The Bioscan 15(2):257-259, 2020
AN INTERNATIONAL QUARTERLY JOURNAL OF LIFE SCIENCES
www.thebioscan.com

MULTIVARIATE ANALYSIS IN BLACKGRAM (*Vigna mungo* (L.) HEPPER)

DUDDUKUR RAJASEKHAR[1]*, B. NANDINI PRIYA YADAV[1], K. HEMALATHA[1] ,GANGISHETTI RANJITHKUMAR[2], AND GABRIYAL M. LAL[1]

[1]Department of Genetics and Plant Breeding, Naini Institute of Agriculture, Sam Higginbottom University of Agriculture, Technology and Sciences, Prayagraj - 211007, INDIA.

[2]Department of Genetics and Plant Breeding, Palli Siksha Bhavana, Visva-Bharati, Sriniketan, West Bengal -731236, INDIA

e-mail: drnaidu.agri@gmail.com

KEYWORDS
Principal component analysis
PCA
Cluster
Blackgram

Received on :
24.03.2020

Accepted on :
05.05.2020

***Corresponding author**

ABSTRACT

The study comprises of thirty-eight blackgram genotypes, which was evaluated for thirteen quantitative traits to undertake the multivariate analysis. The First six principal components (PCs) with eigen values >1 contributed 76.8% of the variability amongst accessions. Among the first six PCs, PC1 accounted for a high proportion of whole variance (21.30%) and the remaining five principal components *viz*., PC2, PC3, PC4, PC5 and PC6 revealed 14.70%, 13.20%, 11.10%, 8.60% and 7.90%. of the entire variance, respectively. The genotypes were categorized into five clusters based on average linkage between groups and clusters I and IV were more clearly separated than II, III and V. The principal traits are pod length (0.531), number of pods per plant (0.311) from PC2, number of primary branches per plant (0.53) from PC3, number of seeds per pods (0.481), seed index (0.401), length of pods (0.397), number of clusters per plant (0.279) from PC4, seed yield (0.383) from PC5 and seed index (0.275) from PC6. it indicates that these identified traits within the first six axes (76.8%) exhibited great influence on the phenotype. So, these traits are considered as key traits for selection criteria to develop high yielding varieties.

INTRODUCTION

The Asian Vigna group of grain legumes consists of six domesticated species, among them black gram is widely grown in South Asia and to a lesser extent in Southeast Asia. In India, blackgram is known as urd bean, minumulu *etc*., and it is a common pulse for idli, pongal and curry dishes. The most important limitations in achieving higher productivity of blackgram crop is lack of genetic variability, absence of suitable cultivars for different cropping systems, poor harvest index and susceptibility to diseases (Soufranamien and Gopalakrishna, 2004). Genetic relationships based on morphological data provide a way of making a relatively rapid assessment of the diversity among genotypes, so that a greater number can be subsequently tested with less cost. An understanding of the genetic relationships among lines can be particularly useful in planning crosses, in assigning lines to specific heterotic groups, and in precise identification with respect to plant varietal protection (Hallauer, Carena, et al., 2010). The multivariate analysis, and in specific, the principal component analysis and cluster analyses have been employed for the assessment of genotypes to examine various traits (MARolA, KBNT, *et al*., 1979). Principal component analysis (PCA) can be used to reveal the resemblances between traits and classify the genotypes, while cluster analysis instead it is concerned with classifying previously unclassified genotypes(Kaufman and Rousseeuw, 2009). PCA provides a roadmap for how to reduce a complex data set to a lower dimension to sometimes hidden, simplified structures that often underlies it. Principal component analysis is appropriate for obtaining measures on a number of observed variables and to develop a smaller number of artificial variables (called principal component) that will account for most of the variance in the observed variables. The first step in PCA is to calculate eigenvalues, which define the amount of total variation that is displayed on the PC axes. The first PC summarizes most of the variability present in the original data relative to all remaining PCs. The second PC explains most of the variability not summarized by the first PC and uncorrelated with the first, and so on (Jolliffe, 1986).

