INVESTIGATION AND COMPARISON OF SOME MORPHOLOGICAL TRAITS OF THE SOYBEAN POPULATIONS USING CLUSTER ANALYSIS

M. FAISAL ANWAR MALIK[1], MUHAMMAD ASHRAF[2], AFSARI S. QURESHI[1] AND MUHAMMAD RASHID KHAN[1]

[1]*Department of Biochemistry, Quaid-i-Azam University, Islamabad.*
[2]*National Agricultural Research Centre, Islamabad,.*

Abstract

A study was performed on 92 genotypes of soybean including 4 checks during the Kharif season (July-Oct) of 2003. High values of the Shannon index for flower colour (0.54), pod colour (1.12), plant type (0.88), seed luster (0.84), seed size (1.08) and seed colour (0.91) revealed greater variability in these qualitative traits. High CVs were recorded in leaf area (44.81%), pods plant[-1] (29.47%), branches plant[-1] (31.72%), 100-seed weight (39.01%) and grain yield plant[-1] (46.55%) with a wide range of 28-146, 26-130, 2.4-11, 4.2-21.5 and 4.04-28.23 respectively indicating a high level of diversity among the accessions for these traits. Grain yield was positively and highly significantly correlated with leaf area, plant height, pods plant[-1], branches plant[-1] and 100-seed weight suggesting thereby that phenotypic selection could be made on the basis of these characters. Means of the clusters showed that the accessions in cluster III were not only late maturing and high yielding but also had more pods plant[-1], branches plant[-1] and 100-seed weight. However, accessions in cluster I showed some promise to earliness with high grain yield and oil. The physical distinctness, on the basis of cluster analysis, between Pakistani and US/AVRDC accessions in this study reflect that the introgression of US and AVRDC accessions to Pakistani breeding programme should broaden Pakistani soybean germplasm diversity. Whereas no clear relationship was found between genetic diversity and geographical origins because accessions from one origin entered into more than one cluster.

Introduction

The soybean (*Glycine max*) is often called the miracle crop. It is the world's foremost provider of protein and oil. The bushy, green soybean plant is a legume related to clover, peas and alfalfa. Genetic variation among traits is important for breeding and in selecting desirable types. 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 (Bretting & Widrlechner, 1995). Introgression of new genetic diversity through hybridization with introduced germplasm is one way to increase genetic variation in breeding populations, the base upon which gain from selection depends (Guedira *et al*., 2000). Sandoval *et al*., (1997) have shown that genetic variation for seed yield and other agronomic traits in soybean populations derived from crosses among elite lines can be limited by a lack of genetic diversity. Das *et al*., (2001) showed that pods plant[-1] and plant height had the maximum contribution towards divergence. Karad *et al*., (2005) assessed high variability for yield plant[-1], plant height and pods plant[-1], and moderate for seeds pod[-1] and branches plant[-1], indicating that these traits have scope for improvement through selection. On the other hand, an analysis of the correlation between seed yield and yield components is essential in determining selection criteria. Khan *et al*., (2000) observed that seed yield was positively and significantly correlated with all

characters except pod height. Amanullah & Hatam (2001) showed significantly positive correlation of grain yield with plant height (0.530), days to maturity, branches (0.541) and pods plant[-1] (0.712). However, the correlation between grain yield plant[-1] and 100-seed weigh was significantly negative (-0.731). Pods plant[-1] had the highest contribution to the seed yield and can be considered as most reliable yield indicator in soybean (Mukhekar *et al.*, 2004). Peluzio *et al.*, (2005) showed that number of pods plant[-1], days to flower and days to maturity were important for the indirect improvement of grain production. Mehetre *et al.,* (1994), Mehetre *et al.*, (1998) and Sihag *et al.*, (2004) found that genetic diversity and geographic distribution were independent of each other and no definite relationship existed between genetic diversity and geographic origin.

The objective of this research was to assess the relationship of some morphological traits in 92 genotypes of soybean that are potential new sources of genetic variation for soybean breeding programs in Pakistan.

