Received: 08 Apr 2021 | Revised: 04 Aug 2021 | Accepted: 08 Aug 2021

Electronic Journal of Plant Breeding

Research Note

Principal component analysis of sorghum landraces for yield contributing traits

K. Geetha* and S. Divya

Regional Research Station, Paiyur, Krishnagiri district, Tamil Nadu - 635 112, India
***E-Mail:** geethakreddy@yahoo.com

Abstract
Principal component analysis was carried out with experimental material comprising of 30 sorghum land races and observations were recorded on eleven biometrical traits *viz.*, days to 50 per cent flowering, days to maturity, plant height, the number of tillers per plant, inflorescence length, inflorescence width stem thickness, hundred seed weight, grain yield per plant, fodder yield per plant (wet) and fodder yield per plant (dry). The results of the principal component analysis indicated that four out of eleven principal components with eigenvalues >1 were obtained. Principal components 1, 2, 3 and 4 contributed 31.82, 18.41, 13.16 and 10.3 per cent towards the variation observed among accessions, respectively. Regarding the comparison of the first four principal components, variation proportional to the first component was positive for days to 50 per cent flowering, days to maturity and inflorescence length. The second principal component was negatively associated with fodder yield per plant (wet) and fodder yield per plant (dry). The third principal component was a positive influence on stem thickness whereas, other traits *viz.,* the number of tillers per plant, hundred seed weight, grain yield per plant had contributed negatively. The fourth principal component was associated with plant height, inflorescence width negatively.

Key words: Sorghum, landraces, PCA, biometrical traits

Sorghum [(*Sorghum bicolor* (L.) Moench)] is a tropical cereal crop belonging to the Poaceae family. It is broadly distributed throughout Ethiopia and it is the most important cereal in the lowland areas because of its drought tolerance. The total area under sorghum in the world is 39.95 million hectares and the production is about 58.06 million metric tonnes with a yield of 1.45 metric tonnes/ha. In India, it is grown in an area of 4.82 million hectares and with annual grain production of about 4.77 million metric tonnes with a yield of 0.99 metric tonnes/ha (Anonymous,2021). Understanding the genetic variation in sorghum accessions is very much needed to obtain promising genetic improvement in sorghum crops. A great level of knowledge on genetic variability among the sorghum landraces will facilitate precision breeding and profiling the genetic diversity of sorghum creates a wider opportunity to select the highly diverged parents with varied traits (Doijad *et al.,* 2016). Hybridization carried out with highly diverged parents creates an opportunity for gene constellation yield in transgressive segregants.

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 technique for reducing the dimensions of datasets and increasing the interpretability but at the same time with the minimized loss of information. The main aim of Principal component analysis is to resolve the total variation of a set of traits into linear, independent composite traits which consecutively maximize variability in the data (Johnson, 2012).

https://doi.org/10.37992/2021.1203.142

In the present investigation, 30 sorghum landraces were used for genetic diversity study of quantitative traits and were raised in Regional Research Station, Paiyur during *Kharif,* 2018 (**Table 1**.). Biometrical observations were recorded on traits *viz*., days to 50 per cent flowering, days to maturity, plant height, the number of tillers per plant, inflorescence length, inflorescence width, stem thickness, hundred seed weight, grain yield per plant, fodder yield per plant (wet) and fodder yield per plant (dry). Principal Component Analysis was conducted to identify the patterns of variation and to estimate the relative contribution of various characters for total variability. Statistical analysis was performed by using Statistical Tool for Agricultural Research (STAR) package.

