Received : 13 Mar 2020 | Revised : 28 Mar 2020 | Accepted : 29 Mar 2020

Electronic Journal of Plant Breeding

Research Article

Multivariate analysis of 102 Indian cowpea (*Vigna unguiculata* (L.) Walp.) germplasm

E. Vijayakumar[1], K. Thangaraj[2], T. Kalaimagal[3*], C. Vanniarajan[4], N. Senthil[5] and P. Jeyakumar[6]

[1] Ph D scholar (PBG), Tamil Nadu Agricultural University, Coimbatore- 641 003, Tamil Nadu, India.
[2] Assistant Professor, Agricultural College and Research Institute, Madurai-625 104, Tamil Nadu, India.
[3] Professor (PBG), Tamil Nadu Agricultural University, Coimbatore- 641 003, Tamil Nadu, India.
[4] Professor and Head (PBG), Agricultural College and Research Institute, Madurai-625 104, Tamil Nadu, India.
[5] Professor (CPMB &B), Tamil Nadu Agricultural University, Coimbatore- 641 003, Tamil Nadu, India.
[6] Professor and Head (CRP), Tamil Nadu Agricultural University, Coimbatore- 641 003, Tamil Nadu, India.
***E-Mail:** kalaimagalt@gmail.com

Abstract

Better understanding of genetic resources available in the crop is the crucial and foremost step in any breeding program. In the present study, 102 cowpea germplasm based on twelve quantitative characters were subjected to Principal Component Analysis (PCA). The PCA analysis dissected the total variation into five major principal components which accounted for 76.53 per cent of total variation. First two PC's were used to construct the biplot in which the genotypes *viz.,* VCP-12-008, PG-CP- 1, TY 1145, ACM 0505 are scattered apart in all the four quadrates representing maximum genetic divergence. Single plant yield followed by number of clusters per plant, number of pods per plant and number of seeds per pod contributes maximum divergence in the first PC. Hence, genetic and phenotypic variations exist among 102 cowpea genotypes could be used in genetic improvement of the cowpea through simple selection and crossing potential parents.

Key words

Cowpea, Germplasm, PCA and Genetic divergence.

INTRODUCTION

Global warming and nutritional security are the two major concern of current and future agriculture. A crop with climate resilience and high nutrition adds more value to these concerns in near future. Cowpea (*Vigna unguiculata* (L.) Walp.) is one such crop, which often referred as a "vegetable meat" and highly tolerant to drought (Carvalho *et al.,* 2019, Fatokun *et al.,* 2018). Cowpea is an economically important legume crop, originated from sub-Saharan Africa where the highest genetic diversity exists (Fatokun *et al.,* 2018, Saxena and Rukam 2020). It plays a major role in income generation of low scale farmers of Asia and Africa (Boukar *et al.,* 2016, Vavilapalli *et al.,* 2013). Cowpea is a dual purpose highly nutritive legume, consumed as a vegetable and grain by both human and animals (Avanza *et al.,* 2013). Its ability to fix atmospheric nitrogen with the help of symbiotic nitrogen fixing bacteria makes it as a nutritional enhancer crop in the poor and marginal soil fertility regions (Ghalmi *et al.,* 2010, Walle *et al.,* 2019). The better understanding of genetic resources available in the crop is the crucial and foremost step in any breeding program which helps in identifying the suitable parents and widening the genetic base of the crop (Prasanthi *et al.,* 2012, Xiong *et al.,* 2016). On the other hand, soaring demand to feed the growing population and development of new varieties led to the way for genetic erosion of valuable germplasm (Fang *et al.,* 2007). This dwindling of genetic resources in crop species possesses a huge danger to agricultural crop production (Muchero *et al.,* 2009). Africa and India being the primary centre of origin for the cowpea (Patel *et al.,* 2016), a diverse germplasm resources is available in India. Hence, with the above considerations a diversity analysis was carried out to identify the potential genotypes for the future breeding programs.

Genetic diversity being the multivariate analysis, several statistical tools like euclidean clustering and principal component analysis (PCA), are available. PCA is a technique or a mathematical algorithm helps in decomposing large number variables into fewer variables without losing much information (Abdi and Williams 2010). The dataset is reduced to fewer variables based on the eigen values by creating new uncorrelated variables called principal components (PC). These principal components helps in minimizing the data lose by maximizing the variance (Jolliffe and Cadima 2016). Principal component analysis is an adaptive data analysis technique which is effectively used to visualize the similarity and difference between the genotypes and helps in identifying the quantitative characters contributing maximum towards genetic divergence (Jindal *et al.,* 2018, Ringnér 2008). Hence, PCA was used to estimate the genetic diversity among the cowpea germplam in the present study.

