International Journal of Agriculture and Plant Science
www.agriculturejournal.in
Online ISSN: 2664-7664, Print ISSN: 2664-7656
Received: 19-02-2025, Accepted: 20-03-2025, Published: 04-04-2025
Volume 7, Issue 2, 2025, Page No. 49-55

Multivariate analysis of diversity of Sorghum germplasm

A Gopikrishnan[1], P Thilagam[2*], P Veeramani[3], T Balaji[4], K P Sivalumar[5]

[1] Assistant Professor, Department of Plant Breeding and Genetics, Agricultural Research Station, Virinjipuram, Vellore, Tamil Nadu Agricultural University, Tamil Nadu, India

[2] Associate Professor, Department of Agricultural Entomology, Agricultural Research Station, Virinjipuram, Vellore, TamilNadu Agricultural University, Tamil Nadu, India

[3] Assistant Professor, Department of Agronomy, Tapioca and Castor Research Station, Yethapur, Selam, Tamil Nadu Agricultural University, Tamil Nadu, India

[4] Assistant Professorr, Department of Soil Science and Agricultural Chemistry, Agricultural College and Research Institute Vazhavachanur, Thiruvannamalai District, Tamil Nadu Agricultural University, Tamil Nadu, India

[5] Assistant Professor, Department of Food Science and Nutrition, Community Science College and Research Institute, Madurai, Tamil Nadu Agricultural University, Tamil Nadu, India

Abstract

Descriptive, Principal Component Analysis and Cluster analysis were performed on twelve morphological and phenological characters to identify the pattern of genetic variability in 50 sorghum genotypes. Descriptive analysis revealed that the largest variation was observed for panicle girth with coefficient of variation of 38.41 % and the least variation was observed in days to maturity with coefficient of variation of 5.97%. Principal component analysis extracted four components contributing to around 78.22% of total variability among twelve characters. Principal component 1, 2, 3 and 4 contributed 35.68 %, 19.99 %, 12.53% and 10.02% of variation respectively in the total variability. Thus the results of principal component analysis used in the study had revealed the high level of genetic variation and the traits controlling for the variation were identified. Hence, these entries can be utilized for trait improvement in breeding programs for the traits contributing for major variation. The biplot analysis indicated node length, days to maturity, panicle girth, days to 50% flowering, single plant yield, plant height, leaf length panicle length exhibited positive associations due to acute angle between them. Cluster analysis acknowledged two major clusters, one cluster with 11 genotypes and other cluster with 39 genotypes.

Keywords: Sorghum bicolor, Diversity, Principal Component Analysis (PCA), Cluster analysis

Introduction

Sorghum [*Sorghum bicolor* (L.) Moench] ranks as the fifth most significant grain crop globally and it is recognized as a C4 plant, known for its high photosynthetic efficiency, enabling it to thrive in various environmental conditions while producing both grain and fodder. Sorghum crop plays a crucial role in feeding 500 million food-insecure individuals in the arid and semi-arid regions of Africa and Asia (Kumar, 2016). In 2024, global sorghum production is projected to reach 58.36 million tons, cultivated over 40.04 million hectares, resulting in a productivity rate of 1460 kg per hectare (https://apps.fas.usda.gov/psdonline/circulars/production).

In India, the production of sorghum is estimated at 4.4 million tons from 4.0 million hectares, yielding a productivity of 1100 kg per hectare, with 43% of the output coming from the Kharif season and 57% from the Rabi season. The worldwide cultivation, yield and efficiency of sorghum remain significantly lower than those of other cereals like maize and rice, falling short of global market requirements (Haywood & Recha, 2022) [3]. This underperformance in sorghum production can be attributed to factors such as drought, water and heat stress, suboptimal soil conditions, the use of inferior sorghum varieties, and inadequate agricultural practices. Therefore, it is essential to develop a climate-resilient sorghum genotype to enhance production levels. Genetic diversity plays a crucial role in enhancing a species' capacity to adapt and thrive. A higher level of genetic diversity provides increased opportunities for resilience against future climate changes. Maintaining genetic diversity is essential for ensuring that diverse populations have the best prospects for long-term survival. Gaining insights into the genetic diversity of sorghum can significantly aid in crop enhancement aimed at developing climate-resilient genotypes with higher yields. Consequently, it is important to assess the existing genotypes for their genetic diversity and identify the most effective ones based on their performance. Principal component analysis is a multivariate method that assists researchers in identifying meaningful relationships among traits and elucidating the variation observed among different genotypes. Cluster analysis is a statistical technique used for categorizing items into groups, or clusters, according to their degree of similarity. Principal component analysis and cluster analysis methods were used here to identify the diverse genotypes for parental selection to use in crop improvement programme.