Multivariate statistical tools are extensively useful to summarize and describe the inherent variation among the genotypes. Principal component analysis (PCA) is one of the technique that identifies the plant traits that characterize the distinctness among the genotypes (Chakravorty, Ghosh, *et al*., 2013). PCA is also used to classify the population into distinct groups based on the similarities in one or more traits and consequently help us in the hybridization for selecting parents (Omokhafe and Alika, 1999). The principal components may be used as criterion variables in consequent analyses (Ray, Dutta, *et al*., 2014). The objective of this study is to identify superior genotypes and important traits using Principal Component Analysis (PCA) and classifying the genotypes into different clusters based on hierarchical grouping.

MATERIALS AND METHODS

The thirty-eight genotypes were evaluated in the experimental farm of department of Genetics and plant breeding, SHUATS, prayagraj. The experiment was laid using randomized block design (RBD) design with three replications. Each genotype was represented by five rows of one-meter length with a spacing of 30 cm x l0cm. A fertilizer dose of 20:40:20 kg NPK/ ha was applied and need based plant protection measures were followed. The observations were recorded on the randomly taken five plants in each entry. The observations were recorded on thirteen quantitative traits *viz*, days to 50% flowering (DFF), days to 50% pod setting (DFP), plant height (PH), number of primary branches per plant (PP), number of clusters per plant (CP), number of pods per plant (PP), number of seeds per plant (SP), pod length (PL), days to maturity (DM), seed index (SI), biological yield (BY), harvest index (HI) and seed yield per plant (SY). Cluster analysis done based on the hierarchical grouping and clustering of genotypes in to similarity groups using average linkage between groups method (Peeters and Martinelli, 1989) by SPSS (IBM Corp. Released 2013. IBM SPSS Statistics for Windows, Version 22.0. Armonk, NY: IBM Corp.). Principal component analysis was done as suggested (Nadarajan, 2005) using OPSTAT(Sheoran, 1998).

RESULTS AND DISCUSSION

In present study, PCA was performed for thirteen quantitative traits of blackgram. Out of thirteen principal components (PCs), only six PCs exhibited more than 1.00 eigen value (Table 1) *viz*., PC1 (2.77), PC2 (1.914), PC3 (1.712), PC4 (1.44), PC5 (1.123) and PC6 (1.033) showed about 76.80% variability among the traits studied for each genotype. Hence, these six principal components were given due importance for the further explanation. The PC1 had 21.30%, PC2 showed 14.70%, PC3 13.20% exhibited and PC4 showed 11.10% PC5 showed 8.60%, PC6 7.90%. Rotated component matrix (Table 2) revealed that the first principal component was more negatively related to seed yield and its contributing traits such as seed yield (-0.421), harvest index (-0.506), seed index (-0.327), number of pods per plant (-0.325) and positively related to vegetative trait like days to maturity (0.288). The second principal component was more related to length of pod (0.531), plant height (0.358), number of pods per plant (0.311) and negatively related to number of clusters per plant (-0.444). The third principal component was more related to vegetative traits such as days to 50% flowering (0.479), days to 50% maturity (0.387), plant height (0.478) and number of primary branches per plant (0.53). These results suggest that PC3 reveals that the tendency of each genotype to emphasize vegetative, as opposed to reproductive growth and incline to have few large reproductive organs. The fourth principal component was more positively related to yield contributing traits such as number of seeds per pods (0.481), seed index (0.401), length of pods (0.397), number of clusters per plant (0.279) and biological yield (0.442) but harvest index (-0.263) is negatively related. The fifth principal component was negatively related to days to 50% maturity (-0.517) and biological yield (0.565) and seed yield (0.383) are positively related. The sixth principal component was negatively related to number of clusters per plant (-0.435), number of seeds per pod (-0.367) and days to maturity (-0.55) and seed index (0.275) was positively related. Rotated component matrix revealed that first six PCs are representing maximum variability (76.80%) hence, the traits falling in these six PCs must be given due importance in breeding programme. The positive and negative loading shows the presence of positive and negative correlation trends between the components and the variables. (Mohanlal, Saravanan, *et al*., 2018) observed 79.12% of cumulative variance among the first four axes where eigen values more than 1. (Jeberson, Shashidhar, *et al*., 2019) studied principal component and cluster analysis in black gram, the eigen values more than 1 have contributed 84.52% of variability among the 25 and they were grouped into five clusters based on the

