, B. N. Motagi and H. L. Nadaf (2016). Diversity in Soybean (*Glycine max*) Accessions Based on Morphological Characterization and Seed Longevity Characteristics. *Progressive Research – An International Journal*, **11(3)** : 377-381.

Panse, V. G. and P. V. Sukhatme (1985). *Statistical Methods for Agricultural Research Workers*. Indian Council of Agricultural Research, New Delhi.

Parameshwar, M. G., B. Sujata and G. T. Basavaraja (2011). Studies on genetic divergence in soybean (*Glycine max* (L.) Merrill). *Green Farming*, **2(3)** : 276-278.

Patil, S. S., M. R. Naik, A. B. Patil and U. R. Ghod (2011). Genetic Diversity in Soybean. *Legume Research*, **34(1)** : 68-70.

Reni, P. Y. and Y. K. Rao (2013). Genetic variability in soybean [*Glycine max* (L) Merrill]. *International Journal of Plant, Animal and Environmental Sciences*, **3(4)** : 35-38.

Rao, C. R. (1952). *Advanced statistical methods in biometrical research*, John Wiley and Sons, New York, pp. 357-369.