

# Electronic Journal of Plant Breeding



## Research Note

### Multivariate analysis in sorghum (*Sorghum bicolor*) germplasm for yield contributing quantitative traits

V. Santhiya<sup>1\*</sup>, D. Kavithamani<sup>1</sup>, B. Selvi<sup>1</sup>, T. Reavanth<sup>2</sup> and P. Vinoth<sup>1</sup>

<sup>1</sup>Department of Millets, Center for Plant Breeding and Genetics, Tamil Nadu Agricultural University, Coimbatore – 641003

<sup>2</sup>Department of Oilseeds, Center for Plant Breeding and Genetics, Tamil Nadu Agricultural University, Coimbatore – 641003

\*E-Mail: santhiyatnau@gmail.com

#### Abstract

Multivariate analyses was carried out for 13 yield-contributing quantitative traits observed in 102 sorghum germplasm accessions evaluated at Department of Millets, TNAU, Coimbatore during Summer, 2021. The Euclidean distances obtained were subjected to Ward clustering algorithm which grouped the genotypes into eight clusters. The clusters VI and VII were observed as the largest and cluster VIII was the smallest. Clusters II and VIII were found to possess elite germplasm with superior traits namely, flag leaf length, flag leaf breadth, flag leaf area, number of leaves, 100-seed weight and grain yield per plant. Principal component analysis revealed that yield contributing traits were separated into 13 principal components of which four components had eigen values more than one. Seventy-three percentage of total variation was observed in the first four principal components. This suggests that the first principal component is the major contributor to the total variation in the population. Traits viz., plant height, days to 50% flowering, stem diameter, flag leaf length, flag leaf breadth, flag leaf area, number of leaves, leaf length, leaf breadth and days to maturity contributed for maximum of 34 % variation in the PC1 with 4.45 eigen value. Divergence of second PC was mainly contributed by plant height, days to 50% flowering, leaf length, panicle length and days to maturity. Eigen value for PC2 was 2.63 and contributed about 20% of total variation. Contribution of divergence by PC3 was observed to be 10% with eigen value of 1.31. Traits responsible for variation in PC3 were leaf length, leaf breadth, days to maturity and hundred seed weight. Total variation contributed by PC4 was 9% and had 1.13 eigen value. Major variation contributing traits in PC4 were flag leaf length, flag leaf breadth and flag leaf area. Results of this study revealed the presence of variability for all the traits under study and identified accessions from clusters II and VIII as elite sources for trait specific improvement.

**Keywords:** Sorghum, genetic diversity, clustering, PCA, Euclidean distance

Sorghum (*Sorghum bicolor* (L.) Moench) is one of the most important dietary staple food crops in the world and was domesticated in Ethiopia, Africa. India is the largest producer and consumer of Sorghum and is popularly known to be Cholam, Jowar and Indian millet. It is a diploid crop ( $2n=2x=20$ ) belonging to Poaceae family. It grows best on heavier soils and tolerates high temperature throughout its life cycle. It can grow well in rainfed as well as in drought condition. Sorghum is referred to

as “The King of coarse cereals” due to its nutritional qualities. As an edible starchy grain, sorghum is rich in iron, protein, minerals, dietary fiber and is gluten free (Derese *et al*, 2018). In India, sorghum ranks fifth among the cereals with an acreage of approximately 4.10 million hectares and 1.16 metric tons per hectare productivity (USDA, 2021). Tamil Nadu occupies fifth position in India with the cultivable area of 0.38 million hectares, production of 0.49 million tons and productivity of 1,268

kg/ha (Agricultural Statistics at a glance, 2021). However, increasing food demand with increasing population is the major challenge for plant breeders nowadays. The average productivity of sorghum is so far lower than its potential (Gebregergs and Mekbib, 2020).

Understanding the genetic variation in sorghum accessions is very much needed to obtain promising genetic improvement in sorghum. Greater use of such diversity in sorghum breeding programs to develop cultivars with a broad genetic base will result in sustainable sorghum production (Karadi and Kajjidoni, 2019). A great level of knowledge on genetic variability among the sorghum germplasm will facilitate precision breeding and profiling the genetic diversity of sorghum creates a wider opportunity to select the highly diverged parents (Doijad *et al.*, 2016). Hybridization with substantially divergent parents creates an opportunity for gene constellation of yield in transgressive segregants. (Olsen and Wendel 2013).