Materials and Methods

A sum of fifty sorghum germplasm accessions (Table 1) were assessed at Agricultural Research Station, Virinjipuram (Latitude: 12° 93′N and Longitude: 79° 01′ E) located in the Vellore district of Tamil Nadu during *Rabi* season of 2021-22. The genotypes were cultivated with a spacing of 60 cm between rows and 15 cm between individual plants. Standard agronomic practices were implemented to ensure the crops remained healthy. Two phenological traits, namely days to 50% flowering and days to maturity, along with ten morphological traits including plant height, number of leaves, number of nodes, leaf length, leaf width, node length, node girth, panicle length, panicle girth, and single plant yield were recorded from five plants of each entries. The gathered data underwent statistical analysis. The Statistical Tool for Agricultural Research (STAR) software was employed for basic statistics and Principal Component Analysis, while cluster analysis was conducted using NTSYs software.

Result and Discussion

Descriptive Analysis

The first-order statistical events, including maximum, minimum, mean, standard deviation (SD), and coefficient of variation (CV), skewness and kurtosis for the measured traits are presented in Table 2. The days to 50% flowering among the germplasm accessions ranged from a minimum of 42 days to a maximum of 60 days with an average of 48 days. Maturity duration varied between 75 days and 100 days, yielding a mean of 84 days. Plant height measurements ranged from a minimum of 61 cm to a maximum of 180 cm, resulting in an average height of 131 cm. The number of leaves per plant varied from 7.70 to 9.20 with a mean of 6.66. The number of nodes per plant had a minimum of 4.7 and a maximum of 9.20 with an average of 6.58. The mean leaf length was recorded at 49.20 cm with minimum and maximum values of 34.00 cm and 70.0 cm respectively. Leaf width ranged from 3.90 cm to 8.20 cm with a mean of 5.77 cm. Nod length had a mean of 21.91 cm with minimum and maximum values of 13.40 cm and 31.00 cm, respectively. Nod girth measurements varied from 2.10 cm to 6.10 cm with an average of 3.76 cm. Panicle length ranged from 7.90 cm to 27.00 cm with a mean of 14.68 cm while panicle girth varied from 5.20 cm to 23.6 cm, averaging 11.20 cm. The single plant yield exhibited a wide range from 9.80g to 30.30 g with a mean of 15.00 g. Among the traits, panicle girth exhibited the greatest variation with a CV of 38.41 followed by panicle length (31.92) and single plant yield (28.90). In contrast, days to maturity showed the least variation with a CV of 5.97. The most significant deviation from the population mean (27.71) was noted in plant height followed by leaf length (7.95) days to maturity (5.01), days to 50% flowering (4.76), panicle length (4.69), single plant yield (4.33) and panicle girth (4.30).

Research on skewness elucidates the degree of deviation from a normal distribution and sheds light on the gene actions that affect a specific trait. Kurtosis provides information about the distribution's flatness or peakedness and helps estimate the number of genes involved in the trait's expression (Ajay *et al*. 2016, Shamini and Selvi 2023) [8] When the frequency distribution is asymmetrical, it signifies a skewed distribution, suggesting that the trait is governed by non-additive gene action and may be influenced by environmental factors. A skewed distribution of certain traits implies the presence of dominant gene action, where positive skewness points to complementary gene interactions, while negative skewness indicates duplicate gene interactions (Trikoesoemaningtyas *et al.,* 2024) [11].

