Biological Forum – An International Journal **14(3): 1318-1323(2022)**

ISSN No. (Print): 0975-1130
ISSN No. (Online): 2249-3239

Principal Component Analysis (PCA) in Pigeonpea (*Cajanus cajan* (L.)Millsp.) germplasm

***M. Dhanushasree*[1*] *and A. Thanga Hemavathy*[2]**
[1]*PG Scholar Department Department of Plant breeding and Genetics University,*
Tamil Nadu Agricultural University, Coimbatore (Tamil Nadu), India.
[2]*Assistant Professor Department Department of Plant breeding and Genetics,*
Tamil Nadu Agricultural University, Coimbatore (Tamil Nadu), India.

(Corresponding author: M. Dhanushasree)*
(Received 26 June 2022, Accepted 17 August, 2022)
(Published by Research Trend, Website: www.researchtrend.net)

ABSTRACT: The research was conducted using sixty eight pigeonpea genotypes based on yield and its contributing traits using principal component analysis (PCA) in order to determine the genetic diversity among genotypes. Analysis of Variance (ANOVA) revealed significant differences among accessions for all the quantitative traits *viz.*, days to fifty per cent flowering and days to maturity, plant height (cm), number of primary branches per plant, pod length (cm), number of pods per plant, number of seeds per pod, pod weight per plant (g) and hundred seed weight (g) and single plant yield (g). Principal component analysis (PCA) revealed that the first three principal components had eigen value greater than 1 which accounted for 68.72 per cent of total variation. PCA biplot revealed that the traits days to maturity and days to fifty per cent flowering had weak positive correlation with hundred seed weight, zero correlation was observed with traits *viz.*, pods per plant, single plant yield and pod weight. Negative correlation was observed between number of seeds per pod and the traits days to maturity and fifty per cent flowering. Thus, the above results were utilized to select the more diverse germplasms and could be used as parents in heterosis breeding programmes.

Keywords: Pigeonpea, eigen vector, eigen value, loadings, PCA, Biplot.

INTRODUCTION

Pigeonpea (*Cajanus cajan* (L.) Millsp.) is a photosensitive, highly nutritious tropical grain legume crop belongs to the family Fabaceae with chromosome number (2n=22). Introduction of short duration pigeonoea varieties have been developed to accommodate in diverse cropping systems but couldn't have comparable yield as long duration ones. Hence, breeding of varieties with high yield and early maturing is the need of the hour. Genetic diversity studies are a prerequisite for any crop improvement program. Hybridization and selection are two crucial steps in the process of plant breeding. Thus, study of many morphological characters in germplasm is important for assessment of the differences among populations in order to identify population differences and breeding potential. Plant breeders frequently measure a huge number of characteristics, some of which may not have enough discriminatory power for evaluation, characterization and management germplasm (Maji and Shaibu 2012). In such case, principal component analysis (PCA) may be used to reveal patterns and eliminate redundancy in data sets (Mahalingam *et al.*, 2020) as morphological and physiological variations routinely occur in crop species. PCA simplifies a large data series into smaller set of variables without losing any vital information of original data set by seeking for groups that have very strong inter-correlation with per cent variation explained for each component towards the total variability (Das *et al.*, 2017).. Biplot Analysis is an exploratory graph method used in statistic with two-dimensional chart that represents of multidimensional dataset. The research aim was planned to estimate the diversity through classification of genotype through Principal Component Analysis (PCA).

MATERIALS AND METHODS

Sixty eight pigeonpea genotypes along with five checks (Table 1) were evaluated on the basis of ten agro-morphological traits at Department of Pulses, Tamil Nadu Agricultural University, Coimbatore. The genotypes were raised in Augmented Block Design II (ABD II) during *rabi*, 21. Ten quantitative traits used for diversity analysis viz., days to fifty per cent flowering and days to maturity, plant height (cm), number of primary branches per plant, pod length (cm), number of pods per plant, number of seeds per pod, pod weight per plant (g) and hundred seed weight (g) and single plant yield (g). Principal Component Analysis (PCA) was performed using the mean data of pigeonpea accessions. The analysis was carried out using R software package.