Table 1. List of selected Sorghum landraces used for PCA analysis

S.No.	Sorghum land races
1.	PYR-(RS)-16-01-Palacode local
2.	PYR-(RS)-16-02-Papparapatti local
3.	PYR-(RS)-16-04-Karimangalam local
4.	PYR-(RS)-16-05-Kochamalai local
5.	PYR-(RS)-16-06-Kochamalai local
6.	PYR-(RS)-16-07-Puliyambatti local
7.	PYR-(RS)-16-08-Irumathur local
8.	PYR-(RS)-16-09-Kochamalai local
9.	PYR-2 (RS)
10.	TNAUR00402457-IS 30536
11.	TNAUR00402482-IS 8012
12.	TNAUR00402488-IS 9113
13.	TNAUR00402493-IS 1233
14.	TNAUR00402504-IS 12804
15.	TNAUR00402508-IS 12965
16.	TNAUR00402519-IS 14861
17.	TNAUR00402520-IS 14862
18.	TNAUR00402657-IS 30536
19.	TNAUR00404263-PGR Red cholam
20.	TNAUR00404266-PGR Red cholam 63
21.	TNAUR00404267-PGR Red cholam 92
22.	TNAUR00404269-PGR Red cholam 82
23.	TNAUR00404416-TKSV 1008
24.	TNAUR00404418-TKSV 1023
25.	TNAUR00404720-PGR Red cholam 108
26.	PYR-(TV)-16-01-Palacode local
27.	PYR-(TV)-16-02-Papparapatti local
28.	PYR-(TV)-16-03-Kambainallur local
29.	PYR-(TV)-16-04-Karimangalam local
30.	PYR-1 (TV)

The PCA scores for 30 selected sorghum landraces of sorghum in the first eleven principal components were calculated and were measured as PC1, PC2, PC3, PC4, PC5, PC6, PC7, PC8, PC9, PC10 and PC11. Data were considered in each component with eigenvalues more than 1 as per the suggestions given by Brejda *et al.* (2000), which determines as a minimum 10 per cent of the variation. With regard to PCA, four out of eleven principal components with eigenvalues > 1 were extracted. Among the thirty sorghum land races, these four components contributed 73.70 per cent of the total variation. The first principal component takes out the maximum amount of variation among yield characters. Principal components 1, 2, 3 and 4 contributed 31.82, 18.41, 13.16 and 10.3 per cent towards the variation observed among accessions, respectively.

This suggests that up to the principal component four more variations were observed. The first principal component is the major contributor to the total variation in the population followed by subsequent components. The criteria were used by Clifford and Stephenson (1975).

The first component was positive for days to 50 per cent flowering, days to maturity and inflorescence length. The second principal component was associated with fodder yield per plant (dry) and fodder yield per plant (wet) negatively. The third principal component was positively associated with stem thickness whereas, other traits *viz.,* the number of tillers per plant, hundred seed weight, grain yield per plant had contributed negatively. The fourth principal component was associated with plant height, inflorescence width negatively (**Table 2**.).

The percentage of variation associated with each principal component was obtained by drawing a graph between eigenvalues and principal component numbers through a scree plot. PC1 showed 31.82 per cent variability with the eigen value of 3.50. The eigen values are gradually declined from PC1 to PC11 (**Table 3**.). The Eigenvalues for left out principal components were 2.02, 1.44, 1.13, 0.92, 0.83, 0.60, 0.36, 0.13, 0.03, and 0.01, respectively (**Fig.1**).

The distribution of sorghum landraces reported by different variables from PC 1, PC2 and PC 3 were distributed in different groups, which clearly showed genetic diversity among sorghum landraces (**Fig. 2 to 4**).

Karadi *et al.* (2019) carried out the PCA analysis using 24 selected mini core and 16 promising varieties in *Rabi* sorghum. The first three principal components exerted variations of 30.55, 19.74, and 16.07 per cent, respectively and the traits *viz.,* plant height and true density traits contributed positively in mini core accessions. Similarly, the principal component analysis was carried out in sorghum accessions by Kavithamani *et al.* (2019). A minimum threshold of eigenvalue of one, the first four principal components accounted for a cumulative of about 73.70 per cent of the whole phenotypic diversity observed among the germplasm accessions. The results were in line with the findings of Abraha *et al.* (2015).

