

Principal Component Analysis of Nine Morphological Traits in Forty Five Accessions of Soybean [*Glycine max* (L.) Merrill]

Manav*, R. N. Arora and Krishan Kumar

Department of Genetics and Plant Breeding, College of Agriculture, CCS
Haryana Agricultural University, Hisar-125 004, India

*Corresponding Author E-mail: manavhisar@gmail.com

Received: 15.11.2017 | Revised: 22.12.2017 | Accepted: 26.12.2017

ABSTRACT

An experiment was carried out using forty five accessions of soybean laid in a randomized block design with three replications at the research farm of CCSHAU, Hisar in 2014 to determine the contribution of nine morphological traits to the total variability in soybean using Principal component analysis. Results revealed that the first 3 principal components accounted for 79.37% of the total variation. PC₁ contributed 52.85% of the total variation and dominated by 100-seed weight, number of days to maturity, number of branches per plant, pods per plant, seed yield per plant and plant height. The second principal component explained an additional 14.94% of the total variation dominated by highly positive pod length and number of seeds per pod while the PC₃ explained an additional 11.58% of the total variation and only number of days to 50% flowering showed positive relation with the third component (PC₃) signifying its importance in divergence.

Key words: Principal Component Analysis, Soybean.

INTRODUCTION

Soybean [*Glycine max* (L.) Merrill] is the most widely grown leguminous crop in the world and is an important source of protein and oil for food and feed^{8,14}. Soybean seeds are rich in proteins, unsaturated fatty acids, minerals (e.g. Ca and P) and vitamins (e.g. A, B, C and D), which meet the nutritional needs of humans and other animals^{6,7}. Soybean seed is referred to the “protein hope” of the future, because of its high nutritive value, containing about 42-45% protein¹². Yield and yield contributing parameters are the most widely targeted traits

for soybean improvement programme worldwide. In view of achieving higher production and productivity, genetic improvement through restructuring of plant, both morphologically and physiologically, is needed for developing high-yielding varieties and bridging the gap in the declining production. The success of breeding programme in any crop for economic characters often depends on the availability of a large germplasm representing a diverse genetic variation.

Cite this article: Manav, Arora, R.N. and Kumar, K., Principal Component Analysis of Nine Morphological Traits in Forty Five Accessions of Soybean [*Glycine max* (L.) Merrill], *Int. J. Pure App. Biosci.* 5(6): 938-941 (2017). doi: <http://dx.doi.org/10.18782/2320-7051.6138>

In order to ensure the efficient and effective use of crop germplasm, its characterization is imperative and multivariate analysis provides a good evaluation of landraces by identifying those that should be further evaluated at the genetic level¹¹.

For the choice of diverse parents in any hybridization programme, multivariate analysis (Principal component analysis) has been extensively used. Dasgupta and Das² considered multivariate analysis best for choosing parents for hybridization. It involves a mathematical procedure that transforms a number of (possibly) correlated variables into a (smaller) number of uncorrelated variables called principal components¹. The first principal component accounts for as much of the variability in the data as possible and each succeeding component accounts for as much of the remaining variability as possible. The objectives of PCA are to discover or to reduce the dimensionality of the data set and to identify new meaningful underlying variables⁹. In the present study, we carried out a PCA to identify agronomic attributes whose selection would lead to improvement in seed yield of soybean genotypes.

MATERIALS AND METHODS

The experiment was carried out at the experimental area of Pulses Section, Department of Genetics and Plant breeding, CCSHAU, Hisar during kharif 2014. Forty five accessions including 3 checks (PS 688, PS 1347 and PK 416) were evaluated for some important agronomic traits. The experiment was planted in randomized block design having three replications in single row of 3m length, keeping row to row distance of 45 cm and plant to plant distance of 10 cm. Five plants randomly selected from each plot were used to record data for 9 different morphological characters viz., days to 50% flowering, days to maturity, plant height, number of branches per plant, number of pods per plant, number of seeds per pod, pod length, 100-seed weight (g) and seed yield per plant (g). The principal component analysis was performed by using the computer software "SPSS" for windows.