MATERIALS AND METHODS

The experimental materials constituted 102 genotypes (100 germplasm and two checks) which were obtained from National Pulses Research Centre (NPRC), Vamban, Tamil Nadu, India. The field experiments were carried out at Agricultural College and Research Institute, Tamil Nadu Agricultural University (TNAU), Madurai which is geographically located at of 9° 54' N latitude and 78° 54' E longitude at an elevation of 147 m above mean sea level. Annual average rainfall is about 856 mm. Randomized Completely Block Design (RCBD) with two replications was followed as an experimental design. Each germplasm line was planted in three rows of 5 m with spacing of 30 ×15 cm. The observations on twelve quantitative characters *viz.,* plant height, days to fifty per cent flowering, days to maturity, number of primary branches, peduncle length, number of clusters per plant, number of pods per cluster, number of pods per plant, pod length, number of seeds per pod, hundred seed weight and single plant yield on 15 plants per replication were taken based on the cowpea descriptor developed by the International Board for Plant Genetic Resources (IBPGR 1983). The statistical analysis was carried out using the software R version 3.3.2 and R Studio 1.0.136.

RESULTS AND DISCUSSION

Principal component analysis is the oldest and most admired method developed by Karl Pearson way back in 1901, but still rules the data analytics because of its recent advancements in visualization and ability to reduce the multiple variables to fewer un correlated variables (Chiquet *et al.,* 2018). In the present study PCA was carried out with twelve quantitative characters of 102 cowpea germplasm. Similarly PCA was used as a genetic divergence and similarity measures by various researchers (Aremu *et al.,* 2007, Fang *et al.,* 2007, Sousa *et al.,* 2015, Walle *et al.,* 2019). The total variation was splited into twelve principal components equaling to the number of variables used in the analysis. The eigen values serves as potential criteria in selection of the critical principal components that contributed maximum to the variation (Gerrano *et al.,* 2019). Hence, in the present study first five principal components are the major contributors towards the total variation whose eigen values are more than one. Variables with eigen values less than one can be eliminated as variation caused by them will be non-significant and negligible (Walle *et al.,* 2019). First principal component contributes 26.37 per cent of total variation, while PC 12 contributes to only 0.21 per cent of total variation. First five PC's cumulatively contributes 76.53 per cent of variation (**Table 1**). Gixhari *et al.,* (2014) suggested that more than 75 per cent of total variation is acceptable for the genetic characterization of pulse crops.

Table1. Eigen values and contribution of twelve quantitative characters towards divergence.

Principal components	eigenvalue	Variance per cent towards divergence	Cumulative per cent variance towards divergence
PC 1	3.16	26.37	26.37
PC 2	2.12	17.67	44.03
PC 3	1.81	15.09	59.12
PC 4	1.12	9.30	68.43
PC 5	1.01	8.40	76.83
PC 6	0.79	6.54	83.37
PC 7	0.67	5.62	88.99
PC 8	0.53	4.41	93.40
PC 9	0.42	3.49	96.89
PC 10	0.29	2.40	99.30
PC 11	0.06	0.50	99.79
PC 12	0.02	0.21	100.00

https://doi.org/10.37992/2020.1101.031

First two PC's were used to construct the biplot in which 102 genotypes are scattering apart (**Fig. 1**). Genotypes which are closer to the origin and closer to each other are said to have more similarity and genotypes apart from each other are more divergent (Sharma *et al.,* 2016). In the present study, genotypes *viz.,* VCP-12-008, PG-CP-1, TY 1145, ACM 0505 are scattered apart in all the four quadrates of the biplot representing maximum genetic divergence among the genotypes. Genotypes like ACM 008, CP 30, CP 211 were closer to the origin and closer to each other indicates that low genetic divergence among them. Contribution of various PC towards total variation was portrayed in the **Fig. 2.**