Table 2. Factor loading by different biometrical traits

Principal Components	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8	PC9	PC10	PC11
Y1	0.4418	0.2918	-0.2769	0.0676	-0.0148	0.0179	-0.1194	-0.0071	0.3485	-0.0016	-0.7086
Y2	0.444	0.2635	-0.2816	0.0346	0.0548	-0.0411	-0.2097	0.1245	0.3226	0.1004	0.689
Y3	0.1545	0.2243	0.2851	-0.3846	0.4456	0.5561	0.2816	0.3298	-0.0213	0.0014	-0.0159
Y4	0.121	-0.19	-0.5261	-0.0124	-0.4179	0.3074	0.6223	0.0241	-0.1042	0.0268	0.0618
Y5	0.4106	-0.1315	-0.0534	-0.3705	-0.2196	-0.1036	-0.3746	0.2593	-0.6296	-0.0865	-0.0597
Y6	-0.1271	-0.4623	0.0166	-0.6041	-0.1925	-0.0439	-0.1296	0.1354	0.5738	0.0276	-0.0281
Y7	0.1126	-0.0496	0.4881	0.4346	-0.5289	0.2729	-0.0971	0.4132	0.1522	0.0103	0.0054
Y8	-0.3523	-0.0633	-0.3878	0.208	0.2529	-0.1648	-0.0487	0.7575	-0.0227	0.0393	-0.0945
Y9	-0.2711	-0.0763	-0.2981	0.072	0.0193	0.6897	-0.5531	-0.1997	-0.0572	0.0385	0.0069
Y10	0.2885	-0.5048	-0.0105	0.2562	0.3255	0.0644	0.0147	-0.0343	0.0878	-0.6896	0.038
Y11	0.3016	-0.5107	0.0718	0.2023	0.2972	0.0125	0.0132	-0.0635	-0.0525	0.7088	-0.0659

Y1- Days to fifty per cent flowering, Y2- Days to maturity, Y3- Plant height, Y4- No of tillers per plant, Y5- Inflorescence length, Y6- Inflorescence width, Y7- Stem thickness, Y8- Hundred seed weight, Y9- Grain yield per plant, Y10- Fodder yield per plant (Wet), Y11-Fodder yield per plant (dry)

Table 3. Principal component analysis for different biometrical traits

Principal Components	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8	PC9	PC10	PC11
Standard deviation	1.8709	1.4231	1.2033	1.0649	0.9618	0.9141	0.7748	0.6010	0.3649	0.1768	0.0778
Proportion of variance	0.3182	0.1841	0.1316	0.1031	0.0841	0.076	0.0546	0.0328	0.0121	0.0028	0.0006
Cumulative proportion	0.3182	0.5023	0.6339	0.737	0.8211	0.8971	0.9517	0.9845	0.9966	0.9994	1
Eigen Values	3.5002	2.0251	1.4479	1.1339	0.9251	0.8356	0.6004	0.3612	0.1332	0.0313	0.006

