Received: 23 Apr 2021 | Revised: 13 May 2021 | Accepted: 21 Jun 2021

Electronic Journal of Plant Breeding

Research Article

Genetic divergence and interrelationships for yield and yield attributing traits in horsegram

G. Neelima[1*], C. V. Sameer Kumar[2], K. Avil Kumar[3] and V. Nataraj[4]

[1]Regional Agricultural Research Station, Palem, Nagarkurnool (District), 509215 Telangana State. PJTSAU, India
[2]Department of Genetics & Plant Breeding, Professor JayashankarTelangana State Agricultural University, Hyderabad, Telangana, India
[3]Department of Agronomy, Professor JayashankarTelangana State Agricultural University, Hyderabad, Telangana, India[1]
[4]ICAR-Indian Institute of Soybean Research, Indore, 452001, M.P, India
[*]**E-Mail:** neelimaneelu107@gmail.com

Abstract

The present experiment was carried out to estimate the genetic diversity by cluster and Principal Component analyses for seven yield and yield attributing traits in 50 horsegram germplasm accessions and two checks during the late *kharif,* 2019. In correlation analysis, the grain yield was positively and significantly associated with test weight and number of pods which suggested the possibilities of improving grain yield by simultaneous improvement of these traits. Using cluster analysis, the accessions were grouped into five clusters. Based on the average cluster mean, the first cluster had high mean values for grain yield, the number of pods, test weight and plant height. The results of PCA revealed that, PC1, PC2, PC3 having eigenvalue greater than 1 contributed to the total variability of 82.69 per cent. Horsegram genetic improvement is receiving less attention from researchers across the globe. Using the results of this study, there is a tremendous space for genetic enhancement in grain yield of horsegram.

Key words

Genetic divergence, PCA, Correlation, Eigenvalue, Horsegram

INTRODUCTION

Horsegram [*Macrotyloma uniflorum* Lam. (Verdc.)] is an underutilized and unexplored food legume (Reddy *et al*., 2008). It is considered as a good source of protein, carbohydrates and energy (Bravo *et al*., 1998). It is tolerant to drought (Bhardwaj and Yadav, 2012), salinity (Reddy *et al*., 1998) and heavy metal stresses (Reddy *et al*., 2005). In addition, due to its genome plasticity the crop is suitable for its cultivation in low input lands (Priyanka, *et al*., 2019). Horsegram mainly grown in India, Africa, Australia, Myanmar, Malaysia, Mauritius, and the West Indies under low soil fertility status with few inputs (Witcombe*et al*., 2008). It is adapted to wide range of temperature regimes (Smartt, 1985) where other crops invariably fail to survive. In India, it is generally sown late in the rainy season by resource-poor farmers in marginal and drought-prone condition.

M. uniflorum is a short day and day neutral plant (Cook *et al*., 2005). In India, Horsegram is grown annually on nearly 4.6 l. ha (*Indiastat*, 018–2019) of land and constitutes approx. 1–2 per cent of total pulses area mainly in states such as Karnataka, Andhra Pradesh, Tamil Nadu, Odisha, Maharashtra, Chhattisgarh, Bihar, Jharkhand and Madhya Pradesh(Singh, 2013). It serves as nutritious food, green leafy vegetable, sprouts, feed and fodder (Bhartiya*et al.,* 2015) owing to high protein (18–29%) content, carbohydrate (57.2%), minerals (3.2%) comprising calcium (287 mg/100 g), iron (8.4 mg/100 g) and phosphorus (311 mg/100 g), crude fibre (5.3%) and vitamins like thiamine (0.42 mg), riboflavin (0.2 mg), niacin (1.5 mg) and vitamin C (1.0 mg/100 g) (Gopalan*et al*., 2006). Despite the presence of many significant properties in the species, the area and production could

https://doi.org/10.37992/2021.1202.065

not be increased due to its poor plant architecture with many wild characteristics such as indeterminate twining growth habit, late flowering, asynchronous maturity and poor harvest index resulting in unsuitability of this crop for modern farming system (Henry *et al*., 2006). The tapering diversity in crop species contributing to the world's food supplies has been considered a potential threat to food security (Khoury*et al*., 2014) and in the present era, this dependency on few crops for ensuring food and nutritional security has been challenged with a greater need for diversification of rural cropping systems through nonconventional underutilized crops as potential future crops (Mabhaudhi*et al*., 2017).

In the process of genetic improvement of any crop, genetic diversity among germplasm plays a major role, since it opens the way to determine the most divergent parents based on the contribution of different qualitative and quantitative traits for further utilization in any breeding programme. Therefore, exploitation of variability in the available germplasm is a pre-requisite in a breeding programme for effective selection of the superior genotypes. The present experiment was conducted to appraise the Horsegram germplasm for correlation variability, cluster analysis and principal component analysis based on the genetic analysis of yield and attributing traits of 50 Horsegram germplasm accessions.

MATERIALS AND METHODS

The present investigation consisting of fifty (50) Horsegram germplasm accessions and two local checks (Palem 1 and Palem 2) was conducted at Regional Agricultural Research Station, Palem, Telangana during the late *kharif,* 2019 in Augmented Block Design. Each genotype was raised in two rows of 4 m length with a spacing of 10 cm between the plants and 45 cm between the rows. Recommended agronomic practices were adopted to raise good crop.