RESULTS AND DISCUSSION

Analysis of Variance (ANOVA) performed for ten quantitative traits revealed presence of significant variation among genotypes and can be subjected to further analysis (Table 2).

Eigen vectors and eigen values. The Principal Component Analysis (PCA) resulted in the formation of three principal components with eigen value greater than 1. The eigen value for PC1, PC2 and PC3 was 3.41, 2.24 and 1.21, respectively. Other principal components (PC4 to PC10) had discriminatory power. The first three principal components which accounted for 68.72 per cent of total variation which was depicted in the scree plot (Fig. 1) and Table 3. Hamid *et al*, (2011) reported 76.2% of total variance during evaluation of hundred germplasm lines for various morphological and agronomic traits. Component matrix showing latent vectors associated with first three principal components was given in the Table 4. Maximum amount of variation was contributed by PC1 (34.16%) to the total variance followed by second principal component (PC2) (22.44%) and third principal component (PC3) (12.10%). The first principal component (PC1) contributed maximum variation to the total variation showed positive loadings for all the studied traits. Similar findings were reported by Rekha *et al*. (2013); Tharageshwari and Hemavathy (2020) using pigeonpea germplasm accessions. The eigen vector value > 0.4 may be considered as large loadings. The trait with eigen vector values greater than 0.4 could be effectively utilized for selection which contributes greater variation to the principal component. Loadings with maximum value was observed in traits *viz.,* number of pods per plant (0.973) followed by single plant yield (0.905), pod weight (0.87), number of branches per plant (0.229) and number of seeds per pod (0.207). The results were found accordance with findings of Hemavathy *et al*. (2017), where PC1 contributed maximum to total variance with positive correlation with all traits except single plant yield.

The second principal component (PC2) showed 22.44 per cent of variation which exhibited positive loading for all traits except number of seeds per pod (-0.027). The traits *viz.*, days to maturity (0.984), days to 50 % flowering (0.98), plant height (0.337) and number of branches per plant (0.108) which showed large positive loading and provide maximum contribution to percent of variance. Similar findings were reported by Yohane *et al*. (2020) for traits *viz.*, plant height, number of primary branches, number of pods per plant, hundred seed weight and grain yield per plant.

The third principal component (PC3) contributed 12.10 per cent of total variation. It showed positive loadings for traits *viz.,* number of seeds per pod (0.974), single plant yield (0.233), pod weight (0.185), plant height (0.058), pod length (0.052) and number of branches per plant (0.027). Conversely, negative loading was present in traits *viz.,* days to fifty per cent flowering (-0.012), days to maturity (-0.019), number of pods per plant (-0.001) and hundred seed weight (-0.04) which depicted that these traits showed negative contribution to the genetic divergence. The results were supported by the findings of Zavinon *et al*. (2019) where, the traits days to 50 per cent flowering and days to maturity showed negative contribution to divergence.

PCA Biplot analysis. PCA biplot was indicated with two-dimensional chart that represents multidimensional dataset (Fig. 3). The correlation between the variables can be detected by measuring the angle between two variables. If the angle between two variables was 90 degrees, it was represented that there was no correlation between two variables. Positive correlation was detected if the angle was less than 90 degree and if the angle was more than 90 degrees, it was found that negative correlation between two variables. Vector line length was used for representation of variance level of the variable. The distance between genotypes and the variable were used to detect the contribution of variable to the genotype. Closer the distance higher the value and vice versa for farther distance. Considering the above conditions, PCA biplot was concluded that the traits days to maturity and days to fifty per cent flowering had weak positive correlation with hundred seed weight, zero correlation with pods per plant, single plant yield and pod weight. Negative correlation was observed between number of seeds per pod and the traits days to maturity and fifty per cent flowering. High positive correlation was observed among the traits *viz.*, days to maturity, days to fifty per cent flowering, plant height, pod length and number of branches per plant. The traits *viz.*, days to maturity, days to fifty per cent flowering, single plant yield (g), number of pods per plant and pod weight showed maximum variance as these variables were represented to have maximum vector length. Contrarily, the traits pod length, hundred seed weight and number of seeds per pod showed minimum variance with shorter vector length. The results were in concurrence with the findings of Yohane (2020) while analysing phenotypic divergence of eighty one pigeonpea genotypes for both quantitative and qualitative traits. In the PCA biplot, the best performing genotypes were concentrated in the right quadrant of the biplot. The genotype AL 2184 which observed to have minimum value for all the variables was plotted in the negative left quadrant. Likewise, the genotypes CO(Rg) 7, DPP 3-81 and ICP 9808 which showed maximum performance of all the variables plotted on the right quadrant.