The process of categorization is generally done with the help of multivariate analysis *viz.*, Principal Component Analysis, cluster analysis and discriminate analysis. Among these, principal component analysis (PCA) is a multivariate method used to study the relationship between different quantitative variables. Upadhyaya *et al.* (2019) mentioned it as the best tool to exploit the diversity in sorghum. PCA is useful for grouping accessions based on their similarity (Leonard and Peter, 2009) and it reduces large data set into the smaller ones without much loss of genetic information (Maji and Shaibu, 2012). The main objective of this research was to classify the sorghum accessions into clusters based on their similarity features with respect to the quantitative characters under study and also to generate data on their performance for plant breeders for further evaluation of the crop in particular. Hence, the current investigation was carried out to identify key traits contributing variation and to identify diverse germplasm by adopting principal component analysis and Ward's clustering method through among the sorghum accessions for yield contributing quantitative traits.

Out of 102 sorghum germplasm evaluated, 50 were obtained from Department of Plant Genetic Resources, TNAU, Coimbatore, 23 from ICAR-Indian Institute of Millets Research (ICAR-IIMR), Hyderabad and 22 from International Crops Research Institute for the Semi-Arid Tropics (ICRISAT), Patancheru along with seven checks (improved varieties of sorghum) which were obtained from the Department of Millets, TNAU, Coimbatore.

The experiment was conducted during *Summer*, 2021 in research plots of Department of Millets, Tamil Nadu Agricultural University, Coimbatore. The trial was laid out in Augmented design II. The genotypes were sown in rows of 4m length adopting a spacing of 45 cm x 15 cm. The experimental plot was divided into three blocks.

All the recommended package of practices for better crop production was followed.

Thirteen quantitative traits related to yield and its contributing characters *viz.*, plant height (cm), days to fifty percent flowering (days), stem diameter (cm), flag leaf length (cm), flag leaf breadth (cm), flag leaf area (cm), number of leaves, leaf length, leaf breadth (cm), panicle length (cm), days to maturity (days), hundred seed weight (gm) and grain yield per plant (gm) were recorded at different growth stages of sorghum.

In the present study, the dissimilarity matrix was constructed using Euclidean method and the clustering was done by following ward's minimum variance linkage method in Graphical Genotyping Software (GGT 2.0). Principal Component Analysis was conducted to identify the patterns of variation and to estimate the relative contribution of various characters towards total variability. Principal component analysis was done in R studio using the package "Factoshiny".

Genetic diversity analysis is useful in identifying diverse sorghum parental lines for hybrid breeding program. The assessment of the diversity among the germplasm may provide essential information in utilizing the genotypes in various crop improvement programs.

Diversity study on 102 sorghum germplasm for 13 quantitative traits associated with grain yield: In the present study, 13 quantitative traits recorded from 102 germplasm were subjected to cluster analysis. The genotypes were grouped into eight clusters and they are presented in **Fig. 1**. Similarly, Elangovan and Babu (2015) grouped 99 landraces into 16 clusters and Mofokeng, (2019) grouped 98 sorghum accessions into three main groups.

The clustering pattern of sorghum germplasm for 13 quantitative traits is provided in **Table 1**. Clusters, VI and VII were the largest and cluster VIII was the smallest among eight clusters. Cluster VI had 21 germplasm and this cluster was further divided into two sub-clusters. Sub-cluster VIA had 19 germplasm and sub-cluster VIB had two germplasm. Cluster VII was composed of 21 germplasm, out of which 14 were grouped in sub-cluster VII A and seven in sub-cluster VII B. Cluster VIII had two germplasm *viz.*, SOR 11218 and SPV 462. The germplasm, SOR 11218 was one of the superior germplasm with superior characters like flag leaf length, flag leaf breadth, stem diameter, number of leaves, leaf length, leaf area, leaf area index and hundred seed weight.