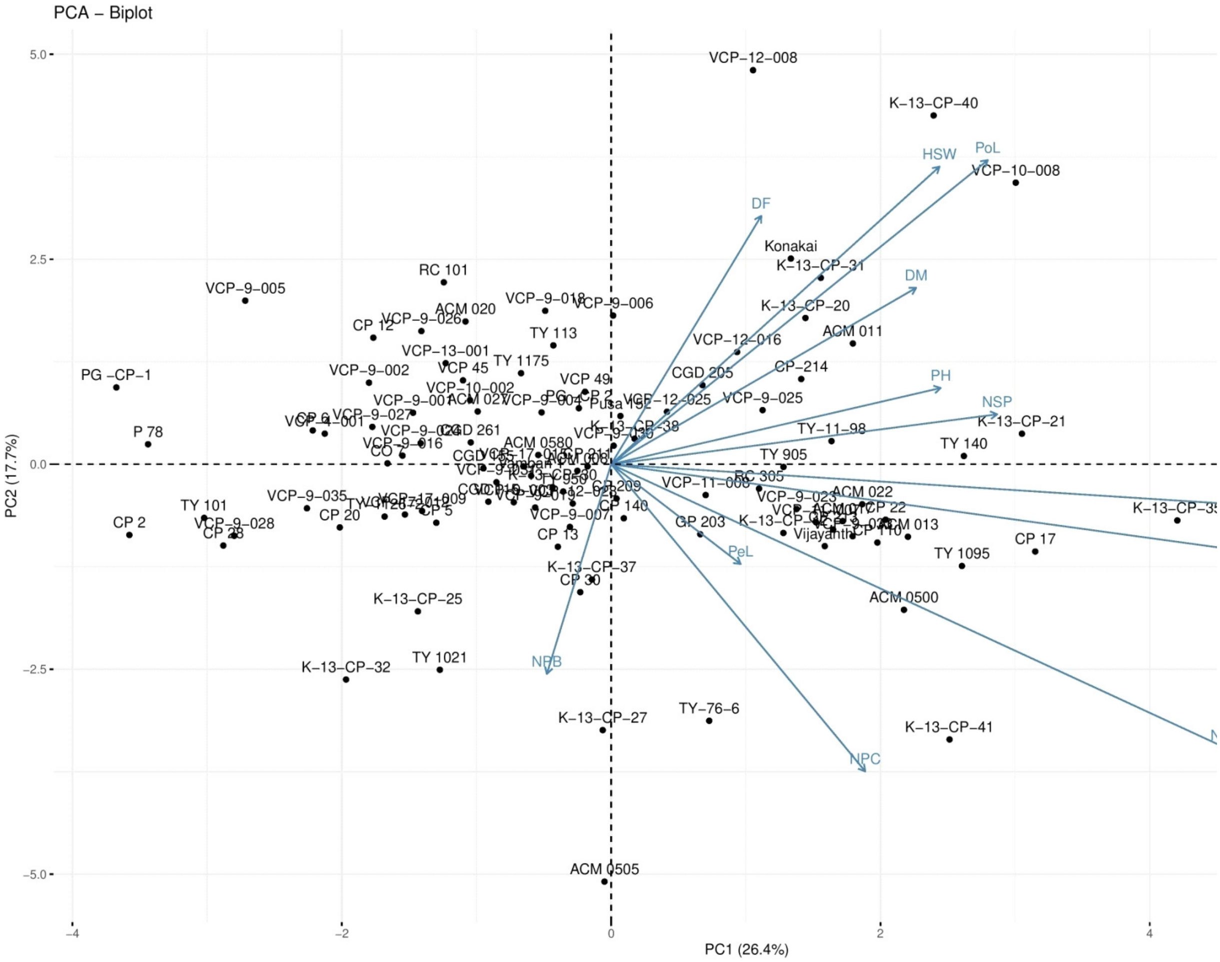

Fig. 1. Biplot showing variation among 102 cowpea genotypes along with 12 quantitative characters.

Interaction between the two variables can be well studied using the squared cosine values (Balestriero, 2017). The cosine values helps in capturing the common variables based on correlation and covariance and represent it geometrically (Shi *et al.,* 2018). In the present study, squared cosine variables based on twelve quantitative characters and five major principal components are presented in the **Fig. 3**. Single plant yield posses the highest absolute value in the first principal component depicts genetic divergence among the genotypes was mainly based on single plant yield. Higher the absolute value in the principal components higher the contribution of characters towards the divergence (Singh *et al.,* 2017). Single plant yield followed by number of clusters per plant, number of pods per plant, number of seeds per pod, pod length, plant height, hundred seed weight, days to maturity contributes maximum towards divergence in the first PC (Arora 2018, Walle *et al.,* 2019). Second PC which accounted for 17.7 per cent of variation and it was contributed by characters like number of clusters per plant, pod length, hundred seed weight, number of pods per plant and days to fifty per cent flowering. Third PC accounted for 15.1 per cent of total variation and contributed by characters like peduncle length, days to fifty per cent flowering and number of primary branches. Fourth and fifth PC accounted for 9.3 and 8.4 per cent of variation and contributed maximum by number of primary branches and number of seeds per pod respectively. Per