Materials and Methods

A study was performed on 92 genotypes of soybean including 4 checks obtained from Oilseed Department and Institute of Agro-biotechnology and Genetic Resources of National Agricultural Research Centre (NARC), Islamabad consisting of 5 different origins/sources (Table 1). The experiment was conducted in the Research Area of the NARC during the Kharif season (July-Oct) of 2003, with single repeat. Each genotype was sown in three rows of 5 m in length, with 60 cm row spacing. The collection was characterized assessing flower colour, pod colour, plant type, seed luster, seed size, seed colour, leaf area, days to flowering initiation, days to flowering completion, days to maturity, plant height (cm), branches plant[-1], pods plant[-1], 100 seed-weight (g), grain yield plant[-1] (g), and oil content (%). Leaf area was noted by leaf area meter. Oil contents (%) were recorded with the help of NMR spectroscopy. The genetic diversity (Shannon) index of qualitative traits (Tang, 1997) and basic statistics and CV (Dong, 1998) were calculated using computer programme MS-Excel. Correlation coefficients and cluster analysis were analyzed with the Statistica programme as described by Sneath & Sokal (1973).

Results and Discussion

The germplasm exhibited high variability for all the 6 qualitative traits. High values of the Shannon index (Table 2) were observed for flower colour (0.54), pod colour (1.12), plant type (0.88), seed luster (0.84), seed size (1.08) and seed colour (0.91). The majority of the genotypes had purple flower (77%) and brown pod colour (59%). Out of 7 types of seed colour, genotypes having yellow seed colour (72.83%) dominated the germplasm. Dong *et al.,* (2001) also showed high variability in qualitative traits using Shannon index. Means, CV's and ranges for 10 characters studied in the soybean accessions are presented in Table 3. High CV's were recorded in leaf area (44.81%), pods plant[-1] (29.47%), branches plant[-1] (31.72%), 100-seed weight (39.01%) and grain yield plant[-1] (46.55%) alongwith a wide range of 28-146 cm², 26-130, 2.4-11, 4.2-21.5 g and 4.04-28.23 g respectively revealing a high level of diversity among the accessions for these traits. Selection on the basis of these traits can be useful. Basavaraja *et al.,* (2005) also observed that seed yield varied from 4983 to 2033 kg ha[-1]. Small CV's were observed for the other characters indicating lesser scope of selection. Sihag *et al.*, (2004) & Chettri *et al.*, (2005) studied the genetic diversity among soybean genotypes and observed a wide genetic diversity among the genotypes. Similar results were also presented by Muhammad & Shah (2003) & Malik *et al.*, (2006).

Table 1. Genotypes showing their origin/source.

S. No.	Origin/Source	Genotypes
1.	Pakistan	30
2.	USA	30
3.	Japan	08
4.	North Korea	04
5.	AVRDC	20

Table 2. Frequency distribution and Shannon index of six qualitative traits.

S. No.	Plant character	Category	Frequency	%age	Shannon index
1.	Flower colour	Purple	71	77	0.54
		White	21	23	
2.	Pod colour	Blackish brown	12	13	1.12
		Brown	54	59	
		Dark brown	16	17	
		Light brown	10	11	
3.	Seed colour	Black	2	2.17	0.91
		Brown	16	17.39	
		Dark brown	1	1.09	
		Green	2	2.17	
		Light brown	3	3.26	
		Reddish brown	1	1.09	
		Yellow	67	72.83	
4.	Seed size	Small	30	32.61	1.08
		Medium	38	41.3	
		Large	24	26.09	
5.	Seed lustre	Dull	4	4.35	0.84
		Intermediate	40	43.48	
		Shiny	48	52.17	
6.	Plant type	Errect	60	65.22	0.88
		Semi erect	13	14.13	
		Spreading	19	20.65	

Table 3. Means, range, CV% of ten quantitative traits studied in ninety-two genotypes of soybean (*Glycine max*).