RESULTS AND DISCUSSION

Principal component analysis is useful technique as it gave information about the groups where certain traits are more important allowing the breeders to conduct specific breeding programs¹³. In the present study, the 3 principal components with eigen values more than one (eigen values 4.76, 1.34 and 1.04 respectively Table 1) contributed 79.37% of the total variation amongst 45 genotypes evaluated for 9 morphological traits in soybean (Table1). These eigen values were greater than one and represented exact linear dependency, but the rest PCs (4-9) had eigen values <1. The principal components (eigen value greater than one), eigen values (Root), per cent variability, cumulative per cent variability and component loading of different characters are presented in Table 1. It was therefore inferred that the essential features of data set had been represented in the first three principal components. It was found that PC₁ contributed 52.85% of the total variation and dominated by 100-seed weight, number of days to maturity, number of branches per plant, pods per plant, seed yield per plant and plant height suggesting that this component reflected the yield potential of each genotype through some yield component aspects. The second principal component explained an additional 14.94% of the total variation dominated by highly positive pod length and number of seeds per pod while the PC₃ explained an additional 11.58% of the total variation and only number of days to 50% flowering showed positive relation with the third component (PC₃) signifying its importance in divergence. Ghafoor *et al.*⁵ evaluated chickpea accessions by using multivariate techniques. The first three PCs with eigen values greater than one contributed 83.3% of the variability amongst genotypes. Ghafoor *et al.*⁵ studied genetic diversity in blackgram germplasm accessions; quantitative traits were analyzed for cluster and principal component analysis. The first four PCA with eigen values greater than one contributed 79.5% of that total variability among accessions. Iqbal *et al.*¹⁰ evaluated soybean genotypes by using multivariate

techniques. The first 3 PCs with eigen values greater than one contributed 69.77% of the total variability among the soybean genotypes. He observed that PC₁ were implicated by number of filled pods, grain yield per plant, biological yield per plant and harvest index (%). Elizabeth et al.³ 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.

Table 1: Eigen values, proportion of the variance represented by nine principal components, cumulative per cent variance and component loading of different characters in soybean

	PC ₁	PC ₂	PC ₃	PC ₄	PC ₅	PC ₆	PC ₇	PC ₈	PC ₉
Eigen value (Root)	4.757	1.344	1.042	0.706	0.455	0.292	0.193	0.114	0.096
%Var. Exp.	52.854	14.937	11.58	7.843	5.057	3.247	2.144	1.272	1.066
Cum. Var. Exp.	52.854	67.791	79.371	87.214	92.271	95.518	97.663	98.934	100.00
DTF	0.255	0.517	0.754	0.126	0.272	-0.01	0.083	0.043	-0.020
DTM	0.888	-0.259	-0.133	0.193	-0.073	-0.050	0.096	0.254	-0.090
PH	0.773	-0.251	-0.137	0.494	0.124	-0.085	0.130	-0.190	-0.016
BP	0.862	-0.33	0.164	0.100	0.115	0.085	-0.225	0.043	0.193
PP	0.843	-0.006	0.013	-0.420	-0.096	0.119	0.275	-0.031	0.112
SP	0.418	0.517	-0.608	-0.128	0.411	-0.001	-0.035	0.030	0.004
PL	0.427	0.732	-0.13	0.312	-0.384	0.125	-0.048	-0.008	0.043
100-SW	0.899	-0.091	0.118	-0.204	-0.007	0.261	-0.136	-0.08	-0.187
SY/P	0.828	0.162	0.099	-0.258	-0.146	-0.421	-0.106	-0.045	-0.018

DTF = Days to 50% flowering, DTM = Days to maturity, PH = Plant height (cm), BP = Number of branches/plant, PP = Number of pods/plant, SP = Number of seeds/pod, PL = Pod length (cm), 100-SW = 100-seed weight (g), SY = Seed yield/plant (g)

CONCLUSION

In the present study, the highly correlated variables (e.g. 100-seed weight, number of days to maturity, number of branches per plant, pods per plant and seed yield per plant) contributed greater variation to the total variance in yield traits as shown by the first principal component analysis (Table 1). A total of 79.37% of the total variation was contributed by these traits. Hence selection for yield using these traits will be effective.