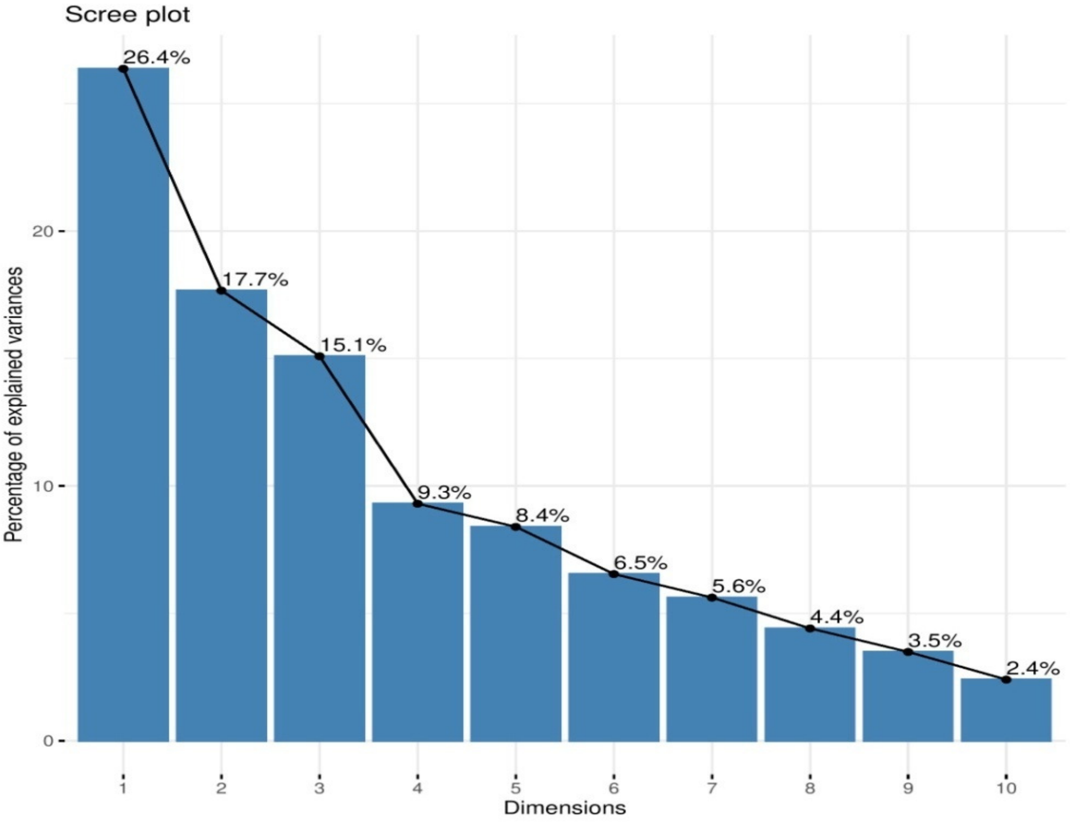

Fig. 2. scree plot showing contribution of various principal components towards divergrnce.

cent contributions of twelve quantitative characters are presented in the **table 2** which represented the contribution of characters towards the divergence.

Loadings score above ± 0.3 are considered as the significant contributors towards the divergence (Walle *et al.,* 2019).

Table 2. Per cent contributions of twelve quantitative characters towards principal components.