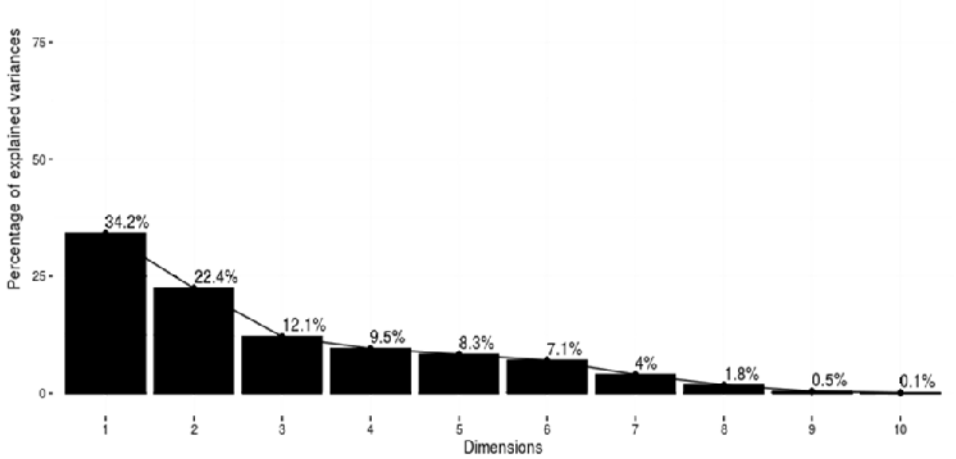


Fig. 1. Scree plot showing variances for ten principal components under PCA analysis.