S. No.	Character	Means	Range	CV%
1.	Leaf area (cm^2)	77.46	28-146	44.81
2.	Days to flowering initiation	44.86	28-54	12.12
3.	Days to flowering completion	67.55	47-79	9.70
4.	Days to maturity	101.18	85-117	6.60
5.	Plant height (cm)	68.05	41.4-106.4	17.46
6.	Pods plant^{-1}	84.78	26-130	29.47
7.	Branches plant^{-1}	5.733	2.4-11	31.72
8.	100-seed weight (g)	11.46	4.2-19.5	39.01
9.	Grain yield plant^{-1} (g)	12.86	4.04-28.23	46.55
10.	Oil content (%)	19.07	13.8-23.45	14.55

Positive and significant correlations were observed between days to flowering initiation and days to flowering completion and days to flowering initiation and days to maturity. These results are well confirmed with the findings of Chamundeswari & Aher (2003). There was a trend to increase pods plant^{-1} and branches plant^{-1} with increased plant height as these were also positively correlated with each other. Malik *et al*., (2006) also revealed that pods plant^{-1} and branches plant^{-1} can be increased by selecting tall plants. Grain yield plant^{-1} was positively and highly significantly correlated with leaf area, plant height, pods plant^{-1}, branches plant^{-1} and 100-seed weight suggesting thereby that phenotypic selection could be made on the basis of these characters. Similar results were reported by Mukhekar *et al*., (2004), Chandel *et al*., (2005) & Turkec (2005). While it showed negative and significant correlation with days to flowering initiation, days to flowering completion and days to maturity. These results revealed that longer duration in soybean is not desirable. Malik *et al.,* (2007) also showed negative association of these characters with yield.

Fig. 1 showed that the UPGMA procedure defined 3 clusters (6 sub-clusters). Cluster I consisted of only one accession, cluster II of 47 and cluster III of 50 accessions. The control cultivars (NARC-I, NARC-III, NARC-IV and NARC-V) were included in cluster III (sub cluster IV). The analysis was able to identify related groups of accessions and many of the soybean plant introductions included in the study were found to be genetically distinct from the accessions of Pakistani origin. For example, Clusters II was composed mainly of plant introductions of American and AVRDC origin, while clusters III consisted of most of the Pakistani and all North Korean accessions. However, one Japanese accession formed the cluster I independently while rest of the accessions of Japanese origin were spread in all clusters. The physical distinctness between Pakistani and US/AVRDC accessions in this study reflect that the introgression of US and AVRDC accessions to Pakistani breeding programme should broaden Pakistani soybean germplasm diversity. Whereas no clear relationship was found between genetic diversity and geographical origins because accessions from one origin entered into more than one cluster. Similarly, accessions from different geographical origins tended to be clustered in one part of the dendrogram. Das *et al.,* (2001) revealed that grouping pattern of the accessions suggested no parallelism between genetic divergence and geographical distribution of the accessions. Genetic patterns obtained from this study of soybean germplasm can help soybean breeders make better choices when selecting among the large numbers of accessions available. Sihag *et al*., (2004) observed that clustering pattern revealed no definite relationship between genetic diversity and geographic diversity. The genotypes from the same eco-geographic region were classified in different clusters, and genotypes from different eco-geographic regions were classified into one cluster.

Means of the clusters presented in Table 3 showed that the accessions in cluster III were not only late maturing and high yielding but also had more pods plant^{-1}, branches plant^{-1} and 100-seed weight. However, accessions in cluster I showed some promise to earliness with high grain yield and oil content. Das *et al*., (2001) also showed that pods plant^{-1} and plant height had the maximum contribution towards divergence. The information on the clusters to which particular accessions with traits of interest belong will assist in researching more accessions with similar traits (Li & Nelson, 2001).
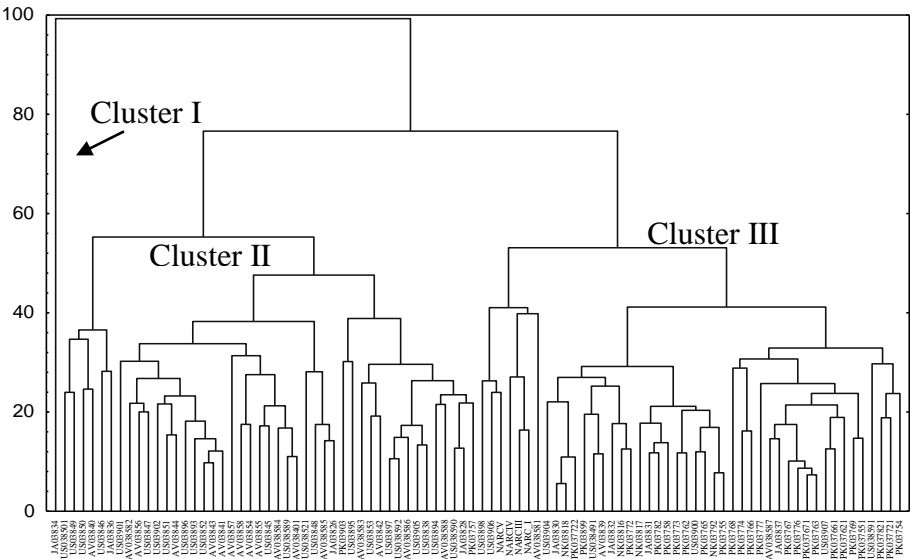
Fig. 1. Dendrogram based on agronomic data of soybean germplasm during 2003 by UPGMA method.