The analysis of fifty sorghum accessions across twelve traits demonstrated various distribution patterns, as illustrated in Fig 1. None of the traits exhibited a normal distribution. The traits including days to 50% flowering, days to maturity, number of leaves, number of nodes, leaf length, leaf width, node girth, panicle length, panicle girth, and single plant yield displayed positive skewness, indicating complementary gene interactions. Conversely, negative skewness was observed in plant height and node length, suggesting the presence of duplicate gene interactions.

The classification of leptokurtic and platykurtic distributions is determined by positive and negative kurtosis values respectively. Positive kurtosis indicates the presence of gene interactions while negative kurtosis suggests their absence. Traits that display a leptokurtic distribution are generally controlled by a small number of genes, whereas those with a platykurtic distribution are influenced by a greater number of genes. Specifically, the traits such as days to maturity, plant height, number of leaves, number of nodes, node girth, panicle length, panicle girth and single plant yield were identified as leptokurtic, while days to 50% flowering, leaf length, leaf width and node length were categorized as platykurtic.

Principal Component Analysis

Principal component analysis (PCA) serves as a fundamental technique in multivariate data analysis and has been extensively applied in breeding programs. Its primary purpose is to reduce the dimensionality of extensive datasets, enhance interpretability, and minimize information loss, all while retaining all samples and variables (Sonal Chavan *et al*., 2024) [10]. In the study PCA has extracted four components having eigen value above 1.0 and exhibited total variance of 78.22% among twelve characters. (Table 3, Fig.2). The largest contribution to the total variation in the population is from principal component 1 (35.68 %) followed by component 2 (19.99 %), component 3 (12.53 %) and component 4 (10.02 %). All the collected sorghum entries were widely scattered across different quarters (Fig 3). Principal component 1 had the contribution from the traits such as plant height (0.3673), leaf length (0.3571), panicle girth (0.3529), days to 50% flowering (0.3237), single plant yield (0.3253) and panicle length (0.3099) which accounted to 35.68 % of the total variability (Ali *et al*. 2011) [2]. The principal component 2 explained 19.99 % of total variability from number of nods (0.4630), number of leaves (0.4561), nod girth (0.4113) and leaf width (0.3203). The Principal Component 3 accounted 12.53% variation of total variability from days to maturity (0.5143), days to 50% flowering (0.4966), Single plant yield (-0.3609) and panicle girth (-0.3148). Principal component 4 contributed 10.02% variation from node length (0.5002), node girth (-0.4542), number of leaves (0.3747), number of nodes (0.3615) and panicle length (-0.3068). Thus the results of principal component analysis used in the study had revealed the high level of genetic variation and the traits controlling for the variation were identified. Hence, these entries can be utilized for trait improvement in breeding programs for the traits contributing for major variation.

Biplot analysis

A genotype by trait biplot is utilized to examine the interrelationships between various traits and their associated genotypes (Jain and Patel. 2016) [4] The cosine of the angle formed between traits reflects their correlation: a positive correlation exists when the angle is less than 90°, while an angle exceeding 90° indicates a negative correlation. An angle of precisely 90° signifies the absence of correlation between the traits (Weikai & Judith 2018) [12]. The biplot analysis revealed that node length, days to maturity, panicle girth, days to 50% flowering; single plant yield, plant height, leaf length, and panicle length demonstrated positive associations, as evidenced by the acute angles between them. Similarly, a positive association was also noted among node girth, number of nodes, number of leaves, and leaf width. Conversely, zero correlation was observed between node girth and plant height, as well as between the number of nodes and panicle girth. The angle created between a genotype and a trait indicates the genotype's relative performance regarding that trait. An acute angle, which measures less than 90°, indicates that the genotype performs above average for the trait, while an obtuse angle, exceeding 90°, suggests that the genotype performs below average. A right angle, precisely 90°, signifies that the genotype is average in relation to the trait. The vectors extending from the origin to each trait illustrate the extent of variation associated with that trait (Shushay, et al., 2023) [9]. A longer vector indicates greater variation, while a shorter vector reflects lesser variation. In this investigation, the number of nodes, number of leaves, days to 50% flowering, and plant height were analyzed in detail, revealing a notable level of variability (Fig 3).