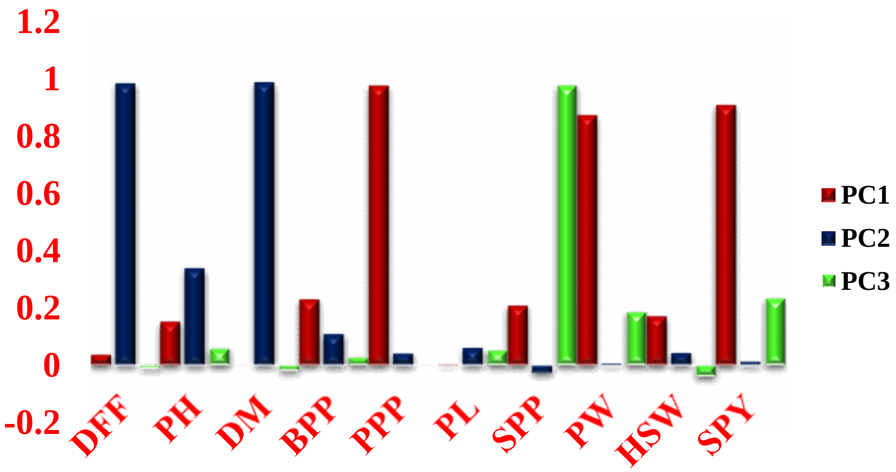
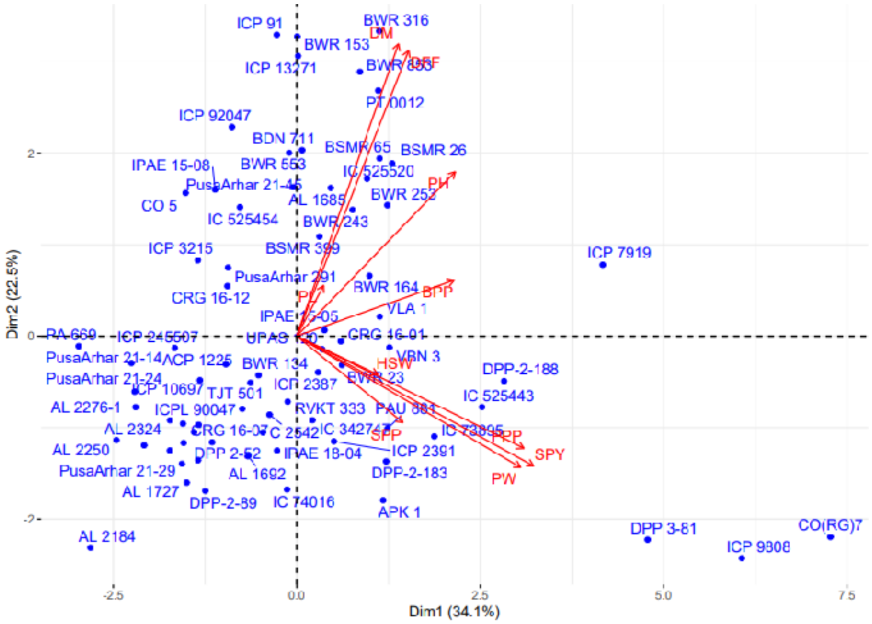


Fig. 2. Component matrix showing latent vectors associated with first three principal components.



Dim 1= dimension 1, Dim 2=dimension 2.

Fig. 3. PCA biplot showing variance exhibited by pigeonpea genotypes and ten quantitative traits.

Table 1: List of pigeonpea genotypes and checks used in the study

Sr. No.	Genotypes	Source of origin
1.	AL 1685	PAU, Ludhiana
2.	AL 1692	PAU, Ludhiana
3.	AL 1727	PAU, Ludhiana
4.	AL 2184	PAU, Ludhiana
5.	AL 2250	PAU, Ludhiana
6.	AL 2276-1	PAU, Ludhiana
7.	AL 2324	PAU, Ludhiana
8.	BSMR 26	IIPR, Kanpur
9.	BSMR 399	IIPR, Kanpur
10.	BSMR 65	IIPR, Kanpur
11.	BWR 134	IIPR, Kanpur
12.	BWR 153	IIPR, Kanpur
13.	BWR 164	IIPR, Kanpur
14.	BWR 23	IIPR, Kanpur
15.	BWR 243	IIPR, Kanpur
16.	BWR 253	IIPR, Kanpur
17.	BWR 316	IIPR, Kanpur
18.	BWR 553	IIPR, Kanpur
19.	BWR 853	IIPR, Kanpur
20.	CRG 16-01	TNAU, Coimbatore
21.	CRG 16-07	TNAU, Coimbatore
22.	CRG 16-12	TNAU, Coimbatore
23.	DPP 2-52	IIPR, Kanpur
24.	DPP 3-81	IIPR, Kanpur
25.	DPP-2-183	IIPR, Kanpur
26.	DPP-2-188	IIPR, Kanpur
27.	DPP-2-89	IIPR, Kanpur
28.	DPP-3-2	IIPR, Kanpur
29.	IC 342747	ICRISAT, Telangana
30.	IC 525443	ICRISAT, Telangana
31.	IC 525454	ICRISAT, Telangana
32.	IC 525520	ICRISAT, Telangana
33.	IC 73895	ICRISAT, Telangana
34.	IC 74016	ICRISAT, Telangana
35.	ICP 10697	ICRISAT, Telangana
36.	ICP 13271	ICRISAT, Telangana
37.	ICP 2387	ICRISAT, Telangana
38.	ICP 2391	ICRISAT, Telangana
39.	ICP 245507	ICRISAT, Telangana
40.	ICP 3215	ICRISAT, Telangana
41.	ICP 7919	ICRISAT, Telangana
42.	ICP 91	ICRISAT, Telangana
43.	ICP 92047	ICRISAT, Telangana
44.	ICP 9808	ICRISAT, Telangana
45.	ICPL 11301	ICRISAT, Telangana
46.	ICPL 20325	ICRISAT, Telangana
47.	ICPL 90047	ICRISAT, Telangana
48.	IPAE 15-05	IIPR, Kanpur
49.	IPAE 15-08	IIPR, Kanpur
50.	IPAE 18-04	IIPR, Kanpur
51.	PA 509	GBPAU&T, Pantnagar
52.	PA 669	GBPAU&T, Pantnagar
53.	PT 0012	IIPR, Kanpur
54.	PusaArhar 21-14	Pusa, New Delhi
55.	PusaArhar 21-24	Pusa, New Delhi
56.	PusaArhar 21-27	Pusa, New Delhi
57.	PusaArhar 21-29	Pusa, New Delhi
58.	PusaArhar 21-45	Pusa, New Delhi
59.	PusaArhar 21-57	Pusa, New Delhi
60.	PusaArhar 21-61	Pusa, New Delhi
61.	PusaArhar 291	Pusa, New Delhi
62.	C 2542	IIPR, Kanpur
63.	ACP 1225	IIPR, Kanpur
64.	RVKT 333	IIPR, Kanpur
65.	TJT 501	BARC & Khargone
66.	CO 5	TNAU, Coimbatore
67.	PAU 881	PAU, Ludhiana
68.	BDN 711	ARS, Badnapur