	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8	PC9	PC10	PC11	PC12
PH	0.979	-0.185	-0.072	0.029	-0.019	0.033	0.000	0.016	0.013	-0.005	0.018	-0.002
NPB	-0.016	0.010	-0.037	0.034	-0.048	-0.001	-0.133	-0.032	0.148	-0.907	0.363	0.001
DF	0.067	-0.006	0.453	-0.186	0.119	-0.843	0.065	0.145	-0.041	-0.059	-0.013	0.011
DM	0.076	0.190	0.841	0.172	-0.286	0.365	0.000	-0.079	-0.010	-0.006	0.014	-0.002
NC	0.041	0.236	-0.037	-0.019	-0.026	-0.079	0.287	-0.237	0.868	0.057	-0.137	0.142
NPC	0.006	0.015	-0.014	-0.029	-0.032	0.015	-0.039	0.004	-0.156	-0.031	-0.034	0.985
NPP	0.105	0.533	-0.187	-0.410	-0.425	-0.058	0.176	-0.220	-0.325	-0.175	-0.305	-0.096
PeL	-0.015	0.103	-0.192	0.784	-0.466	-0.331	0.049	0.049	-0.069	0.041	-0.007	0.001
PoL	0.033	0.069	0.036	0.238	0.430	-0.085	0.165	-0.793	-0.257	0.005	0.146	-0.004
NSP	0.026	0.061	0.021	0.083	0.083	-0.087	-0.810	-0.213	0.117	-0.058	-0.504	-0.027
HSW	0.020	0.083	0.043	0.280	0.434	0.142	0.366	0.309	-0.096	-0.332	-0.596	-0.014
SPY	0.130	0.750	-0.066	0.092	0.350	0.019	-0.209	0.306	-0.002	0.154	0.352	0.008

PH- Plant height, **DF**- Days to fifty per cent flowering, **DM**- days to maturity, **NPB**- number of primary branches, **PeL**- peduncle length, **NC**- number of clusters per plant, **NPC**- number of pods per cluster, **NPP-** number of pods per plant, **PL**- pod length, **NSP**- number of seeds per pod, **HSW**- hundred seed weight and **SPY** - single plant yield

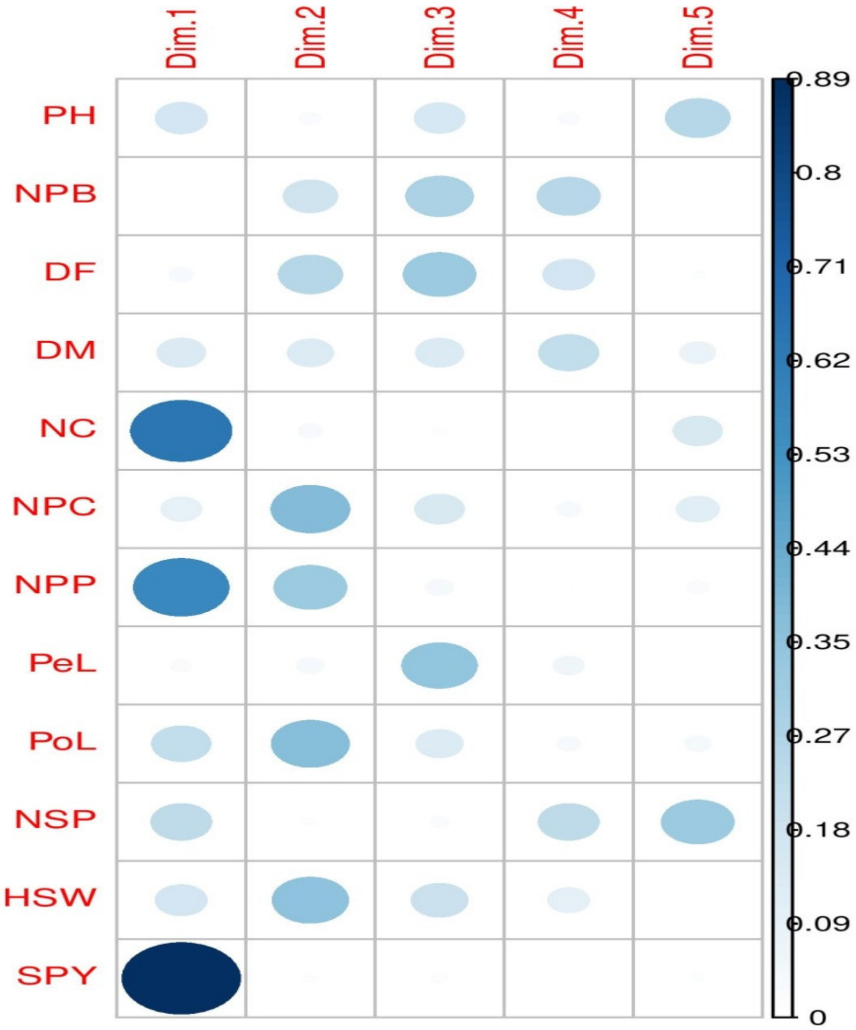

Fig. 3. Squared cosine variables of major principal components

PH- Plant height, **DF**- Days to fifty per cent flowering, **DM**- days to maturity, **NPB**- number of primary branches, **PeL**- peduncle length, **NC**- number of clusters per plant, **NPC**- number of pods per cluster, **NPP-** number of pods per plant, **PL**- pod length, **NSP**- number of seeds per pod, **HSW**- hundred seed weight and **SPY** - single plant yield