Data were recorded on five randomly selected plants in each entry for seven quantitative traits *viz*., days to 50% flowering, days to maturity, plant height (cm), the number of branches per plant, the number of pods per plant, test weight (g) and grain yield (kg/ha). After harvest, hundred seed weight was calculated by weighing of counted 100 seeds for each accession. Seed yields were calculated after harvesting collectively the 3.6 m^{2}plot and the data were transformed to kg/ ha.

Augmented design analysis, correlation, PCA and K means clustering were analyzed in R version 3.6.0. Adjusted means for all the traits under considerations of all the 50 germplasm accessions were estimated by using the R package 'augmented RCBD' (Aravind*et al*., 2018). Correlation plots were obtained by using the R packages 'corrplot' (Wei and Simko,2017) and 'Performance Analytics' Peterson and Carl(2018). Principal Component Analysis (PCA) was performed by using R packages, 'factoextra' (Kassambara and Mundt, 2017), 'FactoMineR' (Le *et al*., 2008), 'devtools' (Wikham*et al*.,2019) and 'ggbiplot' (Vu, 2011). K means clustering was carried out using R packages, 'factoextra' (Kassambara and Mundt,2017), 'tidyverse' (Wickham,2017) and 'cluster' (Maechler*et al*., 2019). Optimum number of clusters was determined by using elbow method.

RESULTS AND DISCUSSION

Correlation matrix showed positive correlations in sky blue colour while the negative associations in red color. The color intensity and the size of the circle are proportional to the correlation coefficients (**Fig.1**). The correlation matrix of grain yield showed a significant positive association with the number of pods per plant (0.96), test weight (0.52), the number of branches per plant (0.39), plant height (0.38) and days to 50% flowering (0.15) and no trait was negatively associated with grain yield and the results were in accordance with (Rakeshalle*et al*., 2016). Similarly, the significant positive association was observed for traits like days to maturity with days to 50% flowering (0.74), test weight with the number of pods per plant (0.53), plant height with days to flowering (0.49) and days to maturity (0.38), the number of branches per plant with plant height (0.45) which are similar to the results of (Mukesh*et al*., 2020). However, the negative association was observed for a number of branches per plant with days to maturity (-0.051).

In (**Fig.2**), the distribution of each variable is exhibited on the diagonal. On the lower side, the vicariate schematic model plots with a line of best fit are shown. Above the diagonal, the values of the correlation are marked with the significance level as stars. The correlation matrix is used to investigate the dependence between multiple variables at the same time. The result showed the correlation coefficients between each variable with the others. The present study also testifies the robustness of similar methods and in line with the previous findings obtained by (Shivakumar*et al.,*2019).

A heat model map diagrammatically represents the data values in different colors. Actually, the heat map was originated by trademarked software, namely designer Cormac Kinney in (1991), to demonstrate a 2D matrix portraying market information, through the likely plots *viz*., as shading matrices have subsist for over a century. In the heat map of correlation, grain yield was highly correlated with the number of pods per plant similarly days to flowering was highly significant with days to maturity. The traits like days to 50% flowering, days to maturity, plant height, the number of pods per plant, the number of branches per plant, test weight and grain yield were significantly correlated with each other **(Fig.3).**

Cluster analysis is a method that aims to clarify and classify a sample of objects based on a set of deliberately chosen variables into a number of different groups

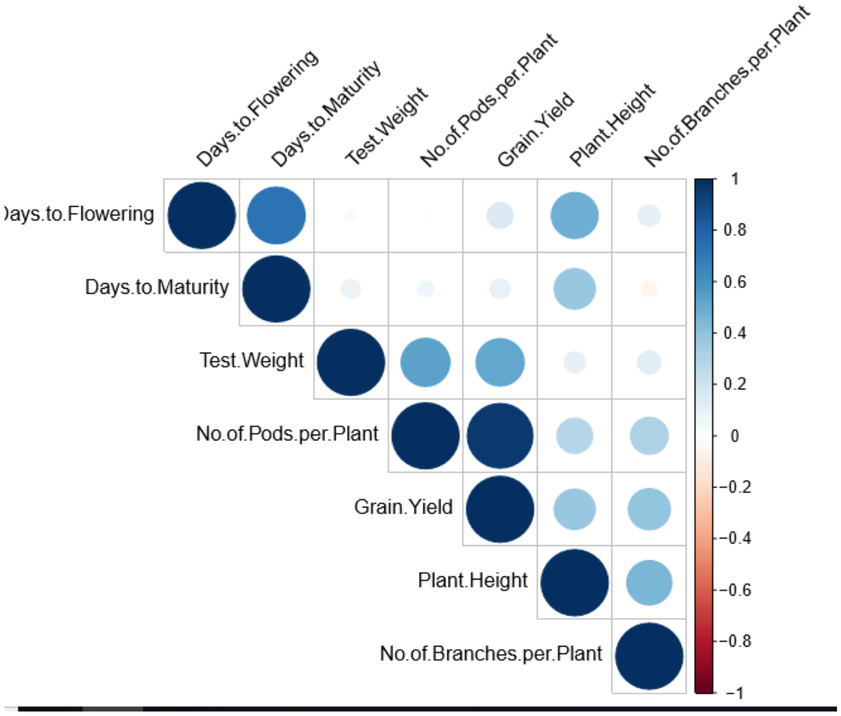

Fig. 1. Correlation Matrix showing character's correlation with each other. A . Positive correlation is shown in blue and negative correlations in shown in red color. The color intensiveness and the size of the circle are relatively proportional to the correlation.

