Pak. J. Bot., 40(6): 2323-2328, 2008.

EVALUATION OF SOYBEAN [*GLYCINE MAX* (L.) MERRILL] GERMPLASM FOR SOME IMPORTANT MORPHOLOGICAL TRAITS USING MULTIVARIATE ANALYSIS

ZAFAR IQBAL, MUHAMMAD ARSHAD, MUHAMMAD ASHRAF[1], TARIQ MAHMOOD[2] AND ABDUL WAHEED

Department of Botany, Pir Mehr Ali Shah Arid Agriculture University, Rawalpindi, Pakistan.
[1] Oilseed Program, National Agriculture Research Center, Islamabad
[2]Department of Environmental Science, Pir Mehr Ali Shah Arid Agriculture University, Rawalpindi, Pakistan.
Corresponding author's e-mail: zafar_narc@yahoo.com

Abstract

One hundred thirty nine soybean genotypes were evaluated for yield and yield associated traits at National Agriculture Research Centre, Islamabad. The traits viz., days to maturity, plant height at maturity, number of branches per plant, number of filled pods per plant, number of unfilled pods per plant, 100-seed weight, oil content, grain yield per plant, biological yield per plant and harvest index were studied. The results revealed significant differences among genotypes for all studied traits. Cluster analysis based on these traits, 139 genotypes were divided into five clusters. Cluster I showed maximum number of filled pods per plant, 100-seed weight, grain yield per plant, biological yield and harvest index. Cluster II showed least oil content, grain yield per plant and harvest index. Cluster III showed maximum plant height and number of branches per plant. Cluster IV comprised of accessions having high oil content and least number of unfilled pods per plant while early maturity was observed in cluster V. Principal component analysis (PCA) showed that first three components accounted 69.77% of the total variance. PC1 indicated that number of filled pods per plant, grain yield and biological yield per plant were important traits for classification. While 100-seed weight, oil content and harvest index were important in PC2. In the PC3 days to maturity and number of branches per plant were important. (Principal component and clustering obtained from this study can make better choice for soybean breeders for selecting genotypes among large number of accessions.

Introduction

The soybean [*Glycine max (*L.) Merrill] is a papilionoid legume that has a fairly wide range of adaptation involving a wide array of climatic, soil and growth conditions though it is mostly grown on rain-fed land (Lackey, 1977; Fageria *et al*., 1997). It is now cultivated throughout East and South East Asia for food, animal feed and medicine. It is a miraculous crop due to its extraordinary qualities, it contains about 37-42% good qualities protein, 6% ash, 29% carbohydrate and 17-24% oil comprising 85% poly unsaturated fatty acid with two essential fatty acids (lenoleic and linolenic acid) which are not synthesized by the human body (Antalina, 2000; Balasubramaniyan & Palaniappan, 2003). At present, in Pakistan it is grown on a very small area only in NWFP. Average yield of soybean is low as compared to other soybean growing countries. On average of 5 years (2001-2005), the yield was 1333 kgha^{-1} (Anon., 2005-2006).

With increase in the sample sizes of breeding materials and germplasm accessions used in crop improvement programs, methods to classify and order genetic variability are assuming considerable significance. The use of established multivariate statistical algorithms is an important strategy for classifying germplasm, ordering variability for a large number of accessions, or analyzing genetic relationships among traits in any breeding materials. Multivariate analytical techniques, which simultaneously analyze multiple measurements on each individual under investigation, are widely used in analysis of genetic diversity irrespective of the dataset (morphological, bio-chemical, or molecular marker data). Among these algorithms, cluster analysis, principal component analysis (PCA), principal coordinate analysis (PCA) and multidimensional scaling (MDS) are at present most commonly applied and appear particularly useful (Melchinger, 1993; Johns *et al*., 1997; Thompson *et al*., 1998; Brown *et al*., 2000).

Materials and Methods

The experiment was carried out during July 2004 at the experimental area of Oilseeds Program, National Agricultural Research Center, Islamabad (33.40 º N and 73.07 º E). One hundred and thirty nine accessions including 3 approved varieties viz. NARC-I, NARC-II and Ajmeri were evaluated for some important agronomic traits. The experiment was planted in augmented field design (Federer & Ragavarao, 1975); each plot consisted of two rows of 5 meter long, row to row spacing was 45 cm and maintaining 20 plants per meter. Ten plants randomly selected from each plot were used to record data viz., days to maturity, plant height at maturity (cm), number of branches per plant, number of filled pods per plant, number of unfilled pods per plant, 100-seed weight (g), oil content (%), grain yield per plant (g), biological yield per plant (g) and harvest index. Average values of each trait were analysed for multivariate analysis (cluster analysis and principal component analysis) by using the SPSS-12 and STATISTICA-6 programme as described by Sneath & Sokal (1973).