Table 1: Eigenvalues and proportion of variation of principal components

	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8	PC9	PC10	PC11	PC12	PC13
Eigenvalues	2.77	1.914	1.712	1.44	1.123	1.033	0.822	0.67	0.483	0.39	0.332	0.249	0.061
Proportion of variation	0.213	0.147	0.132	0.111	0.086	0.079	0.063	0.052	0.037	0.03	0.026	0.019	0.005
Cumulative Proportion of variation	0.213	0.36	0.492	0.603	0.689	0.769	0.832	0.883	0.921	0.951	0.976	0.995	1

Table 2 : Loadings (Eigenvectors) of thirteen quantitative charcters for principal components

	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8	PC9	PC10	PC11	PC12	PC13
Days to 50% flowering	0.205	-0.209	0.479	0.128	-0.092	0.247	-0.468	-0.165	-0.037	-0.344	-0.332	0.339	0.089
Days to 50% pod setting	0.249	-0.004	0.387	-0.068	-0.517	-0.242	-0.048	0.043	0.558	0.298	0.175	-0.15	-0.001
Plant height	0.016	0.358	0.478	0.012	0.211	-0.01	0.133	-0.506	-0.32	0.381	-0.13	-0.245	0.007
Number of primary branches	-0.158	0.21	0.53	0.04	0.051	0.127	0.408	0.356	-0.043	-0.451	0.348	0.022	-0.095
Clusters per plant	-0.146	-0.444	0.046	0.279	-0.217	-0.435	0.122	-0.319	-0.341	-0.07	0.395	0.134	0.229
Pods per plant	-0.325	0.311	-0.05	-0.106	0.184	-0.219	-0.511	-0.335	0.272	-0.352	0.319	-0.161	-0.059
Pod length	-0.007	0.531	-0.113	0.397	-0.144	0.14	-0.196	0.13	-0.077	0.296	0.269	0.494	0.203
Seeds per pod	-0.248	0.255	-0.06	0.481	-0.211	-0.367	0.189	0.024	0.151	-0.228	-0.582	-0.094	0.009
Days to maturity	0.288	0.169	0.08	-0.204	0.028	-0.55	-0.325	0.452	-0.462	-0.044	-0.065	-0.084	-0.013
Seed index	-0.327	-0.226	0.105	0.401	-0.069	0.275	-0.364	0.302	-0.155	0.203	0.072	-0.54	-0.063
Biological yield	0.241	-0.185	0.086	0.442	0.565	-0.241	-0.013	0.052	0.239	0.145	0.089	0.17	-0.457
Harvest index	-0.506	-0.059	0.108	-0.263	-0.241	-0.076	-0.068	0.008	-0.127	0.221	-0.114	0.375	-0.61
Seed yield per plant	-0.421	-0.158	0.234	-0.179	0.383	-0.173	-0.027	0.242	0.231	0.248	-0.151	0.192	0.548

Table 3: Distribution of genotypes into different clusters

Clusters	Number of Clusters	Genotypes
Cluster I	10	AZAD-1, IC-140816 , IPU-96-1, PGRV-99022, PLU 826, PLU-277, PU-31, SHEKHAR-3, SPS-33, UH-85-5
Cluster II	18	IC-106194, IC-24129, IC-250188, IC-56048, IC-91567, IPU-199-60, IPU9 94-10, MDU8-5-7, P-1, PANT U-30, PLU-648, PLU-710, STNP2, UH 81-89, UH-10, UH-82-35, UH82-83, UTTARA
Cluster III	1	IC-250187
Cluster IV	8	IC-250190, IPU 7-3, PKG U3, T-9, U-5, U-9, UG-27, UH82-15
Cluster V	1	IPU 99-16

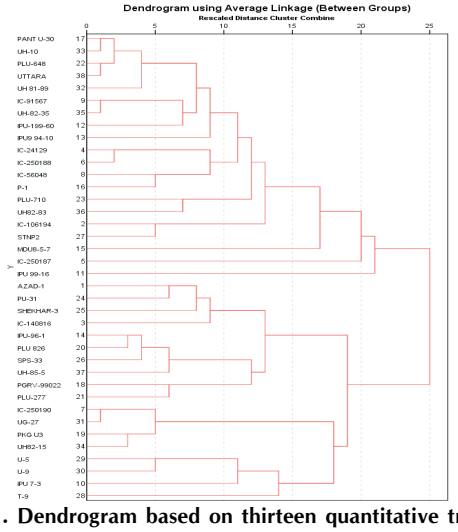

Figure 1. Dendrogram based on thirteen quantitative traits using average linkage between groups.

average linkage. (Rajasekhar, Lal, *et al*., 2017) observed in their study that Characters such as number of primary branches per plant, number of clusters per plant, pod length, pods per plant, seed index, seeds per pod and harvest index were having positive correlation with direct and dominant role in higher seed yield.