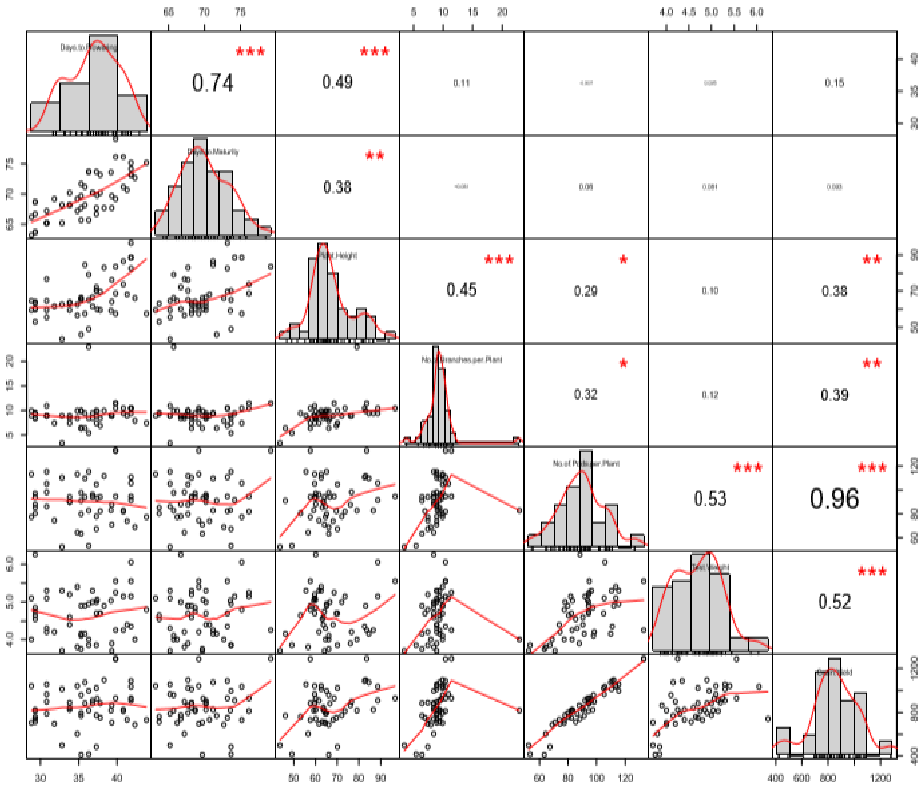

Fig. 2. Schematic model of plot showing the significance of traits with each other as well as other traits

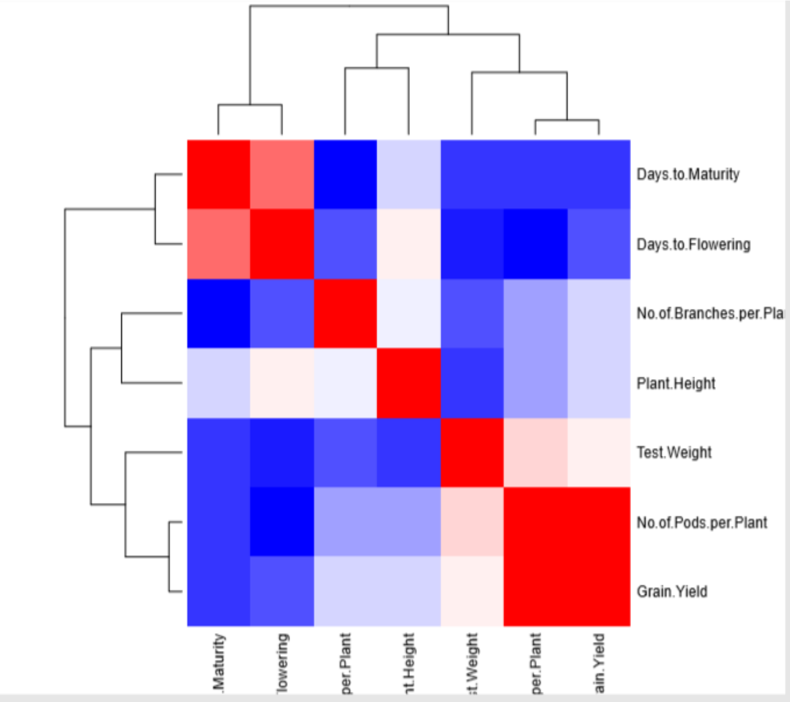

Fig. 3. A heat map of horse gram germplasm accessions showing the relationships between yield and yield related traits.

keeping alike objects in the similar group. Diversity among the genotypes was also estimated based on the considerable amount of variation in cluster means for different characters. Different clusters exhibited the distinct mean values for almost all the seven characters which reflect the genetic differences between the clusters (**Table 1**). Using cluster analysis, all the 50 germplasm accessions and two checks were grouped into 5 clusters based on different parameters (**Fig.4**). The highest number of accessions (25) were distributed into cluster 3 followed by cluster 1 (10), cluster 2 (9), cluster 4 (7) and cluster 5 (1). Cluster-1 had genotypes with significant values for plant height (1.19), grain yield (1.16) and the no. of pods per plant (1.09). In similar manner cluster-2 was composed of genotypes holding significant values for days to flowering (1.09) while cluster-5 had highest values for the number of branches (5.62) and for days to flowering (1.12). However, in cluster-4 low values were observed for grain yield (-1.61) followed by the no. of pods per plant (1.39) and test weight (1.27) and the similar results were obtained by Ameer*et al*. (2019). Comparative assessment of cluster means showed for improving specific characters, the genotypes should be selected from the cluster having high mean value for that particular character. This comparison indicates that clusters I had better cluster means for most of the characters, therefore, this cluster might be considered better for selecting genotypes as divergent parents. The similar results are exhibited with the findings of (Kumar*et al*., 2006; Shrivastava*et al*., 2012;Shalini*et al*.,2019; and Shivakumar*et al*., 2019).