Results and Discussion

Cluster analysis: Cluster diagram using Ward's method based on agro-morphological traits of 139 soybean accessions proposed two major groups A and B. The group A consisted of one cluster, whereas group B constitutes four clusters (Fig. 1). The grouping of different accessions according to clusters for the quantitative traits is presented in Table 1, whereas mean and standard deviation for various plant characters in each cluster are presented in Table 2. Cluster I comprised of 32 accessions, which represented 23.02% of the total accessions (Table 1). It has been observed that characters viz., more number of filled pods per plant (88.28), maximum 100-seed weight (14.69g), highest grain yield per plant (26.20g), maximum biological yield per plant (65.28 g) and higher harvest index (40.25%) were found in cluster 1 (Table 2). The cluster II accounts for 8.63% of the population and includes 12 accessions. This cluster shows least oil content (17.4%), grain yield per plant (7.2g) and harvest index (21.6%). Cluster III represents 15.82% of the total 139 accessions and comprised 22 accessions. The accessions from this cluster could be picked up for maximum plant height (76.3cm) and higher number of branches per plant (5.3) while, second highest in grain yield per plant (18.1g) and medium 100-seed weight (13.1g) was also found in this cluster. The cluster IV contributed 17.25% and comprised of 24 accessions. This cluster comprised of late maturity (105.5), higher oil content (22.9%) and less number of unfilled pods per plant (2.5) accessions. Cluster V contributed 35.25% to the population and comprising of maximum (49) accessions. These accessions showed early maturity (97.1) and second highest oil content (22.3%).

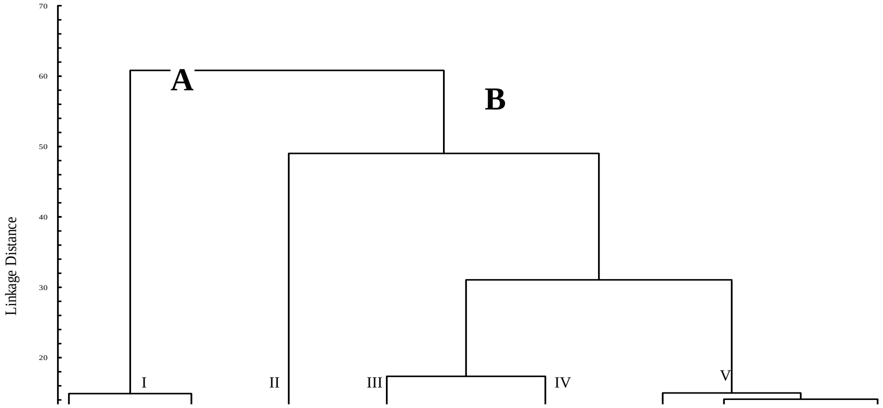

Fig. 1. Phenogram based on ten quantitative traits of 139 soybean accession.

Table 1. Grouping of soybean accessions based on five clusters for various traits.

Cluster	*f*	% Age	Cluster membership
Cluster I	32	23.02	03128, 03138, 03089, 03140, 03077, 03095, 03061, 03096, 03060, 03079, 03118, 03059, 03094, 03130, 03119, 03074, 03041, 03086, AJMERI, 03058, 03137,03036, 03122, 03035, 03092, 03082, 03078, 03056, 03084, 03076, 03080, 03034
Cluster II	12	8.63	03124, 03087, 03121, 03022, 03023, 03019, 03018, 03017, 03028, 03016, 03015, 03011
Cluster III	22	15.82	03141, 03091, 03083, 03052, 03039, 03111, 03127, 03072, 03115, 03126, 03063, 03038, 03067, 03053, 03136, 03116, 03129, 03037, 03125, 03031, 03145, 03030
Cluster IV	24	17.25	03051, 03090, 03048, 03108, 03106, 03045, 03117, 03033, 03026, 03105, 03104, 03075, 03102, 03071, 03142, 03133, 03134, 03064, 03046, 03110, 03085, 03044, 03040, 03013
Cluster V	49	35.25	03143, 03073, 6057, 03131, 03098, 03054, 03135, 03112, 03107, 03093, 03101, 03050, 03042, 03024, 03021, 03088, 03014, 03068, 03032, 03020, 03012, NARC-II, 03081, 03099, 03103, 03070, 03114, 03113, 03120, 03097, NARC-I, 03065, 03100, 03062, 03109, 03047, 03043, 03055, 03027, 03069, 03029, 03132, 03025, 03139, 03066, 03049, 03144, 03123, 03010

f- Frequency, % Age- percentage

Table 2. Mean values and standard deviation for five clusters based on ten quantitative traits of 139 soybean accessions.