Table 4. Correlation coefficients of ten quantitative traits studied in ninety-two genotypes of soybean (*Glycine max*).

	LA	DFI	DFC	DMA	PH	PP	BP	100-SW	GY
DFI	0.04								
DFC	-0.01	0.85**							
DMA	-0.04	0.66**	0.72**						
PH	0.19	-0.02	0.04	0.15					
PP	0.40**	-0.08	-0.09	-0.05	0.27**				
BP	0.35**	0.30**	0.16	0.21*	0.05	0.30**			
100-SW	0.65**	-0.35**	-0.29**	-0.24*	0.23*	0.45**	0.11		
GY	0.56**	-0.25**	-0.21*	-0.07	0.32**	0.73**	0.34**	0.69**	
OC	0.38**	-0.37**	-0.31**	-0.21*	0.23*	0.38**	-0.03	0.70**	0.53**

* Significant at 5%, ** Significant at 1%, LA = Leaf area (cm^2), DFI = Days to flowering initiation, DFC = Days to flowering completion, DMA = Days to maturity, PH = Plant height (cm), PP = Pods plant^{-1}, BP = Branches plant^{-1},100-SW = 100-seed weight (g), GY = Grain yield plant^{-1} (g), OC = Oil content (%)

Table 5. Cluster means of ten quantitative traits studied in ninety-two genotypes of soybean (*Glycine max*).

S. No.	Character	Cluster I	Cluster II	Cluster III	Cluster IV
1.	Leaf area (cm^2)	97.50	49.20	118.96	45.63
2.	Days to flowering initiation	43.18	46.25	45.81	44.33
3.	Days to flowering completion	66.23	69.00	67.96	67.20
4.	Days to maturity	95.27	102.40	101.08	101.13
5.	Plant height (cm)	72.80	61.13	70.53	67.03
6.	Pods plant^{-1}	75.64	52.09	111.20	90.40
7.	Branches plant^{-1}	5.21	5.12	7.05	5.38
8.	100-seed weight (g)	13.95	7.42	15.18	9.12
9.	Grain yield plant^{-1} (g)	13.31	6.07	18.83	11.90
10.	Oil content (%)	20.54	17.12	20.39	18.15

Conclusion

Our results showed that Shannon index for flower colour, pod colour, plant type, seed luster, seed size and seed colour was high revealing greater variability in these traits. High level of diversity among the accessions was recorded in leaf area, pods plant[-1], branches plant[-1], 100-seed weight and grain yield plant[-1]. Phenotypic selection could be made on the basis of leaf area, plant height, pods plant[-1], branches plant[-1] and 100-seed weight as grain yield was positively and highly significantly correlated with these characters. Accessions showing some promise to earliness with high grain yield and oil content were identified. The early maturing genotypes provide an opportunity to avoid terminal drought and produce acceptable yield. The physical distinctness, on the basis of cluster analysis, between Pakistani and US/AVRDC accessions in this study reflects that the introgression of US and AVRDC accessions to Pakistani breeding programme should broaden Pakistani soybean germplasm diversity. Whereas no clear relationship was found between genetic diversity and geographical origins because accessions from one origin entered into more than one cluster.

References

Amanullah and M. Hatam. 2001. Performance of AVRDC vegetable soybean germplasm under Peshawar valley conditions [Pakistan]. *Sarhad J. Agric.*, 17(1): 27-30.

Basavaraja, G.T., G.K. Naidu and P.M. Salimath. 2005. Evaluation of vegetable soybean genotypes for yield and component traits. *Karnataka J. Agric. Sci.*, 18(1): 27-31.