Extend of variation and relation among the quantitative characters are represented in the Fig. 4. Characters like peduncle length and number of primary branches are closer to the origin considered to have lower loading score with least contribution towards divergence and characters away from origin (single plant yield and number of pods per plant) are considered to have the highest loading score with maximum contribution towards the divergence. Characters placed in the opposite quadrants are considered to have opposite association and characters placed in the same quadrants said to have positive association (Molosiwa *et al.,* 2016). In the present study, twelve quantitative characters are placed only in three quadrants and number of primary branches had negative association with days to fifty per cent flowering which lies in opposite quadrant. Genotypes placed in the first quadrant were similar for days to fifty per cent flowering, hundred seed weight, pod length, days to maturity, plant height and number of seeds per pod as they were placed in the same quadrant. Genotypes in the second quadrants are different from each other for all the characters. Genotypes present in the third quadrant are similar for number of primary branches alone. Genotypes present in the fourth second quadrant were similar for single plant yield, number of clusters per plant, number of pods per plant, number of pods per cluster and peduncle length. Similar findings was obtained by (Lazaridi *et al.,* 2017).

Hence, the present investigation proved the existence of genetic and phenotypic variation among 102 cowpea genotypes obtained from NPRC, Vamban. This genetic

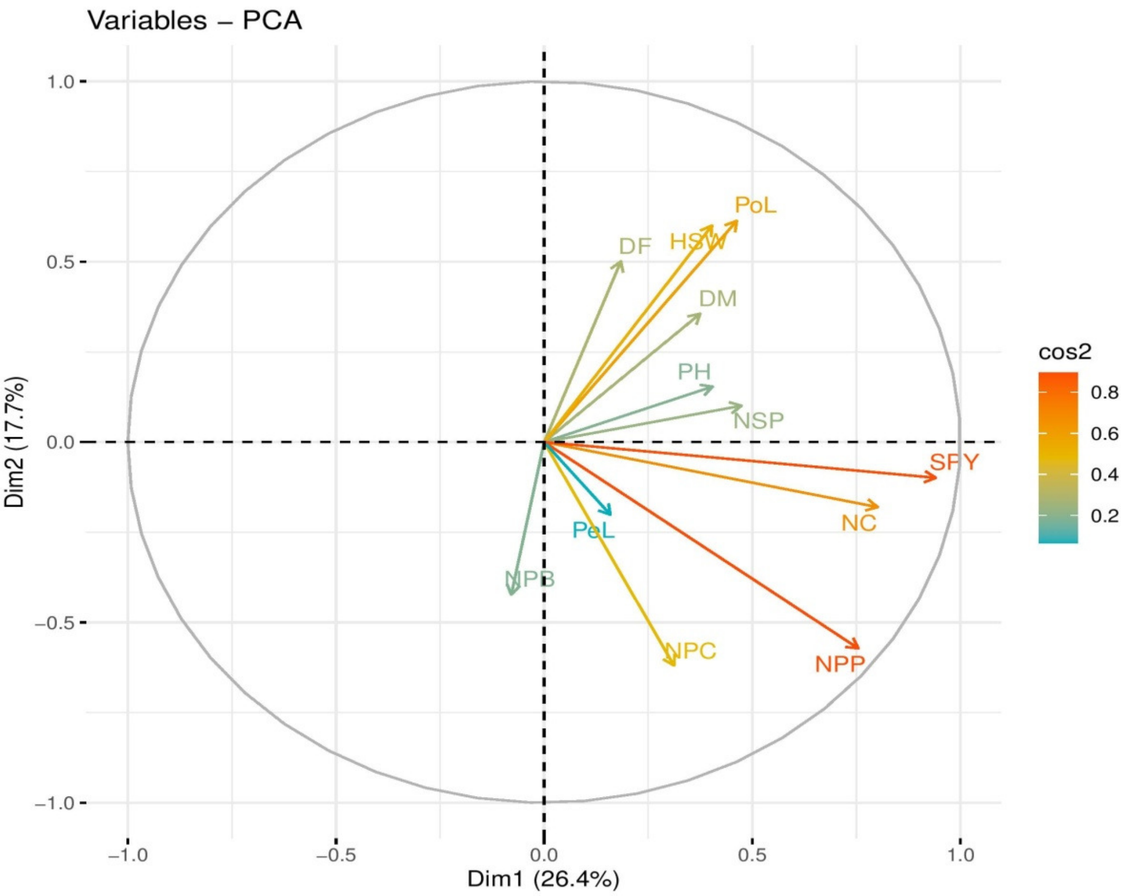

Fig. 4. Loadings plot of twelve quantitative characters based on principal components

variation promotes plenty opportunities for the genetic improvement of the cowpea through simple selection based on the novel traits and crossing potential parents.