It is one of a series of techniques for collecting high-dimensional information and using the dependence between the variables in a more tractable form without any loss of information. It represents the major contributor to the total difference in each differentiation axis (Shalini*et al.,*2019) and (Shivakumar*et al.,*2019). The most important factor is to distribute the total variance into the different factors, which is considered as a potential tool for applicable parental lines for a useful breeding programme (Akter*et al*., 2009). The genetic diversity plays a vital role in selection of varieties for hybridization (Vishwanatha*et al.,* 2016). Based on the analysis, three out of five component factors (PCs) were found with an eigenvalue of >1 (**Table.2**).These three principal factors (PCs) contributed to a total of 82.69 per cent of variance. The first component (PC1) accounted for 40.9 per cent of variation which was largely through grain yield (27.36), the number of pods (23.69) and plant height (15.26); PC2 accounted for 26.63 per cent of variation contributed through days to flowering (35.13) and days to maturity (32.17) and PC3 contributed 15.16 per cent through the number of branches (49.75) and test weight (20.57). Whereas, PC4 and PC5contributed to7.74 and 5.67 per cent, respectively. Similarly, biplot diagram of PC1 and PC2 (**Fig.5& 6**) demonstrated that grain yield, the number of pods per plant and days to flowering contributed significantly towards the diversity in germplasm. The contribution of traits (**Fig.7**) like grain yield, the number of pods per plant, days to flowering towards variability was highest (approx. 18% each) and

Table 1. Cluster mean values for seven traits

Cluster number	Daysto50% flowering	Days to maturity	Plant height	Number of branches	Number of pods	Test weight	Grainyield
1	0.90	0.99	1.19	0.43	1.09	0.56	1.16
2	1.09	0.71	0.21	-0.19	-0.69	-0.39	-0.44
3	-0.68	-0.64	-0.34	-0.02	0.22	0.32	0.15
4	-0.26	0.10	-0.94	-1.09	-1.39	-1.27	-1.61
5	1.12	0.09	-1.09	5.62	-0.47	-1.11	-0.18

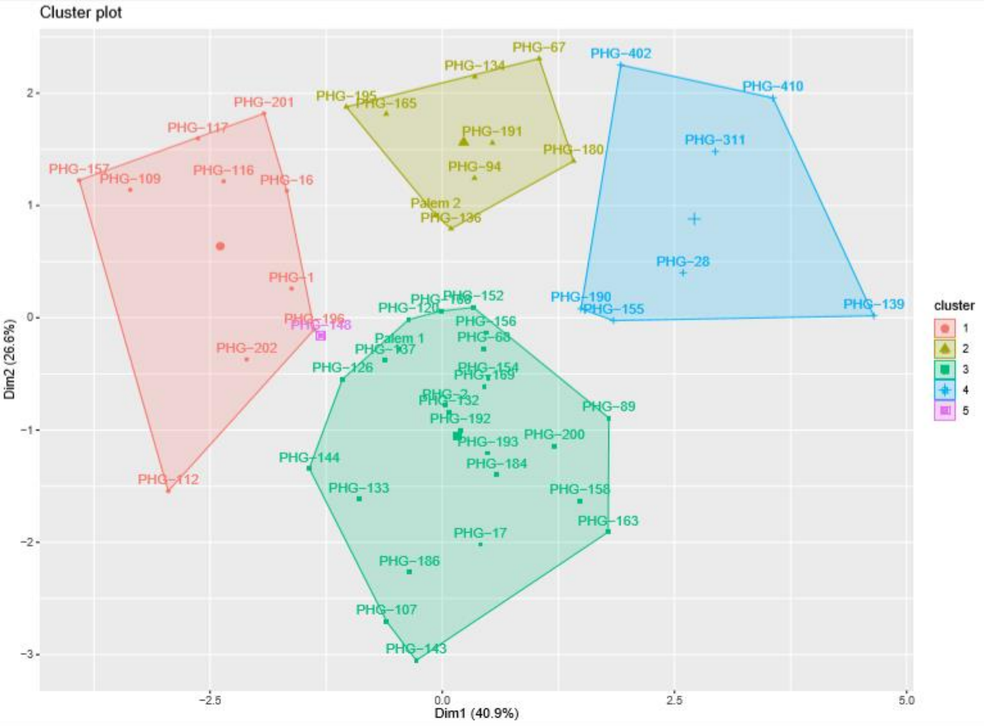

Fig. 4. K-means clustering pattern of genotypes (k=5)

Table 2. Eigenvalues, variance contribution (%), variables coordinates and percent contribution of each variable on individual five principal components of 50 horse gram germplasm.