Cluster I had 12 germplasm with two sub-cluster and most of the check varieties were grouped in this cluster. Likewise, Cluster II had 14 germplasm with two sub-cluster in which EJM 20 recorded maximum plant height and EJM 7 recorded the longest flag leaf length and maximum flag leaf area. The germplasm EJM 7 and EJM

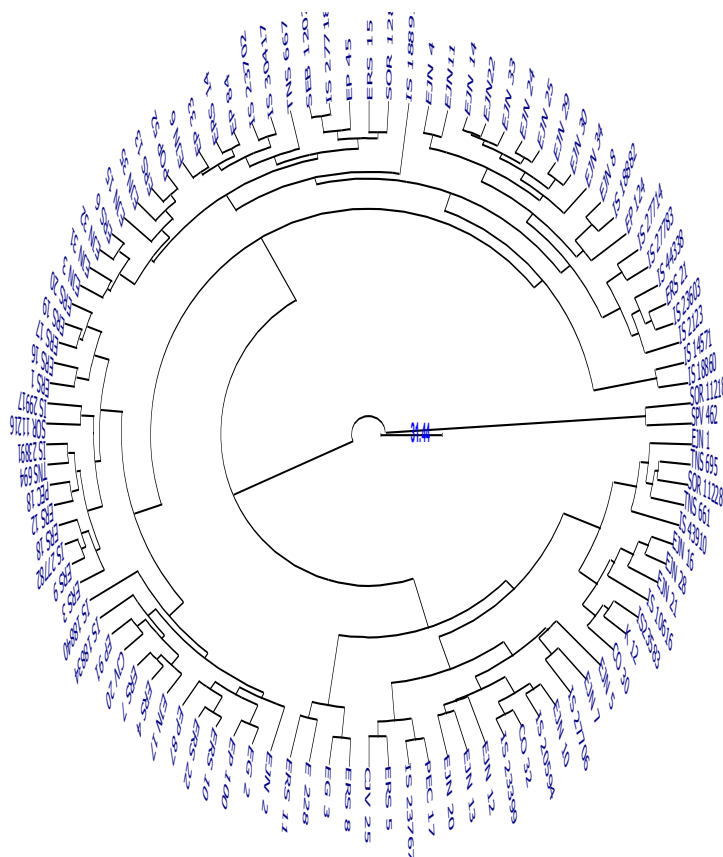


Fig. 1. Cluster classification of sorghum genotypes for thirteen morphological traits

Table 1. Cluster classification for thirteen yield contributing traits among 102 germplasm

CLUSTER	SUB CLUSTER	NO OF GERmplasm	GERMPLASM
I	IA	6	EJN 1, TNS 695, SOR 11228, TNS 661, IS 43910, EJN 16.
	IB	6	EJN 28, EJN 21, IS 10616, IS 23583, K12, CO 30
II	II A	10	EJN 5, EJN 7, IS27706, EJN 10, IS 20594, CO 32, IS 23399, EJN 12, EJN 13, EJN 20
	II B	4	PEC 17, IS 23767, ERS 5, CJV 25
III	III	4	ERS 8, EG 3, E 228, ERS 11
IV	IVA	10	EJN 2, EG 2, EP 100, ERS 10, ERS 22, EP87, EJN17, ERS 4, ERS 7, CJV 20.
	IV B	3	EP 91, IS 18834, IS 18840
V	VA	5	ERS 3, ERS 9, IS 27782, ERS 18, ERS 12
	VB	10	PEC 18, TNS 694, IS 23891, SOR 11216, IS 29917, ERS 1, ERS 16, ERS 17, ERS 19, ERS 20
VI	VIA	19	EJN 3, EJN 31, EJN 32, ERS 6, EJN 15, EJN 35, ERS 13, POP 52, EJN 6, EP 33, ERS 14, EP 84, IS 23702, IS 30417, TNS 667, SEB 12020, IS 27718, EP 45, ERS 15
	VI B	2	SOR 12829, IS 18891
VII	VII A	14	EJN 4, EJN 11, EJN 14, EJN 22, EJN 33, EJN 24, EJN 25, EJN 29, EJN 30, EJN 34, EJN 8, IS 18882, EP 124, IS 27714
	VII B	7	IS 27783, IS 44338, ERS 21, IS 23603, IS 2123, IS 14571, IS 18860
VIII	VIII	2	SOR 11218, SPV 462

20 were grouped under sub-cluster II B. Cluster III had four germplasm, out of which E 228 had maximum plant height, number of leaves and panicle length. Cluster IV had 13 germplasm and the genotype EP 91 had large stem diameter and broad leaf. Fifteen germplasm were grouped under Cluster V and elite germplasm present in Cluster V were ERS 1 and ERS 16 recorded maximum stem diameter, plant height, leaf length and panicle length. Diversity analysis in sorghum has been reported by Rayaprolu *et al.* (2017), Hamidou *et al.* (2018), Derese *et al.* (2018), Subramanian *et al.* (2019), Karadi and Kajjidoni (2019) and Gebrie and Genet (2019).