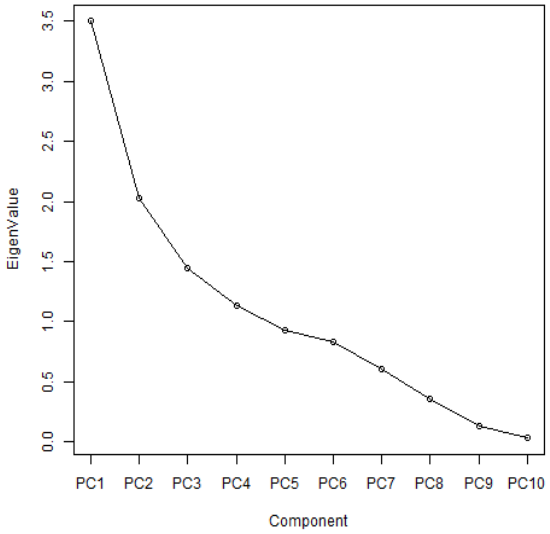

Fig.1.Scree plot for Eigen values and principal components

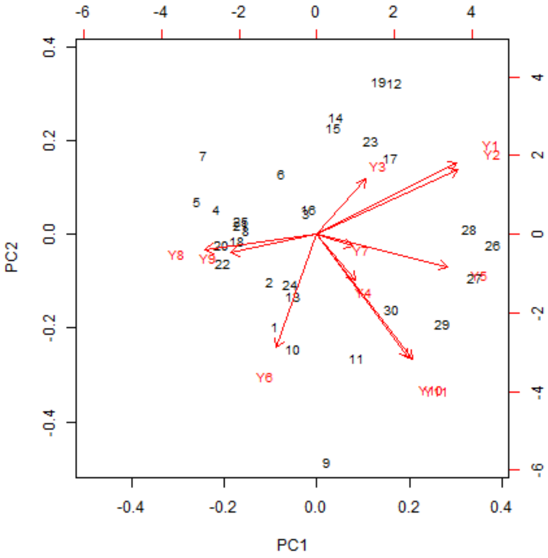

Fig. 2. Distribution of genotypes across first two components based on PCA

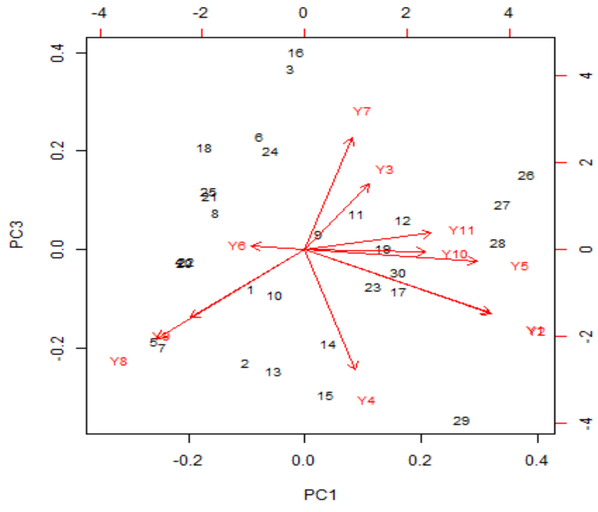

Fig. 3. Distribution of genotypes across first and third components based on PCA

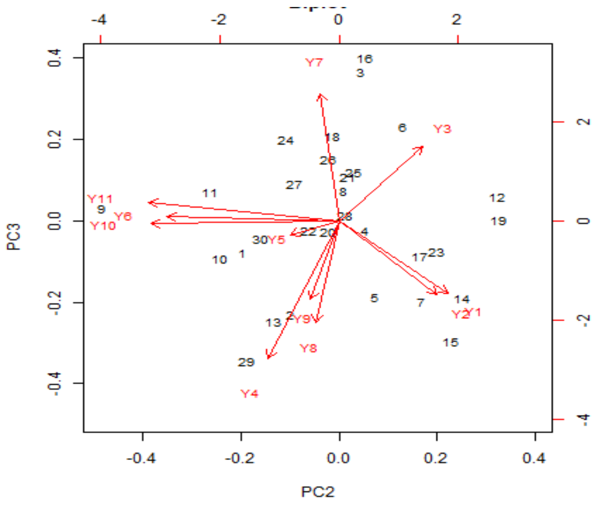

Fig. 4. Distribution of genotypes across second and third components based on PCA

Enhancement towards grain quality characteristics is feasible only when the genetic variability exists for different traits in the crop, which is the basic requirement for crop improvement. The high level of genetic variation prevailing in the population will be analysed with the support of principal component analysis and it also provides information on which characters are contributing to genetic divergence among the genotypes in the population. So, the traits that spread over different principal components and contributing most to explain the variability should be taken into consideration for their utilization in a breeding programme. It provides information that could facilitate in better selection of parental genotypes with specific traits and in formulating breeding strategies for trait improvement.

REFERENCES

Abraha, T., Githiri, S. M., Kasili, R., Araia, W. and Nyende, A. B. 2015. Genetic variation among sorghum (*Sorghum bicolor* L. Moench) landraces from Eritrea under post-flowering drought stress conditions. *Am. J. Plant Sci.,* **6**(09): 1410. [Cross Ref]

Anonymous. 2021. USDA. Foreign Agricultural Service, Circular series, March 2021.

Brejda, J., Moorman, T., Karlen, D. and Dao, T. 2000. Identification of regional soil quality factors and indicators. I. Central and Southern High- *Plains. Soil Sci. Soc. Am. J.,* **64**: 2115-2124. [Cross Ref]

Clifford, H. T. and Stephenson, W.1975. An introduction to numerical classification,New York: Academic press. [Cross Ref]

Doijad, S. B., Bagade, A. B. and More, A. W. 2016. Evaluation of sorghum germplasm for genetic diversity using D2 statistics. *Electron. J. Plant Breed.*, **7(4):** 934-938. [Cross Ref]

Johnson, D.E.2012. Applied multivariate methods for data analysis. New York: Duxbury Press.

Karadi, A. and Kajjidoni, S. T. 2019. Genetic diversity and principal component analysis for grain quality and productivity traits in sorghum [*Sorghum bicolor* (L.) Moench]. *Int. J. Curr. Microbiol. App. Sci*, **8(9):** 839-850. [Cross Ref]

Kavithamani, D., Yuvaraja, A. and Selvi, B. 2019. Principal component analysis and grouping of sorghum (*Sorghum bicolor* L. Moench) gene pool for genetic diversity. *Electron. J. Plant Breed.*, **10(4):** 1426-1434. [Cross Ref]