		Days to 50% flowering	Days to maturity	Plant height	Number of branches per plant	Number of pods per plant	Test weight	Grain yield	Eigen values	Variance (%)
PC1	Coordinates	0.44	0.38	0.66	0.54	0.82	0.58	0.88	2.86	40.90
	Contribution (%)	6.66	5.27	15.26	10.17	23.69	11.58	27.36		
PC2	Coordinates	0.81	0.77	0.41	-0.10	-0.44	-0.35	-0.34	1.86	26.63
	Contribution (%)	35.13	32.17	8.82	0.52	10.59	6.47	6.29		
PC3	Coordinates	0.11	0.34	-0.39	-0.73	0.15	0.48	0.09	1.08	15.16
	Contribution (%)	1.04	11.01	14.63	49.75	2.29	20.57	0.69		
PC4	Coordinates	0.03	-0.03	0.003	0.26	-0.29	0.56	-0.26	0.54	7.74
	Contribution (%)	0.13	0.12	0.002	12.49	15.63	59.08	12.54		
PC5	Coordinates	0.16	0.14	-0.49	0.31	0.04	-0.09	0.06	0.39	5.67
	Contribution (%)	6.42	4.80	60.90	24.33	0.33	2.15	1.05		

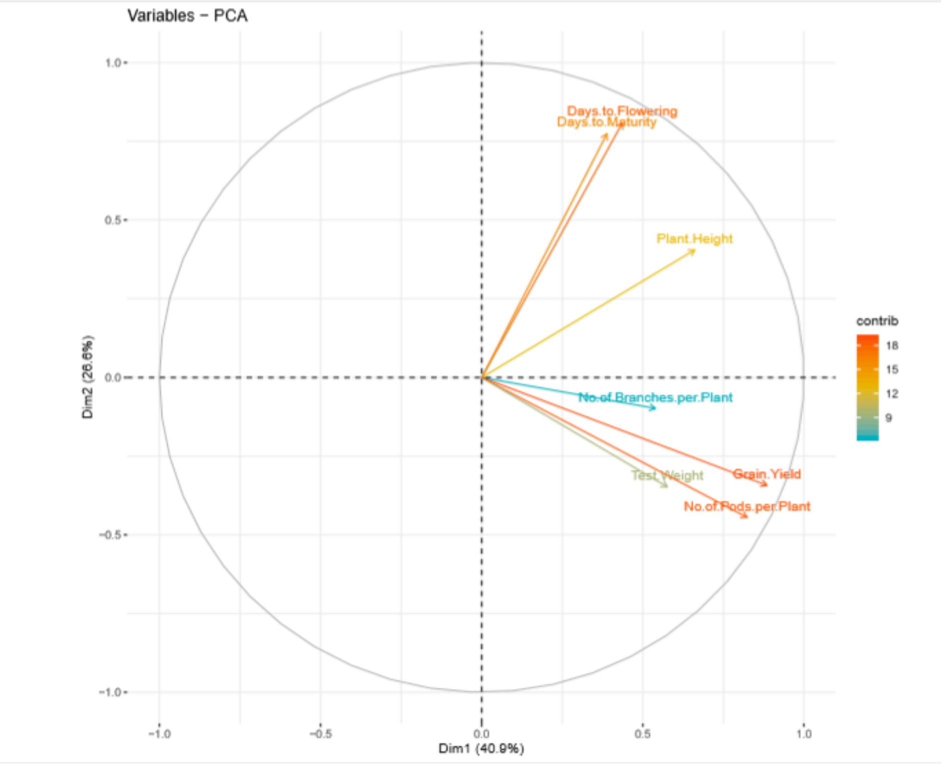

Fig. 5. Biploting of different variables loaded on PC1 and PC2 components.