**Principal Component Analysis:** The degree of differentiation between germplasm is needed to safeguard potentially vital traits. PCA was performed to find out the traits contributing the greatest towards variability. It also reduces the large data set of phenotypic characters into a small meaningful set of characters because the extent of genetic variation was much needed for genetic improvement. In this study, a total of 102 germplasm accessions were subjected to PCA. The yield contributing traits were separated into 13 principal components of which four components had more than one eigen values. Similarly, Derese *et al.* (2018) and Mumtaz *et al.* (2018) reported three principal components with more than one eigen value. Seventy three percentage of total variation was observed from the first four PCs. The first principal component is the major contributor to the total variation in the population followed by subsequent components. Proportion of variance, cumulative proportion and eigen values of 13 PCs are furnished in **Table 2**. Characters *viz.*, hundred seed weight, plant height and grain yield per plant were of frequent occurrence in the first four principal components highlighting their major contribution towards

variability. Variations contributed by each character within the components are depicted in **Table 3**. Biplot and Scree plot of sorghum genotypes for 13 PCs are presented in **Fig. 2**.

Traits *viz.*, days to 50% flowering (0.3156), stem diameter (0.3632), flag leaf length (0.3468), flag leaf breadth (0.3410), flag leaf area (0.3496), number of leaves (0.3028), leaf length (0.2675), leaf breadth (0.3186) and days to maturity (0.3134) contributed for maximum of 34 % variation in the PC1 with 4.45 eigen value.

Divergence of second PC was mainly contributed by plant height (0.3714), days to 50% flowering (0.3545), leaf length (0.2438), panicle length (0.145) and days to maturity (0.3664). Eigen value for PC2 was 2.63 and contributed about 20% of total variation. The second principal component was negatively associated with flag leaf area (-0.3203), hundred seed weight (-0.3907) and grain yield per plant (-0.3235).

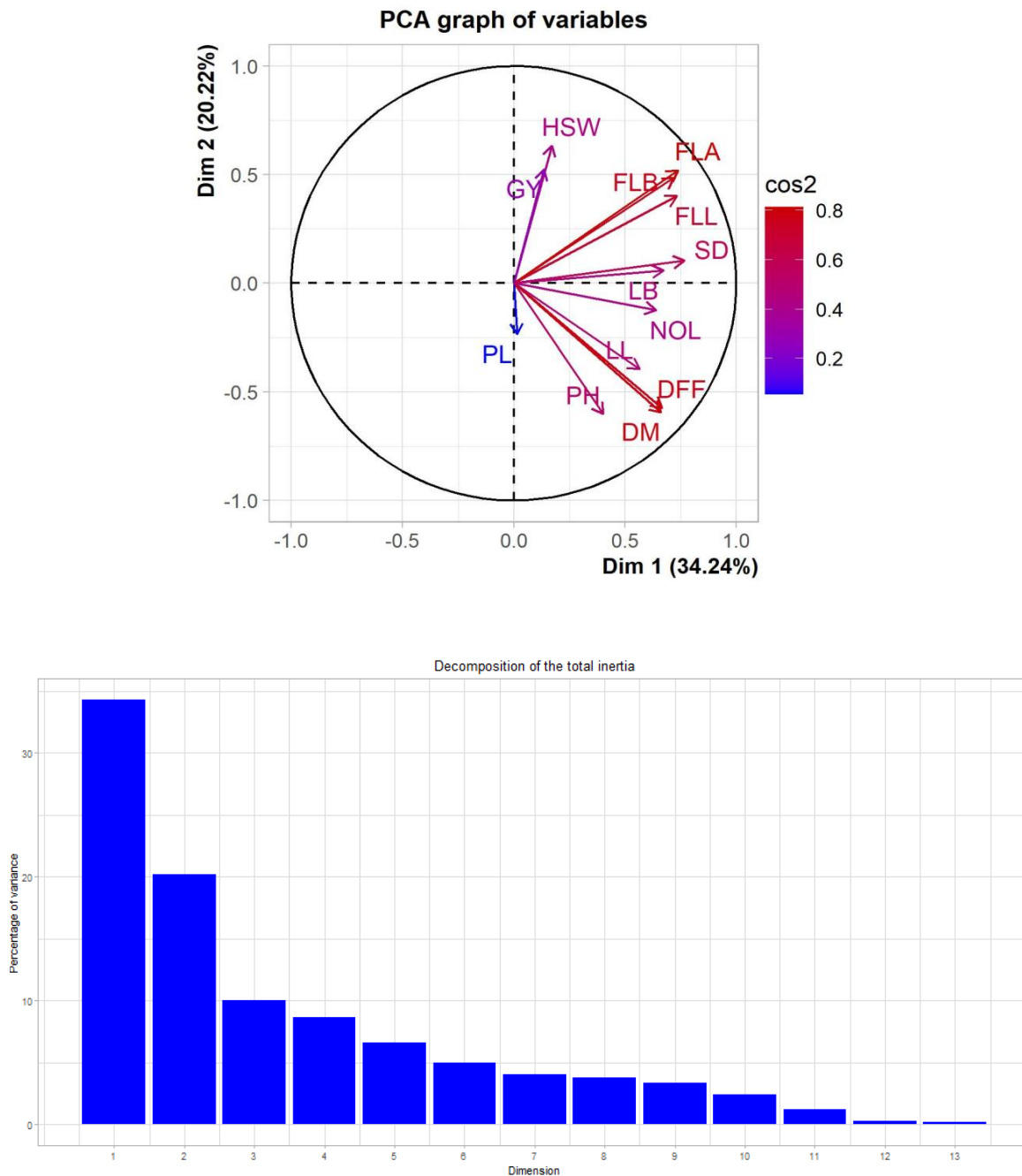
Contribution of divergence by PC3 was observed to be 10% with eigen value of 1.31. Traits responsible for variation in PC3 were leaf length (0.1148), leaf breadth (0.3711) and hundred seed weight (0.2316). The third principal component was negatively associated with panicle length (-0.7221) and grain yield per plant (-0.3067).