ACKNOWLEDGEMENT

We acknowledge the Board of Research in Nuclear Sciences for providing the financial support and Dr. J. Souframanien , Program Officer, BARC for his technical assistance towards this experimental study.

REFERENCE

Abdi, Hervé, and Lynne J Williams. 2010. "Principal component analysis." *Wiley interdisciplinary reviews: computational statistics* **2** (4):433-459.

Aremu, CO, MA Adebayo, OJ Ariyo, and BB Adewale. 2007. "Classification of genetic diversity and choice of parents for hybridization in cowpea Vigna unguiculata (L.) Walp for humid savanna ecology." *African journal of biotechnology* **6** (20).

Arora, RN. 2018. "Principal component analysis in kabuli chickpea (Cicer arietinum L.)." *IJCS* 6 (2):2767-2768.

Avanza, M, B Acevedo, M Chaves, and M Añón. 2013. "Nutritional and anti-nutritional components of four cowpea varieties under thermal treatments: principal component analysis." *LWT-Food Science and Technology* **51** (1):148-157.

Balestriero, Randall. 2017. "Multiscale Residual Mixture of PCA: Dynamic Dictionaries for Optimal Basis Learning." *arXiv preprint arXiv:1707.05840*.

Boukar, Ousmane, Christian A Fatokun, Bao-Lam Huynh, Philip A Roberts, and Timothy J Close. 2016. "Genomic tools in cowpea breeding programs: status and perspectives." *Frontiers in plant science* **7**:757.

Carvalho, Márcia, Manuela Matos, Isaura Castro, Eliana Monteiro, Eduardo Rosa, Teresa Lino-Neto, and Valdemar Carnide. 2019. "Screening of worldwide cowpea collection to drought tolerant at a germination stage." *Scientia horticulturae* **247**:107-115.

Chiquet, Julien, Mahendra Mariadassou, and Stéphane Robin. 2018. "Variational inference for probabilistic Poisson PCA." *The Annals of Applied Statistics* **12** (4):2674-2698.

Fang, Jinggui, Chih-Cheng T Chao, Philip A Roberts, and Jeffrey D Ehlers. 2007. "Genetic diversity of cowpea [Vigna unguiculata (L.) Walp.] in four West African and USA breeding programs as determined by AFLP analysis." *Genetic resources and crop evolution* **54** (6):1197-1209.

Fatokun, Christian, Gezahegn Girma, Michael Abberton, Melaku Gedil, Nnanna Unachukwu, Olaniyi Oyatomi, Muyideen Yusuf, Ismail Rabbi, and Ousmane Boukar. 2018. "Genetic diversity and population structure of a mini-core subset from the world cowpea (Vigna unguiculata (L.) Walp.) germplasm collection." *Scientific reports* **8** (1):1-10.

Gerrano, Abe Shegro, Willem S Jansen van Rensburg, and Funso R Kutu. 2019. "Agronomic evaluation and identification of potential cowpea (Vigna unguiculata L. Walp) genotypes in South Africa." *Acta Agriculturae Scandinavica, Section B—Soil & Plant Science* **69** (4):295-303.

Ghalmi, Naima, Marie Malice, Jean-Marie Jacquemin, Sidi-Mohamed Ounane, Leila Mekliche, and Jean-Pierre Baudoin. 2010. "Morphological and molecular diversity within Algerian cowpea (Vigna unguiculata (L.) Walp.) landraces." *Genetic resources and crop evolution* **57** (3):371-386.

Gixhari, Belul, Michaela Pavelkova, Hairi Ismaili, Hekuran Vrapi, Alban Jaupi, and Petr Smýkal. 2014. "Genetic diversity of Albanian pea (Pisum sativum L.) landraces assessed by morphological traits and molecular markers." *Czech Journal of Genetics and Plant Breeding* **50** (2):177-184.