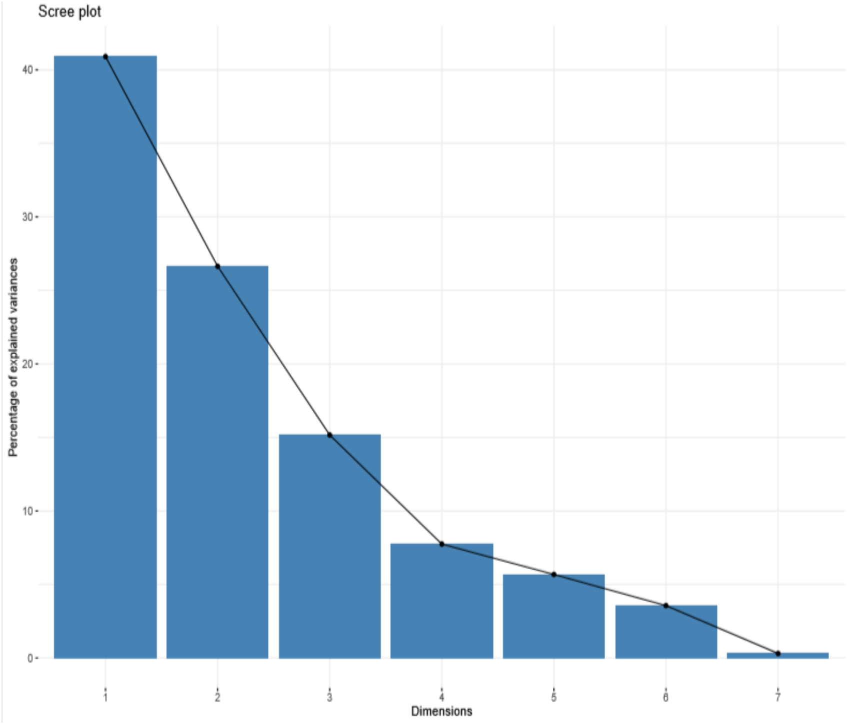

Fig. 6. Screen plot showing variance and along with Principal components

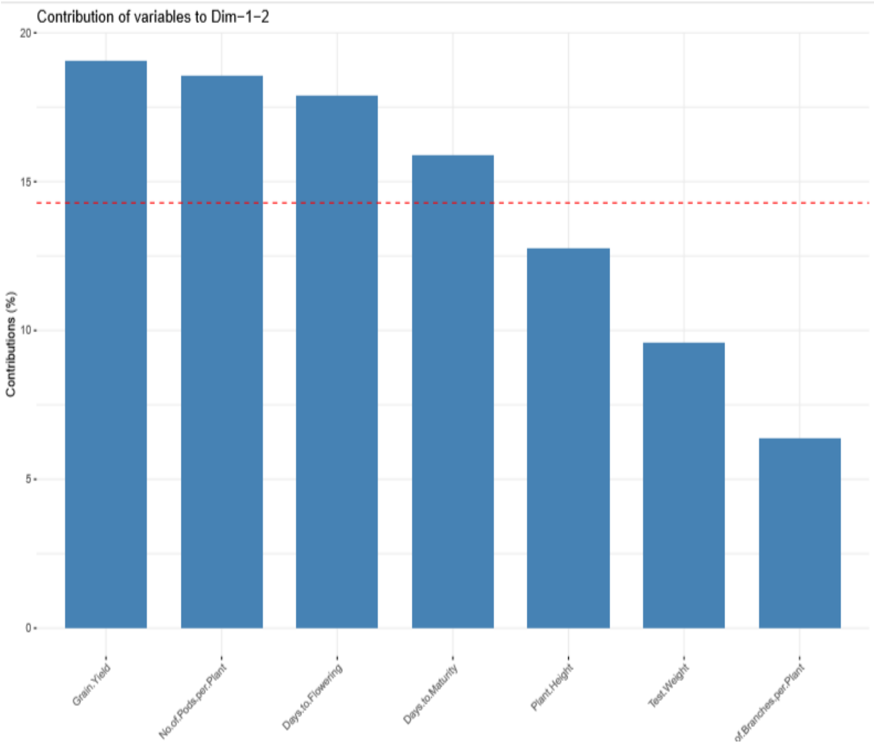

Fig. 7. Contribution of traits to dimensions 1 and 2

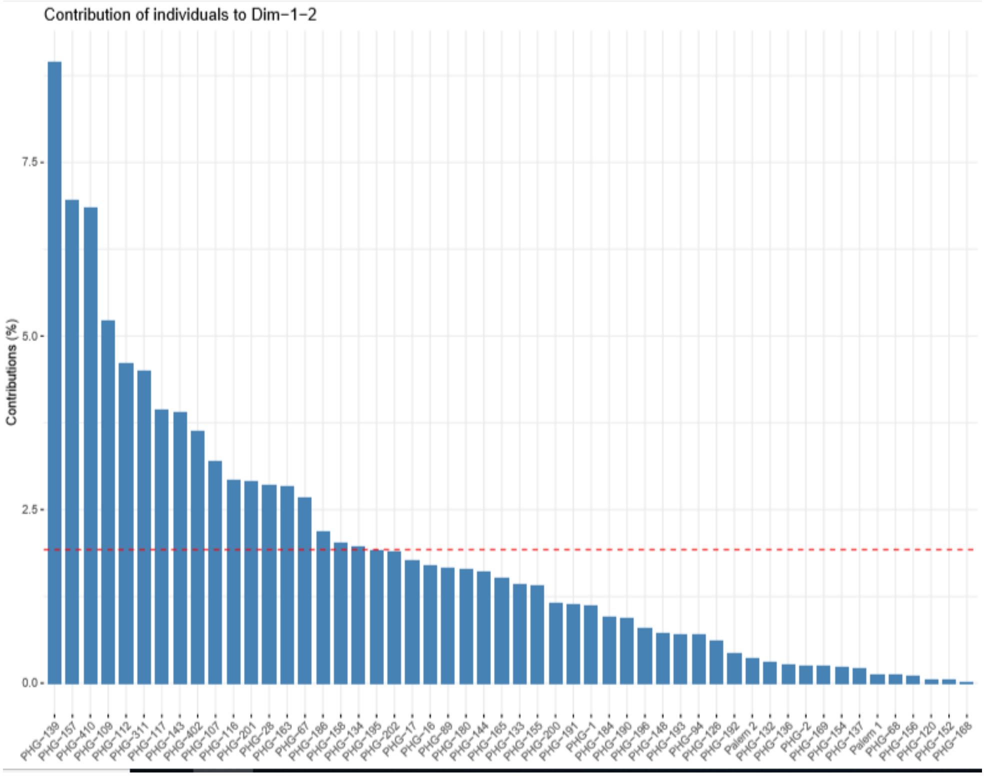

Fig. 8. Contribution of genotypes to dimensions 1 and 2

the genotypes *i.e.,* PHG-139 (>8%) followed by PHG-157 and PHG-410 of approximately 7 per cent each was contributed towards variability to dimension 1 and 2 (**Fig.8**).Thus the variability of traits in the study bearing on the eigenvalue of the PC for the specific traits, such as the maximum eigenvalue is responsible for wide variance for the traits under consideration. These conclusions are also in correspondence with that of (Nazir*et al*., 2013) who concluded that the contribution between first two PCs are very important in the total variation. Exploitation of genotypic resources could be made by assigning the total variation into number of factors; however, it also provides a chance to use very important and appropriate germplasm in future for the crop improvement program to amend the yield production for specific plant traits (Ameer *et al*., 2019).