Total variation contributed by PC4 was 9% and had 1.13 eigen value. Major variation contributing traits in PC4 were flag leaf length (0.2234), flag leaf breadth (0.3567) and flag leaf area (0.3044). Similar results were reported by Tesfaye (2017) and Derese *et al.* (2018). Distribution of germplasm across first and second components based on PCA is presented in **Fig. 3**.

**Table 2. Proportion of variance, Cumulative proportion and Eigen values for thirteen principal components**

Components	Proportion of variance	Cumulative proportion	Eigen values
PC1	0.34	0.34	4.45
PC2	0.20	0.54	2.63
PC3	0.10	0.65	1.31
PC4	0.09	0.73	1.13
PC5	0.07	0.80	0.85
PC6	0.05	0.85	0.65
PC7	0.04	0.89	0.52
PC8	0.04	0.93	0.49
PC9	0.03	0.96	0.43
PC10	0.02	0.98	0.31
PC11	0.01	1.00	0.16
PC12	0.00	1.00	0.04
PC13	0.00	1.00	0.03

PC – Principal Component



**Fig. 2. Biplot and Scree plot of sorghum genotypes for thirteen principal components**

Results of this study revealed that Cluster II and Cluster VIII had more number of elite germplasm. The germplasm from these clusters had maximum flag leaf length, flag leaf breadth, flag leaf area, number of leaves, hundred seed weight and grain yield per plant suggesting their use as one of the parental line for hybridization breeding programme. The germplasm EJN 20 from the cluster II was early maturing. Similarly, the germplasm SPV 462

of cluster VIII was observed to be higher grain yielding. Crossing of these two genotypes might be useful in developing high yielding short duration cultivars.

Enhancement of grain quality features is only possible when there is genetic diversity for distinct qualities in the crop, which is a basic condition for crop development. As a result, the characteristics that are dispersed over

Table 3. Eigen vectors for thirteen yield contributing traits on sorghum genotypes