Hierarchical clustering classified the thirty-eight genotypes into five clusters (Table 3) and dendrogram was formed using average linkage between groups (Figure 1). Cluster II contains eighteen genotypes followed by cluster I contains ten genotypes, cluster VI contains eight genotypes and both Clusters III and V contains single genotype each. (Vyas, Joshi, *et al*., 2016) were also observed five clusters from the twenty two genotypes in their study. It clearly indicates that crossing between the genotypes from different clusters will gives heterotic hybrids. The cluster II contains highest number of genotypes, it indicates that the intragroup or intra-cluster hybridization is useful for the effective population development in the breeding programme. Inter-cluster hybridization is suggested among the genotypes of different clusters, so it may result heterotic hybrids.

REFERENCES

Chakravorty, A., P. Ghosh and P. Sahu. 2013. Multivariate analysis of phenotypic diversity of landraces of rice of West Bengal. *J. Experimental Agriculture International*. PP.110-123.

Hallauer, A. R., M. J. Carena and J.d. Miranda Filho. 2010. Quantitative genetics in maize breedingSpringer Science and Business Media.

Jeberson, M.S., K. Shashidhar and A.K. Singh. 2019. Genetic variability, principal component and cluster analyses in black gram under Foot-hills conditions of Manipur. Legume Research: *An International J*.P. 42.

Jolliffe, I.T. 1986. Principal components in regression analysis. Principal component analysis. *Springer*. p. 129-155.

Kaufman, L. and P.J. Rousseeuw. 2009. Finding groups in data: an introduction to cluster analysisJohn Wiley and Sons.

MARolA, K., J. KBNT and J. Bibly. 1979. Multivariate analysisAcadeInic Press, Londres.

Mohanlal, V., K. Saravanan and T. Sabesan. 2018. Multivariate analysis in blackgram (*Vigna mungo* L. hepper) genotypes. *J. Pharmacognosy and Phytochemistry*. **7:** 860-863.

Nadarajan, N. 2005. Quantitative genetics and biometrical techniques in plant breedingKalyani Publishers.1999.

Omokhafe, K. O. and J. E. Alika. 1999. "Multivariate of agronomic data on Hevea brasiliensis." In Proceeding of SUSANIBS Meeting, pp. 168-173.

Peeters, J. and J. Martinelli. 1989. Hierarchical cluster analysis as a tool to manage variation in germplasm collections. *Theoretical and Applied Genetics*. **78:** 42-48.

Rajasekhar, D., S.S. Lal and G.M. Lal. 2017. character association and path analysis for seed yield and its components in black gram [*Vigna mungo* (L.) Hepper]. *Plant Archives* .**17:** 467-471.

Ray, K., J. Dutta, H. Banerjee, R. Biswas, A. Phonglosa and A. Pari. 2014. Identification of principal yield attributing traits of Indian mustard [*Brassica juncea* (L.) Czernj and Cosson] using multivariate analysis. *The Bioscan*. **9:** 803-809.

Sheoran, O.P.T. D.S., Kaushik, L. S., Hasija, R.C and Pannu, R.S 1998. Statistical Software Package for Agricultural Research Workers. Recent Advances in information theory, Statistics and Computer Applications D.S. Hooda and R.C. Hasija Department of Mathematics Statistics, CCS HAU, Hisar.PP. 139-143.

Souframanien, J. and T. Gopalakrishna. 2004. A comparative analysis of genetic diversity in blackgram genotypes using RAPD and ISSR markers. Theor Appl Genet 109: 1687-1693. doi:10.1007/s00122-004-1797-3.

Vyas, D., Joshi,A., Keder, O., Dodiya,N., Sukla,K. and Kaur,G. 2016. Multivariate analysis on genetic divergence in ysis on genetic divergence in different environments in black gram genotypes. *The Bioscan*. **11:** 3197-3201.