	Cluster I	Cluster II	Cluster III	Cluster IV	Cluster V
Traits	**Mean ± SD**	**Mean ± SD**	**Mean ± SD**	**Mean ± SD**	**Mean ± SD**
Days to maturity	104.13 ± 7.1	102.3 ± 5.5	103.1 ± 3.5	105.5 ± 3.8	97.1 ± 6.3
Plant height	63.51 ± 17.0	61.4 ± 9.8	76.3 ± 18.8	51.8 ± 10.9	57.4 ± 14.8
Number of branches per plant	4.59 ± 1.1	4.7 ± 1.0	5.3 ± 0.8	4.5 ± 1.0	3.3 ± 1.2
Number of filled pods per plant	88.28 ± 16.2	40.4 ± 19.2	65.1 ± 8.4	58.1 ± 8.8	47.3 ± 11.3
Number of unfilled pods per plant	4.48 ± 2.1	3.7 ± 1.9	3.8 ± 1.2	2.5 ± 1.4	3.1 ± 1.1
100-seed weight	14.69 ± 1.7	8.5 ± 1.1	13.1 ± 1.3	14.4 ± 1.5	13.9 ± 2.1
Oil contents %	21.14 ± 1.4	17.4 ± 1.3	21.5 ± 1.4	22.9 ± 1.0	22.3 ± 1.5
Grain Yield per plant per plant	26.20 ± 3.5	7.2 ± 2.7	18.1± 2.5	17.9 ± 2.5	13.5 ± 3.1
Biological yield per plant per plant	65.28 ± 8.4	32.4 ± 7.9	51.3 ± 9.7	46.2 ± 6.6	37.4 ± 6.8
Harvest index	40.25 ± 3.4	21.6 ± 3.5	35.6 ± 3.3	38.9 ± 4.7	36.0 ± 5.4

SD-standard deviation

Table 3. Principal components (PCs) analysis of soybean accessions for ten quantitative traits.

		PC_1	PC_2	PC_3
Eigenvalues		3.51	2.37	1.08
Proportion of variance		35.15	23.77	10.85
Cumulative variance		35.15	58.92	69.77
Traits	**Communalities**	**Eigenvector**		
Days to maturity	0.70	0.23	-0.36	0.72
Plant height (cm)	0.42	0.17	-0.47	-0.41
No. of branches per plant	0.62	0.29	-0.54	0.50
No. of filled pods per plant	0.82	0.88	-0.19	-0.06
No. of unfilled pods per plant	0.54	0.36	-0.54	-0.34
100 seed weight (g)	0.67	0.45	0.69	0.06
Oil contents (%)	0.63	0.19	0.76	0.13
Grain yield per plant (g)	0.95	0.97	0.05	-0.05
Biological yield per plant (g)	0.85	0.90	-0.20	-0.07
Harvest index (%)	0.74	0.66	0.56	-0.02

Though cluster analysis grouped together having accessions with greater morphological similarity, thus representative accessions from a cluster of particular group could be chosen for hybridization programme. Ghatge & Kadu (1993) evaluated 58 soybean genotypes from diverse eco-geographical areas. They observed seven clusters on the basis of yield components. Kumar & Nadarajan (1994) studied genetic diversity for yield components in 64 soybean genotypes revealed 11 clusters. Mehetre *et al*., (1994) evaluated 51 soybean genotypes using D2 analysis and they found 10 clusters. Das *et al*., (2001) reported that grouping pattern of the diverse genotypes suggested no parallelism between genetic divergence and geographical distribution of the genotypes. Some potentially important traits have been identified and could be exploited for specific trait improvement and assemblage of core collection from a bulk genetic stock. Ihsan Ullah *et al*., (2007) reported 5 clusters derived from 10 genotypes on the basis of seed yield in sunflower.