In the correlation analysis, the grain yield was positively and significantly correlated with the number of pods per plant (0.96) and test weight (0.52), which suggested possibilities for improving the grain yield by simultaneous improvement of these traits. In distribution of genotypes into 5 clusters, the cluster I had better cluster means for most of the characters, therefore, this cluster might be considered better for selecting genotypes *i.e.,* PHG-157, PHG-109, PHG-112, PHG-117 and PHG-201 as divergent parents. The cluster II had better mean performance for plant height and the genotype PHG-67 exhibited a maximum variability in dimension 1 & 2. The cluster III had good mean performances for the number of pods, test weight and grain yield and comprises of genotypes like PHG-143, PHG-107, PHG-163, PHG-186 and PHG-158. Among all the PCs, PC1, PC2 & PC3 have contributed to a total of 82.69% variation. Overall, characters including grain yield, the number of pods, days to flowering and days to maturity exhibited higher contribution towards variation and can be further used as a selection criteria to increase the grain yield. Therefore, these positive and significant association and divergent parental genotypes (**Fig.8**) notify through Principal Component Analysis (PCA) would be more effective to get breakthrough in successful breeding program.

The selected genotypes from different clusters with extreme characters will be used as parents for hybridization and developing varieties with superior yields and development of mapping population for linkage mapping and identification and validation of markers linked with higher yields.

REFERENCES

Akter, A., Hasan, M.J., Paul, A.K.,Mutlib, M. and Hossain, M.K. 2009. Selection of parent for improvement of restorer line in rice (*Oryza sativa* L.). *SAARC J. of Agriculture,***7**: 43-50.

Ameer, H. J., Xiaoyan,W.,Muhammad, S.I.,Zareen, S., Long, W.,Qifeng, Q. and Fan, S. 2019.Genetic divergence on the basis of principal component, correlation and cluster analysis of yield and quality traits in cotton cultivars. *Pak.J. of Bot.,***51**: 1143-1148. [Cross Ref]

Aravind, J., Sankar, M.S.,Dhammaprakash, P.W. and Kaur, V. 2018.Augmented RCBD: Analysis of Augmented Randomised Complete Block Designs. R package version 0.1.0, https:// aravind-j.github.io/augmentedRCBD/.

Bhardwaj, J. andYadav, S.K. 2012. Comparative study on biochemical and antioxidant enzymes in a drought tolerant and sensitive variety of horsegram [*Macrotyloma uniflorum* (Lam.) Verdc.].) under drought stress. *American J. of Plant Physiology,***7**:17–29. [Cross Ref]

Bhartiya, A., Aditya, J.P.and Kant, L.2015. Nutritional and remedial potential of an underutilized food legume horsegram [*Macrotyloma uniflorum* (Lam.) Verdc.] :a review. *J. of Animal and Plant Sci.,***25**:908–920.

Bravo, L., Siddhuraju, P.andSaura-Calixto, F. 1998. Effectofvarious processing methods on the in vitro starch digestibility and resistant starch content of Indian pulses. *Journal of Agric. and Food Chem.,***46**:4667–4674. [Cross Ref]

Cook, B.G., Pengelly, B.C., Brown, S.D., Donnelly, J.L., Eagles,D.A.,Franco, M.A., Hanson, J., Mullen, B.F.,Partridge, I.J., Peters, M. and SchultzeKraft, R.2005. Tropical Forages: an interactive selection tool. [CD-ROM], CSIRO, DPI&F (QLD), CIAT and ILRI, Brisbane, Australia.

Gopalan, C., Ramasastry, B.V. and Balasubramanian, S.C. 2006. Nutritive value of Indian foods (revised and updated). National Institute of Nutrition, Hyderabad, pp: 156.

Henry, A.,Vishwanatha, K.P., Sharma, S.K. and Thakur, H.L. 2006. Genetic Improvement. In: Kumar D (ed) Horsegram in India. Scientific Publishers, India pp:29–35.

Kassambara,A. and Mundt, F. 2017.Factoextra: Extract and Visualize the Results of Multivariate Data Analyses. R package version 1.0.5. https://CRAN.R-project.org/package=factoextra.

Khoury, C.K., Bjorkman, A.D.,Dempewolf, H., Villegas, J.R.,Guarino,L.,Jarvis, A.,Rieseberg, L.H.andPaul, C.S. 2014. Increasing homogeneity in global food supplies and the implications for food security. *Proceedings of the National Academy of Sciences,***111**: 4001–4006. [Cross Ref]

Kumar, R., Dhari, R.and Kumar, R. 2006. Divergence studies in pea germplasm (*Pisum sativum* L.). *National Journal of Plant Improvemnet,***8**:122-124.

Le, S., Josse, J. andHusson, F. 2008. Facto Mine R: An R Package for Multivariate Analysis. *Journal of Statistical Software,***25**: 1-18. [Cross Ref]

Mabhaudhi, T., Chimonyo, V.G.P.,Chibarabada, T.P. and Modi, A.T. 2017. Developing a roadmap for improving neglected and underutilized crops: a case study of South africa. *Frontiers in Plant Sci.* 8:2143. [Cross Ref]

Maechler, M., Rousseeuw, P.,Struyf, A., Hubert, M.and Hornik, K. 2019.Cluster:Cluster Analysis Basics and Extensions. R package version 2.1.0.