S .No.	Checks	Source of origin
1.	CO(Rg) 7	TNAU, Coimbatore
2.	VLA 1	ICRISAT, Hyderabad
3.	UPAS 120	GBPAU&T, Pantnagar
4.	APK 1	RRS, Aruppukkottai
5.	VBN 3	NPRC, Vamban

Table 2: Mean sum square from ANOVA for Augmented Block Design II for ten quantitative characters.

S. No.	Characters	Blocks	Total entries	Checks	Genotypes	Checks vs Genotypes	Error
1.	Days to 50% flowering	10.85*	62.5363**	36.7**	64.8909**	8.1225*	2.75
2.	Plant height	674.0294**	390.4178**	1132.0517**	342.9843**	601.9246**	25.2274
3.	Days to maturity	4.3781	55.2631**	30.25**	57.4967**	5.6658	3.25
4.	No of branches/plant	5.9872**	2.2236**	4.325**	1.9569*	11.6834**	0.5917
5.	No of pods/plant	1211.0396**	447.3842**	1343.425**	366.293**	2296.3316**	58.1917
6.	Pod length(cm)	0.2641**	0.1453**	0.05	0.1389**	0.953**	0.0193
7.	No of seeds/pod	1.6025**	0.2166	0.3	0.2004	0.968**	0.095
8.	Pod weight(g)	109.707**	177.8501**	1656.7533**	89.5173**	180.5335**	9.2576
9.	100 seed weight(g)	0.4065**	0.7868**	6.0632**	0.4748**	0.5889**	0.0206
10.	Single plant yield(g)	1203.9458**	151.32**	1256.3768**	64.8273**	1526.1027**	3.9848

*Significant at 5% level **Significant at 1% level

Table 3: Eigen values and percentage of total variances for different principal components.

Components	Eigen value	Percentage of variance	Cumulative percentage of variance
PC1	3.417	34.168	34.168
PC2	2.245	22.446	56.614
PC3	1.21	12.105	68.719

Table 4: Component matrix showing latent vectors associated with first three principal components.