Principal component analysis: The data revealed that 3 principal components having greater than one eigenvalues contributed 69.77% of the total variation among 139 accessions of soybean (Table 3). It was found that principal component 1 (PC1) contributed 35.15%, whereas PC2 and PC3 contributed 23.77% and 10.85%, respectively of the total variation. The traits, which contributed more positively to PC1, were grain yield per plant (0.97), biological yield per plant (0.90), number of filled pods per plant (0.88) and harvest index (0.66), while the rest of traits under the present study in PC1 showed low eigenvector values. This indicated that the populations with greater PC1 values are high yielding and formed by having more grain yield per plant, biological yield per plant, number of filled pods per plant and harvest index. Maximum genetic variance to PC2 was contributed by oil contents (0.76), 100-seed weight (0.69) and harvest index (0.56), while days to maturity, plant height, number of branches per plant, number of filled pods per plant, number of unfilled pods per plant and biological yield per plant had negative value. The second component was strongly associated with oil content, 100-seed weight and harvest index. The PC3 showed that days to maturity (0.72) and number of branches per plant (0.50) contributed more positively. Hundred seed weight and oil content contributed positively but low values. The rest of characters in this component have negative association with low to moderate value. It is evident that quantitative traits contributed positively to first three principal components and hence these could be given considerable importance for the genetic material under investigation. First three components which contributed 69.77 of the total variance were plotted graphically as scattered diagrams to observe the relationship between clusters. When PC1 was plotted against PC2, cluster 1 and 2 only showed complete separation and there was mixing up of clusters 3, 4 and 5 (Fig. 2).

Broschat (1979) considered PCA as powerful technique for data reduction which removes interrelationships among components. Results reported by various researchers showed multivariate analysis as a valid system to deal with germplasm collection. Smith *et al.*, (1995) conducted average linkage cluster and principal component analyses, and reported the utility of these results in preservation and utilization of germplasm. Rabbani *et al.*, (1998) determined the extent of diversity and relationship among *Brassica juncea* germplasm from Pakistan for 35 morphological characters in 52 accessions using cluster and principal component analysis. Ghafoor *et al.*, (2001) studied genetic diversity in blackgram germplasm accessions. Quantitative traits were analysed for cluster and principal component analysis. The first four principal components with eigenvalues >1 contributed 79.5% of the total variability among accessions. Elizabeth *et al.*, (2001) investigated 19 sesbania accessions to characterize them on morphological and agronomic data using multivariate method. Principal component analyses indicated that variance accumulated by the first two components for morphological and agronomic data was 74.4 and 77.0%, respectively. Ghafoor *et al.*, (2003) evaluated chickpea accessions by using multivariate techniques. The first three principal components with eigenvalues >1 contributed 83.3% of the variability amongst genotypes.

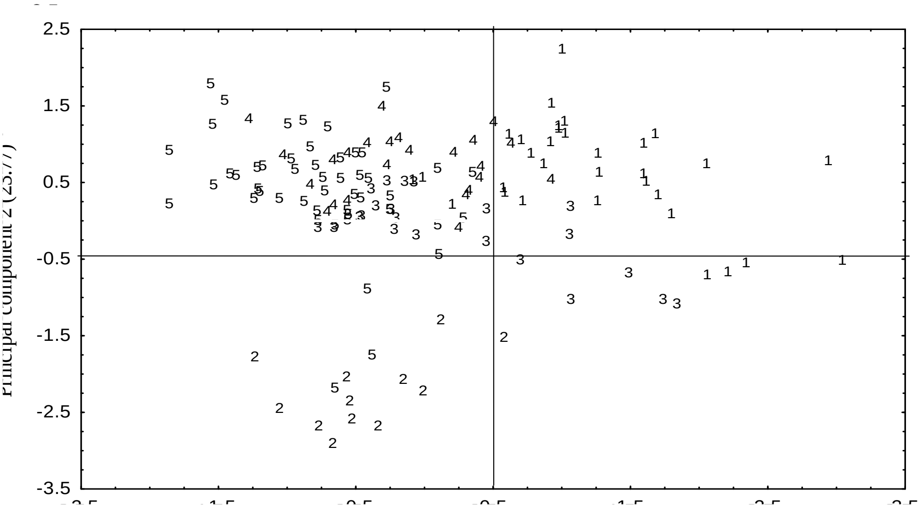

Fig. 2. Scatter diagram on average cluster diversity for first two PCs of 139 soybean accession.