Mukesh, K.S., Ruth, E.E., Sinha, S.K., Vivek, K.S., Ved, P.S. and Birjhu. 2020. Correlation and path analysis in Horsegram[*Macrotyloma uniflorum* (Lam.) Verdc.]. *Indian Journal of Current Microbiology and Applied Sciences,***9**(11):604-610. [Cross Ref]

Nazir, A., Farooq, J., Mahmood, A., Shahid, M. and Riaz, M. 2013. Estimation of genetic diversity for CLCuV, earliness and fiber quality traits using various statistical procedures in different crosses of *Gossypium hirsutum* L. *Vestnik. Orel. Gau.,***43**(4):2-9.

Peterson, B.G. and Carl, P. 2018. PerformanceAnalytics: Econometric Tools for Performance and Risk Analysis.RPackage version 1.5.2. https://CRAN.R-project.org/ package=PerformanceAnalytics.

Priyanka, S., Sudhagar, R., Vanniarajan, C., Ganeshamurthy, K. and Sougframanian, J. 2019.Combined mutagenic ability of gamma ray and EMS in horsegram (Macrotyloma uniflorum(Lam) Verdc. *Electronic Journal of Plant Breeding,***10**(3):1086-1094. [Cross Ref]

Rakesh, A., Hemalatha, V., Eshwari, K.B. and Swarnalatha, V. 2016. Genetic variability, correlation and path analysis for yield and its components in horsegram[*Macrotyloma uniflorum* (Lam.) Verdc.]. *Green farming,***7**: 1-4.

Reddy, A.M., Kumar, S.G.,Jyothsnakumari, G.,Thimmanaik, S. and Sudhakar, C. 2005. Lead induced changes in antioxidant metabolism of horsegram [*Macrotyloma uniflorum* (Lam.) Verdc.] and bengalgram (*Cicer arietinum* L.). *Chemosphere,***60**:97–104. [Cross Ref]

Reddy, P.C.O.,Sairanganayakulu, G., Thippeswamy, M., Sudhakar,R.,Reddy, P.andSudhakar, C. 2008. Identification of stress-induced genes from the drought tolerant semi-arid legume crop horsegram (*Macrotyloma uniflorum* (Lam.) Verdc.) through analysis of subtracted expressed sequence tags. *Plant Science,***175**:372–384. [Cross Ref]

Reddy, P.S., Ramanjulu, S.,Sudhakar, C.and Veeranjaneyulu, K. 1998. Differential sensitivity of stomatal and non-stomatal components of NaCl or Na_2SO_4 salinity in horsegram (*Macrotyloma uniflorum* (Lam.). *Photosynthetica,***35**:99–105. [Cross Ref]

Shalini, S., Rakesh, V.,Vinay, K.R.N., Singh, B. and Mukesh, K. 2019.Phenotypic Diversity of Pea Genotypes (*Pisum sativum* L.) Based on Multivariate Analysis. *Legume Res.*,**4165**:1-7.

Shivakumar, M., Giriraj, K.,Nataraj, V.,Gireesh, C., Sanjay, G.,Gyanesh, K.S.,Milind, B.R.and Devendra, P.Y. 2019. NAM population – a novel genetic resource for soybean improvement: development and characterization for yield and attributing traits. *Plant Genetic Resources Characterization and Utilization,*1–9.

Shrivastava, V., Lal, G.M., Chandra, K., Singh, C.M. 2012. Estimation of genetic distance in field pea (*Pisum sativum* L.). *Environment and Ecology,* **30**: 136-138.

Singh, R.P.2013. Status paper on pulses. Government of India ministry of agriculture, department of agriculture & cooperation, Bhopal, Madhya Pradesh, pp: 215.

Smartt,J. 1985. Evolution ofgrainlegumesII. Oldandnew worldpulses of lesser economic importance. *Experimental Agriculture.,***21**:1–18. [Cross Ref]

Vishwanatha, K.P., Yogeesh, L.N. and Amitha, K. 2016. Morphological diversity studies in horsegram[*Macrotyloma uniflorum* (Lam.) Verdc.] based on principal component analysis. *Electronic journal of plant breeding*, **7**(3): 767-770. [Cross Ref]

Vu,V.Q.2011.ggbiplot: A ggplot2 based biplot. R package version 0.55. http://github.com/vqv/ggbiplot.

Wei, T. and Simko, V. 2017. R package ‘corrplot’: Visualization of a Correlation Matrix (Version 0.84). Available from https:// github.com/taiyun/corrplot.

Wickham, H. 2017.tidyverse: Easily Install and Load the ‘Tidyverse’. R package version 1.2.1. https:// CRAN.R-project.org/package=tidyverse.

Wickham, H., Hester, J.and Chang, W. 2019.devtools: Tools to make developing R packages easier. R package version 2.0.2. https://CRAN.R-project.org/package=devtools.

Witcombe, J.R., Billore, M., Singhal,H.C., Patel, N.B., Tikka, S.B.S., Saini,D.P.,Sharma, L.K.,Sharma, R.Yadav, S.K. and Yadavendra, J.P. 2008. Improving the food security of low-resource farmers: introducing horsegram into maize-based cropping systems. *Experimental Agriculture.***44**:339– 348. [Cross Ref]