Bretting, P.K. and M.P. Widrlechner. 1995. Genetic markers and plant genetic resource management. In: *Plant Breeding Reviews*. (Ed.): J. Janick. John Wiley & Sons, Inc., New York. Vol. 13. pp. 11-87.

Chamundeswari, N. and R. Aher. 2003. Character association and component analysis in soybean [*Glycine max* (L.) Merrill]. *Ann. Biol.*, 19(2): 199-203.

Chandel, K., O.P. Sood, V.K. Sood and S.L. Gartan. 2005. Character associations and path analysis studies in cultivated soybean (*Glycine max*) genotypes restructured through introgression of *G. soja* chromatin. *Crop Res. Hisar.*, 29(3): 469-473.

Chettri, M., S. Mondal and R. Nath. 2005. Studies on genetic variability in soybean (*Glycine max* (L.) Merril) in the mid hills of Darjeeling District. *J. Interacademicia*, 9(2): 175-178.

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

Dong, Y.S. 1998. Exploration on genetic diversity center for cultivated soybean in China. *Chin Crops J.*, 1: 18-19.

Dong, Y.S., B.C. Zhuang, L.M. Zhao, H.Sun and M.Y. He. 2001. The genetic diversity of annual wild soybeans grown in China. *Theor. Appl Genet.*, 103: 98-103.

Guedira, G.L.B., 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.

Karad, S.R., P.N. Harer, D.D. Kadam and R.B. Shinde. 2005. Genotypic and phenotypic variability in soybean (*Glycine max* (L.) Merrill). *J. Maharashtra Agric. Univ.*, 30(3): 365-367.

Khan, A., Hatam, M. and A. Khan. 2000. Heritability and interrelationship among yield determining components of soybean varieties. *Pak. J. Agric. Res.*, 16(1): 5-8.

Li, Z. and R.L. Nelson. 2001. Genetic diversity among soybean accessions from three countries measured by RAPDs. *Crop Sci.*, 41(4): 1337-1347.

Malik, M.F.A., A.S. Qureshi, M. Ashraf and A. Ghafoor. 2006. Genetic variability of the main yield related characters in soybean. *Int. J. Agric. Biol.*, 8(06): 815-819.

Malik, M.F.A., M. Ashraf, A.S. Qureshi and A. Ghafoor. 2007. Assessment of genetic variability, correlation and path analyses for yield and its components in soybean. *Pak. J. Bot.*, 39(2): 405-413.

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

Mehetre, S.S., R.B. Shinde, U.M. Borle and P.P. Surana. 1998. Genetic divergence studies of assimilate partitioning in roots, growth and yield characters in soybean (*Glycine max* (L.) Merril). *Adv. Plant Sci.,* 11(1): 205-208.

Muhammad, A. and P. Shah. 2003. Comparative yield and yield components of land races and improved varieties of soybean. *Sarhad J. Agric.,* 19(4): 459-461.

Mukhekar, G.D., N.D. Bangar and D.B. Lad. 2004. Character association and path coefficient analysis in soybean (*Glycine max* (L.) Merill). *J. Maharashtra Agric. Univ.,* 29(3): 256-258.

Peluzio, J.M., R.D. de-Almeida, R.R. Fidelis, D. Almeida-Jr, E.L. Brito and E.R. Francisco. 2005. Correlations among traits of soybean, in Gurupi, Tocantins. *Rev. Ceres.,* 52(303): 779-786.

Sandoval, M.P., T.E. Carter-Jr, D.M. Webb and J.W. Burton. 1997. RFLP genetic similarity and coefficient of parentage as genetic variance predictors for soybean yield. *Crop Sci.,* 37:698-703.

Sihag, R., J.S. Hooda, R.D. Vashishtha and B.P.S. Malik. 2004. Genetic divergence in soybean [*Glycine max* (L.) Merrill]. *Ann. Biol.,* 20(1): 17-21.

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

Tang, Q.Y. 1997. *Data processing system*. Chinese Agric. Press, Beijing.

Turkec, A. 2005. Correlation and path analysis of yield components in soybean varieties. *Turk. J. Field Crops*, 10(1): 43-48.