References

Anonymous. 2005-2006. *Agricultural Statistics of Pakistan.* Ministry of Food, Agriculture and Live stock (Economic Wing), Islamabad. pp. 107

Antalina, S. 2000. Modern processing and utilization of legumes. Recent Research and Industrial achievement for soybean food in Japan *Proceeding of RILET-JIRCAS*. Workshop on Soybean Research. September 28, Malang-Indonesia.

Balasubramaniyan, P. and S.P. Palaniappan. 2003. *Principles and practices of agronomy*. Pub. By agrbios (India). pp. 45-46. Chapter 1. Field crops. An overview.

Broschat, T.K. 1979. Principal component analysis in horticultural research. *Hort. Sci.*, 14: 114-117.

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

Das, S.P., P.N. Harer and A.B. Biradar. 2001. Genetic divergence and selection of genotypes in soybean. *J. Mah. Agric. Univ*., 25(3): 250-252.

Elizabeth, A.V., E.A. Schammass, R. Vencovsky, P.S. Martins and G. Bandel. 2001. Germplasm characterization of *Sesbania* accessions based on multivariate analyses. *Genet. Resour. Crop Evol*., 48: 79-90.

Fageria, N.K., V.C. Baligar and C.A. Jones. 1997. Growth and mineral nutrition of field crops. 2nd Ed. Marcel Dekker, Inc., New York.

Federer, W.T. and D. Ragavarao. 1975. On augmented designs. *Biometrics*, 31: 29-35.

Ghafoor, A., A. Sharif, Z. Ahmad, M.A. Zahid and M.A. Rabbani. 2001. Genetic diversity in blackgram (*Vigna mungo* (L.) Hepper). *Field Crops Res*., 69: 183-190.

Ghafoor, A., F.N. Gulbaaz, M. Afzal, M. Ashraf and M. Arshad. 2003. Inter-relationship between SDS-PAGE markers and agronomic traits in chickpea (*Cicer arietinum* L.). *Pak. J. Bot*., 35(4): 613-624.

Ghatge, R.D. and R.N. Kadu. 1993. Genetic diversity in soybean. *Ann. Agric. Res.*, 14(2): 143-148.

Ihsan Ullah, M. Ayub, M.R. Khan, M. Ashraf, M.Y. Mirza and M. Yousaf. 2007. Garphical analysis of multi-environment trial (MET) data in sunflower (*Helianthus annuus* L.) through clustering and GGE biplot technique. *Pak. J. Bot*., 39(5): 1639-1646.

Johns, M.A., P.W. Skrotch, J. Neinhuis, P. Hinrichsen, G. Bascur and C. Munoz-Schick. 1997. Gene pool classification of common bean landraces from Chile based on RAPD and morphological data. *Crop Sci*., 37: 605-613.

Kumar, M. and N. Nadarajan. 1994. Genetic divergence studies in soybean (*Glycine max* (L.) Merrill). *Indian J. Genet. Plant Breed*., 54(3): 242-246.

Lackey, J.A. 1977. *A synopsis of Phaseoleae* (*Leguminosae, Papilionoideae*). Ph.D. dissertation. Iowa State University. Ames, Iowa.

Mehetre, S.S., C.R. Mahajan, P.A. Patil and D.N. Hajare. 1994. Genetic divergence in soybean (*Glycine max* (L.) Merrill.). *Indian J. Genet. Plant Breed*., 54(1): 83-88.

Melchinger, A.E. 1993. Use of RFLP markers for analyses of genetic relationships among breeding materials and prediction of hybrid. *International Crop Science Congress*, 1st, Ames, IA. July 1992. CSSA, Madison, WI.

Rabbani, M.A., A. Iwabuchi, Y. Murakami, T. Suzuki and K. Takayanagi. 1998. Phenotypic variation and the relationship among mustard (*Brassica juncea* L.) germplasm from Pakistan. *Euphytica*, 101: 357-366.

Smith, S.E., L. Guarino, A.A. Doss and D.M. Conta. 1995. Morphological and agronomic affinities among Middle Eastern alfalfas accessions from Oman and Yemen. *Crop Sci*., 35: 1188-1194.

Sneath, P.H.A. and R.R. Sokal. 1973. *Numerical taxonomy: the principal and practices of numerical classification*. W.F. Freeman & Co., San Francisco, p. 573.

Thompson, J.A., R.L. Nelson and L.O. Vodkin. 1998. Identification of diverse soybean germplasm using RAPD markers. *Crop Sci*., 38: 1348-1355.

(Received for publication 12 February 2008)