REFERENCES

1. Chatfield, C. and Collis, A., *Introduction to Multivariate Analysis*, CRC Press, Boca Raton, 246 pp. (1980).
2. Dasgupta, T. and Das, P.K., Multivariate analysis and selection of parents for hybridization in blackgram. *Philippine Agriculturist* **57(1)**: 86-92 (1984).
3. Elizabeth, A.V., Schammas, E.A., Vencovsky, R., Martins, P.S. and Bandel, G., Germplasm characterization of *Sesbania* accessions based on multivariate analyses. *Genet. Resour. Crop Evol.*, **48**: 79-90 (2001).
4. Ghafoor, A., Sharif, A., Ahmad, Z., Zahid, M.A. and Rabbani, M.A., Genetic diversity in blackgram (*Vigna mungo* (L.) Hepper). *Field Crops Res.*, **69**: 183-190 (2001).
5. Ghafoor, A., Gulbaaz, F.N., Afzal, M., Ashraf, M. and Arshad, M., Inter-relationship between SDS-PAGE markers and agronomic traits in chickpea (*Cicer arietinum* L.). *Pak. J. Bot.*, **35(4)**: 613-624 (2003).
6. Ghosh, J., Ghosh, P.D. and Choudhary, P.R., An assessment of genetic relatedness between soybeans [*Glycine max* (L.) Merrill] cultivars using SSR markers, *American Journal of Plant Sciences*, **5**: 3089-3096 (2014).
7. He, F.J. and Chen, J.Q., Consumption of soybean, soy foods, soy isoflavones and breast cancer incidence: Differences between Chinese women and women in Western countries and possible mechanisms, *Food Science and Human Wellness*, **2**: 35-38 (2013).
8. Hossain, Z. and Komatsu, S., Potentiality of soybean proteomics in untying the mechanism of flood and drought stress tolerance, *Proteomes*, **2(1)**: 107-127 (2014).

9. Jolliffe, I., *Principal Component Analysis. Springer Series in Statistics* (2nd edition), New York, USA, 487 pp. (2002).
10. Iqbal, Z., Arshad, M., Ashraf, M., Mahmood, T. and Waheed, A., Evaluation of soybean [*Glycine max* (L.) Merrill] germplasm for some important morphological traits using multivariate analysis. *Pak. J. Bot.* **40(6)**: 2323-2328 (2008).
11. Rabbani, M.A., Iwabuchi, A., Murakami, Y., Suzuki, T. and Takayanagi, K., Phenotypic variation and the relationships among mustard (*Brassica juncea* L.) germplasm from Pakistan. *Euphytica* **101**: 357-366 (1998).
12. Sathe, S.K., Sharma, G.M., Kshirsagar, H.H., Su, M. and Roux, K.H., Effects of long-term frozen storage on electrophoretic patterns, immune reactivity, and pepsin in vitro digestibility of soybean [*Glycine max* (L.) Merrill] proteins, *Journal of Agricultural and Food Chemistry*, **57**: 1312-1318 (2009).
13. Salimi, S., Lahiji, H.S., Abadi, G.M., Salimi, S. and Moradi, S., Genetic diversity in soybean genotypes under drought stress condition using factor analysis and cluster analysis. *World Appl. Sci. J.*, **16**: 474–478 (2012).
14. Sharma, S., Kaur, M., Goyal, R. and Gill, B.S., Physical characteristics and nutritional composition of some new soybean [*Glycine max* (L.) Merrill] genotypes, *Journal of Food Science and Technology*, **51**: 551-557 (2013).