IBPGR. 1983. "Descriptors for Cowpea." *International Board for Plant Genetic Resources* Rome, Italy.

Jindal, Y, Rajesh Yadav, and DS Phogat. 2018. "Principal component analysis and determination of the selection criteria in fodder cowpea (Vigna unguiculata (L) Walp.) genotypes." *Range Management and Agroforestry* **39** (2):191-196.

Jolliffe, Ian T, and Jorge Cadima. 2016. "Principal component analysis: a review and recent developments." *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* 374 (2065):20150202.

Lazaridi, E, G Ntatsi, D Savvas, and PJ Bebeli. 2017. "Diversity in cowpea (Vigna unguiculata (L.) Walp.) local populations from Greece." *Genetic resources and crop evolution* **64** (7):1529-1551.

Molosiwa, Odireleng O, Chiyapo Gwafila, Joshuah Makore, and Stephen M Chite. 2016. "Phenotypic variation in cowpea (Vigna unguiculata [L.] Walp.) germplasm collection from Botswana." *International Journal of Biodiversity and Conservation* **8** (7):153-163.

Muchero, Wellington, Ndeye N Diop, Prasanna R Bhat, Raymond D Fenton, Steve Wanamaker, Marti Pottorff, Sarah Hearne, Ndiaga Cisse, Christian Fatokun, and Jeffrey D Ehlers. 2009. "A consensus genetic map of cowpea [Vigna unguiculata (L) Walp.] and synteny based on EST-derived SNPs." *Proceedings of the national academy of sciences* **106** (43):18159-18164.

Patel, UV, VK Parmar, PB Patel, and AV Malviya. 2016. "Correlation and path analysis study in cowpea (Vigna unguiculata (L.) WALP)." *International Journal of Science, Environment and Technology* **5** (6):3.897-893.904.

Prasanthi, L, B Geetha, BN Ramya Jyothi, and K Raja Reddy. 2012. "Evaluation of genetic diversity in cowpea, Vigna unguiculata (L.) Walp gentotypes using random amplified polymorphic DNA (RAPD)." *Current Biotica* **6** (1):22-31.

Ringnér, Markus. 2008. "What is principal component analysis?" *Nature biotechnology* **26** (3):303-304.

Saxena, A, and T S Rukam. 2020. "Assessment of genetic diversity in cowpea (Vigna unguiculata L. Walp.) through ISSR marker." *Research Journal of Biotechnology Vol* **15**:3.

Sharma, Hariom Kumar, Moonmoon Sarkar, Shashi Bhushan Choudhary, A Anil Kumar, RT Maruthi, Jiban Mitra, and Pran Gobinda Karmakar. 2016. "Diversity analysis based on agro-morphological traits and microsatellite based markers in global germplasm collections of roselle (Hibiscus sabdariffa L.)." *Industrial Crops and Products* **89**:303-315.

Shi, Bo, Peng Yu, Chongjin Zhao, Luolei Zhang, and Hui Yang. 2018. "Linear correlation constrained joint inversion using squared cosine similarity of regional residual model vectors." *Geophysical Journal International* **215** (2):1291-1307.

Singh, B, Aakansha Goswami, and Amit Kumar. 2017. "biotechnology approaches for quality improvement of vegetables." *Annals of Horticulture* **10** (1):7-14.

Sousa, CC, KJD SILVA, EA Bastos, and M de M ROCHA. 2015. "Selection of cowpea progenies with enhanced drought-tolerance traits using principal component analysis." *Embrapa Meio-Norte-Artigo em periódico indexado (ALICE).*

Vavilapalli, Sivakumar, VA Celine, Shrishail Duggi, Sanjeev Padakipatil, and Santoshkumar Magadum. 2013. “Genetic variability and heritability studies in bush cowpea (Vigna unguiculata (L.) Walp.).” *Legume Genomics and Genetics* 4.

Walle, Tesfaye, Firew Mekbib, Berhanu Amsalu, and Melaku Gedil. 2019. “Genetic diversity of Ethiopian cowpea [Vigna unguiculata (L) Walp] genotypes using multivariate analyses.” *Ethiopian Journal of Agricultural Sciences* **29** (3):89-104.

Xiong, Haizheng, Ainong Shi, Beiquan Mou, Jun Qin, Dennis Motes, Weiguo Lu, Jianbing Ma, Yuejin Weng, Wei Yang, and Dianxing Wu. 2016. “Genetic diversity and population structure of cowpea (Vigna unguiculata L. Walp).” *PLoS One* **11** (8).