Variables	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8	PC9	PC10	PC11	PC12	PC13
PH	<b>0.1899</b>	<b>0.3714</b>	-0.2234	0.0876	-0.3683	-0.0313	0.2007	-0.6807	0.1284	-0.3175	0.0786	0.0152	0.0668
DFF	<b>0.3156</b>	<b>0.3545</b>	0.0363	-0.0616	0.3928	0.259	0.1426	0.0558	-0.1753	-0.0108	0.1099	0.6546	-0.231
SD	<b>0.3632</b>	-0.0637	0.0719	-0.2439	-0.2647	0.0296	0.4009	-0.0258	0.2292	0.7151	-0.0323	-0.0362	-0.0265
FLL	<b>0.3468</b>	-0.2473	-0.281	<b>0.2234</b>	0.1625	0.0903	-0.2244	-0.1598	0.1352	-0.0069	-0.6373	-0.0615	-0.3886
FLB	<b>0.341</b>	-0.2991	-0.0156	<b>0.3567</b>	-0.0157	-0.175	0.0758	0.0459	-0.1591	-0.0636	0.6127	-0.2063	-0.4252
FLA	<b>0.3496</b>	-0.3203	-0.195	<b>0.3044</b>	0.0983	-0.0077	-0.0889	0.0017	-0.031	0.0575	0.0768	0.2501	0.7444
NOL	<b>0.3028</b>	0.0762	0.0697	-0.0633	-0.5533	0.4297	-0.3036	0.4495	0.1848	-0.2681	0.042	0.0272	-0.0101
LL	<b>0.2675</b>	<b>0.2438</b>	<b>0.1148</b>	-0.2884	0.1192	-0.4433	-0.6782	-0.1427	0.1527	0.1717	0.1694	-0.0132	0.0063
LB	<b>0.3186</b>	-0.0357	<b>0.3711</b>	-0.1155	-0.1093	-0.5477	0.2941	0.2018	-0.1999	-0.364	-0.3573	0.0571	0.0604
PL	0.0072	<b>0.145</b>	-0.7221	-0.2357	-0.2023	-0.1966	-0.0194	0.2482	-0.4956	0.1097	-0.037	-0.0177	-0.0313
DM	<b>0.3134</b>	<b>0.3664</b>	0.0382	-0.0282	0.3783	0.2657	0.1107	0.0849	-0.151	-0.0486	-0.0281	-0.675	0.225
HSW	0.0806	-0.3907	<b>0.2316</b>	<b>-0.419</b>	-0.097	0.3191	-0.1612	-0.4138	-0.5486	-0.0402	0.0072	-0.0223	0.0108
GYPP	0.0642	-0.3235	-0.3067	<b>-0.5718</b>	0.2791	-0.0025	0.1935	0.046	0.4345	-0.3611	0.1816	-0.0233	0.0053

PH-Plant height (cm); DFF- Days to fifty percent flowering; SD-Stem diameter (cm); FLL- Flag leaf length (cm); FLB-Flag leaf breadth (cm); FLA- Flag leaf area (cm); NOL- Number of leaves; LL- Leaf length (cm); LB- Leaf breadth (cm); PL- Panicle length (cm); DM- Days to maturity; HSW-Hundred seed weight (gm); GYPP- Grain yield per plant (gm)

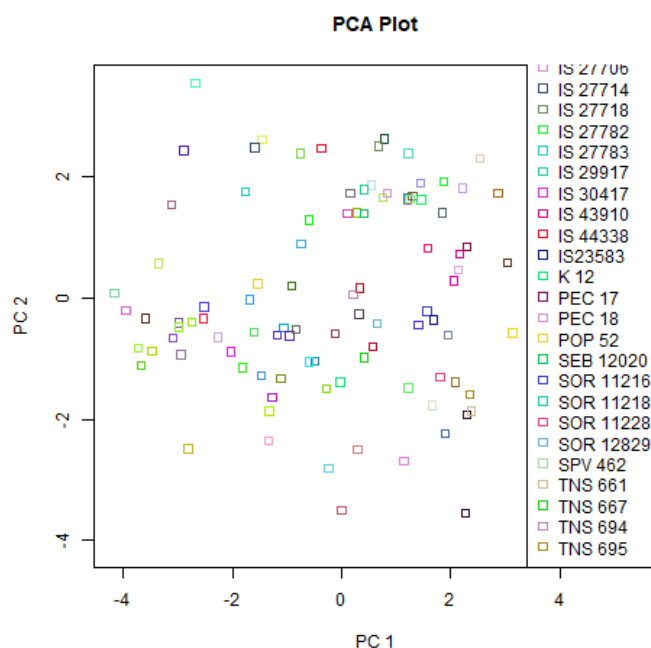


Fig. 3. Distribution of genotypes across first and second components based on PCA

several main components and contribute the most to explaining variability should be considered for use in a breeding effort. It gives information that may aid in the better selection of parental genotypes with certain traits as well as the formulation of breeding strategies for trait enhancement.