Characters	Principal components		
	PC1	PC2	PC3
Days to 50% flowering	0.036	0.98	-0.012
Plant height	0.151	0.337	0.058
Days to maturity	0.001	0.984	-0.019
No of branches/plant	0.229	0.108	0.027
No of pods/plant	0.973	0.04	-0.001
Pod length(cm)	0.003	0.06	0.052
No of seeds/pod	0.207	-0.027	0.974
Pod weight(g)	0.87	0.006	0.185
100 seed weight(g)	0.17	0.042	-0.04
Single plant yield (g)	0.905	0.012	0.233

CONCLUSION

PCA can be used as an important technique to enhance the breeding programme because it extracts all the key components and highlights their contribution to total variability. PCA biplot revealed the high performing genotypes *viz.*, CO(Rg) 7, DPP 3-81 and ICP 9808 can be effectively utilized for crop improvement programmes.

FUTURE SCOPE

Principal Component Analysis (PCA) helps to speed up crop enhancement by sorting the key components and their contribution to variation. Selection of genotypes which showing high contribution to yield can be made which could be utilized as parents in future breeding programmes.

Acknowledgement. I am grateful to all professors of Department of Pulses, Tamil Nadu Agricultural University, Coimbatore for their immense support to complete the research work successfully.

Conflict of Interest. None

REFERENCES

Das, S., Das, S. S., Chakraborty, I., Roy, N., Nath, M. K. and Sarma, D. (2017). Principal component analysis in plant breeding. *Biomol. Rep, 3*: 1-3.

Hamid, A., Husna, A., Haque, M. M. and Islam, M. R. (2011). Genetic variability in pigeonpea (*Cajanus cajan* L. Millspaugh). *Electronic Journal of Plant Breeding*, *2*(1): 117-123.

Hemavathy, A. T., Bapu, J. R. and Priyadharshini, C. (2017). Principal component analysis in pigeonpea *(Cajanus cajan* (L.) millsp.). *Electronic Journal of Plant Breeding*, *8*(4): 1133-1139.

Mahalingam, A., Manivanan, N., Kumar, K. B., Ramakrishnan, P. and Vadivel, K. (2020). Character association and principal component analysis for seed yield and its contributing characters in greengram (*Vigna radiat*a (L.) Wilczek). *Electronic Journal of Plant Breeding*, *11*(01): 259-262.

Maji, A. T. and Shaibu, A. A. (2012). Application of principal component analysis for rice germplasm characterization and evaluation. *Journal of Plant Breeding and Crop Science*, *4*(6) : 87-93.

Rekha, R., Prasanthi, L., Sekhar, M. R. and Priya, M. S. (2013). Principal Component and Cluster Analyses in Pigeonpea [*Cajanus cajan* (L.) MillSp].

Tharageshwari, L. M. And Hemavathy, A. T. (2020). Principal component analysis of pigeonpea {*Cajanus cajan* (L.) Millispaugh} Germplasm. *Applied Biological Research, 22*(2): 102-108.

Yohane, E. N. (2020). Genetic improvement of pigeonpea [*Cajanus cajan* (L.) millspaugh] for yield, earliness and resistance to Fusarium wilt (Fusarium udum Butler) in Malawi.

Yohane, E. N., Shimelis, H., Laing, M., Mathew, I. and Shayanowako, A. (2020). Phenotypic divergence analysis in pigeonpea [*Cajanus cajan* (L.) Millspaugh] germplasm accessions. *Agronomy*, *10*(11): 1682.

Zavinon, F., H. Adoukonou-Sagbadja, A Bossikponnon, H Dossa and C. Ahanhanzo (2019). "Phenotypic diversity for agro-morphological traits in pigeon pea landraces [(*Cajanus cajan* L.) Millsp.] cultivated in southern Benin." *Open Agriculture, 4*(1): 487-499.

How to cite this article: M. Dhanushasree and A. Thanga Hemavathy (2022). Principal Component Analysis (PCA) in Pigeonpea (*Cajanus cajan* (L.) Mill sp.) germplasm. *Biological Forum – An International Journal*, *14*(3): 1318-1323.