## REFERENCES

Agricultural Statistics At a Glance. 2021. "Directorate of

Economics and Statistics, Government of India, Ministry of Agriculture, Department of Agriculture and Cooperation, New Delhi."

Derese, S. A., Shimelis, H., Mwadzingeni, L. and Laing, M. 2018. Agro-morphological characterisation and selection of sorghum landraces. *Acta Agriculturae Scandinavica, Section B—Soil & Plant Science*, **68(7)**: 585-595. [Cross Ref]



- Doijad, S. B., Bagade, A. B. and More, A. W. 2016. Evaluation of sorghum germplasm for genetic diversity using D2 statistics. *Electronic Journal of Plant Breeding*, **7(4)**: 934- 938. [Cross Ref]
- Elangovan, M. and Babu, P. K. 2015. Genetic variability and diversity of sorghum landraces collected from uttar pradesh, India. *Indian Journal of plant genetic resources*, **28(2)**: 213-221. [Cross Ref]
- Gebregergs, G. and Mekbib, F. 2020. Estimation of genetic variability, heritability, and genetic advance in advanced lines for grain yield and yield components of sorghum [*Sorghum bicolor* (L.) Moench] at Humera, Western Tigray, Ethiopia. *Cogent Food & Agriculture*, **6(1)**: 1764181. [Cross Ref]
- Gebrie, G. and Genet, T. 2019. Morphological characterization and evaluation of sorghum [*Sorghum bicolor* (L.) Moench] landraces in Benishangul Gumuz, North-western Ethiopia. *Greener Journal of Agricultural Sciences*, **9(1)**: 37-56. [Cross Ref]
- Hamidou, M., Souley, A. K. M., Kapran, I., Souleymane, O., Danquah, E. Y., Ofori, K., Gracen, V. and Ba, M. N. 2018. Genetic variability and its implications on early generation sorghum lines selection for yield, yield contributing traits, and resistance to sorghum midge. *International Journal of Agronomy*, **32(6)**: 757-765. [Cross Ref]
- Karadi, A. and Kajjidoni, S. 2019. Genetic variability and diversity for productivity traits and grain quality including nutritional quality traits in selected mini core and promising released varieties of sorghum. *Journal of Pharmacognosy and Phytochemistry* **8(4)**: 2091-2097.
- Leonard, K. and R.J. Peter 2009. Finding Groups in data: An Introduction to cluster Analysis, 344.
- 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. [Cross Ref]
- Mofokeng, M. A., Shimelis, H., Laing, M. and Shargie, N. 2019. Genetic variability, heritability and genetic gain for quantitative traits in South African sorghum genotypes. *Australian Journal of Crop Science*, **13(1)**: 1-10. [Cross Ref]
- Mumtaz, A., Hussain, D., Saeed, M., Arshad, M. and Yousaf, M. I. 2018. Estimation of genetic diversity in sorghum genotypes of Pakistan. *Journal of the National Science Foundation of Sri Lanka*, **46(3)**: 271-280. [Cross Ref]
- Olsen, K. M. and Wendel, J. F. 2013. A bountiful harvest: genomic insights into crop domestication phenotypes. *Annual review of plant biology*, **64**: 47-70. [Cross Ref]
- Rayaprolu, L., Ashok Kumar, A., Manohar Rao, D. and Deshpande, S. P. 2017. Genetic variability for agronomic traits in sorghum minicore collection. *International Journal of Agriculture Innovations and Research*, **6(3)**: 533-537.
- Subramanian, A., Raj, R. N. and Elangovan, M. 2019. Genetic variability and multivariate analysis in sorghum (*Sorghum bicolor*) under sodic soil conditions. *Electronic Journal of Plant Breeding*, **10(4)**: 1405-1414. [Cross Ref]
- Tesfaye, K. 2017. Genetic diversity study of sorghum (*Sorghum bicolor* (L.) Moench) genotypes, Ethiopia. *Acta Universitatis Sapientiae, Agriculture and Environment*, **9(1)**: 44-54. [Cross Ref]
- Upadhyaya, H. D., Vetriventhan, M., Asiri, A. M., CRAzevedo, V., Sharma, H. C., Sharma, R., Sharma, R. and Wang, Y. H. 2019. Multi-trait diverse germplasm sources from mini core collection for sorghum improvement. *Agriculture*, **9(6)**: 121. [Cross Ref]
- U.S.D.A. 2021. "World sorghum production". *USDA Foreign Agricultural Service, Washington*.