Cluster Analysis

Cluster analysis useful to identify the similarity and dissimilarities among the tested genotypes (Neladr et al, 2024, Premkumar et al., 2022) [7]. In this study cluster analysis acknowledged two major clusters, one cluster with 11 genotypes and other cluster with 39 genotypes (Fig. 4). High panicle girth genotypes EJ6 and EJ34 are grouped in sub cluster of first major cluster. High single plant yield genotypes ER3, GGUB24 and EJ73 are found sub cluster of second Major cluster. The highest plant height genotype NSJB6683 occupied first major cluster and shortest plant genotype IS4744 reside in second major cluster likewise late maturity genotype E146 placed in first major cluster and early maturity genotypes IS4168 positioned in second major cluster. Selecting genotypes for parents from distant clusters will facilitates to develop lines of high yield with climate resilient crop.

Table 1: Sorghum entries involved in genetic diversity analysis

S. No.	Entries	S. No.	Entries	S. No.	Entries	S. No.	Entries	S. No.	Entries
1	ER3	11	EJ54	21	IS29650	31	IS30409	41	IS4012
2	EJN4	12	EJ56	22	IS30322	32	IS32298	42	IS4078
3	EJN7	13	EJ62	23	IS30339	33	IS40116	43	IS4079
4	EJN15	14	EJ63	24	IS30369	34	EJ6	44	IS4094
5	EJN24	15	EJ69	25	IS30403	35	EJ7	45	IS4103
6	E146	16	EJ70	26	IS917	36	E-9	46	IS4168
7	E158	17	EJ72	27	IS2833	37	EJ33	47	154735
8	E172	18	EJ73	28	IS24694	38	EJ34	48	1S4741
9	E174	19	GGUB24	29	IS29640	39	NSJB6683	49	1S4744
10	EJ49	20	GGUB28	30	ER3	40	CSV17	50	1S4763

Table 2: Descriptive analysis for 12 characters of 50 sorghum genotypes

Characters	Minimum	Maximum	Mean	Std Dev	CV	Skewness	Kurtosis
Days to 50% flowering	42.00	60.00	48.38	4.76	9.84	0.36	-0.80
Days to maturity	75.00	100.00	83.88	5.01	5.97	0.69	0.84
Plant Height	61.00	180.00	130.99	27.71	21.16	-0.92	0.26
Number of leaves	7.70	9.20	6.66	0.97	14.50	0.51	0.31
Number of Nodes	4.70	9.20	6.58	1.00	15.28	0.58	0.46
Leaf length	34.00	70.00	49.20	7.95	16.17	0.46	-0.33
Leaf width	3.90	8.20	5.77	0.91	15.72	0.49	-0.00
Node length	13.40	31.00	21.91	3.88	17.70	-0.23	-0.23
Node Girth	2.10	6.10	3.76	0.88	23.49	0.56	0.23
Panicle length	7.90	27.00	14.68	4.69	31.92	0.93	0.06
Panicle Girth	5.20	23.60	11.20	4.30	38.41	0.94	0.60
Single Plant Yield	9.80	30.30	15.00	4.33	28.90	1.32	2.20

Table 3: Four principal components and their eigen value, percent of total variation and component matrix

Principal Components	PC1	PC2	PC3	PC4
Eigen Values	4.2814	2.3985	1.5037	1.2028
% of Variance	0.3568	0.1999	0.1253	0.1002
Cumulative %	0.3568	0.5567	0.6820	0.7822
	Components matrix			
Days to 50% flowering	0.3237	-0.1994	0.4966	-0.1389
Days to maturity	0.2875	-0.2707	0.5143	-0.1041
Plant Height	0.3673	-0.1527	0.0999	0.2830
Number of leaves	0.2454	0.4561	0.1359	0.3747
Number of Nodes	0.2298	0.4630	0.1785	0.3615
Leaf length	0.3571	-0.0485	-0.0195	-0.2550
Leaf width	0.2650	0.3203	-0.2267	-0.0444
Node length	0.1448	-0.2834	-0.2729	0.5002
Node Girth	0.1398	0.4113	-0.0067	-0.4542

Panicle length	0.3099	0.0957	-0.2699	-0.3068
Panicle Girth	0.3529	-0.2131	-0.3148	-0.0291
Single Plant Yield	0.3253	-0.1793	-0.3609	-0.0146

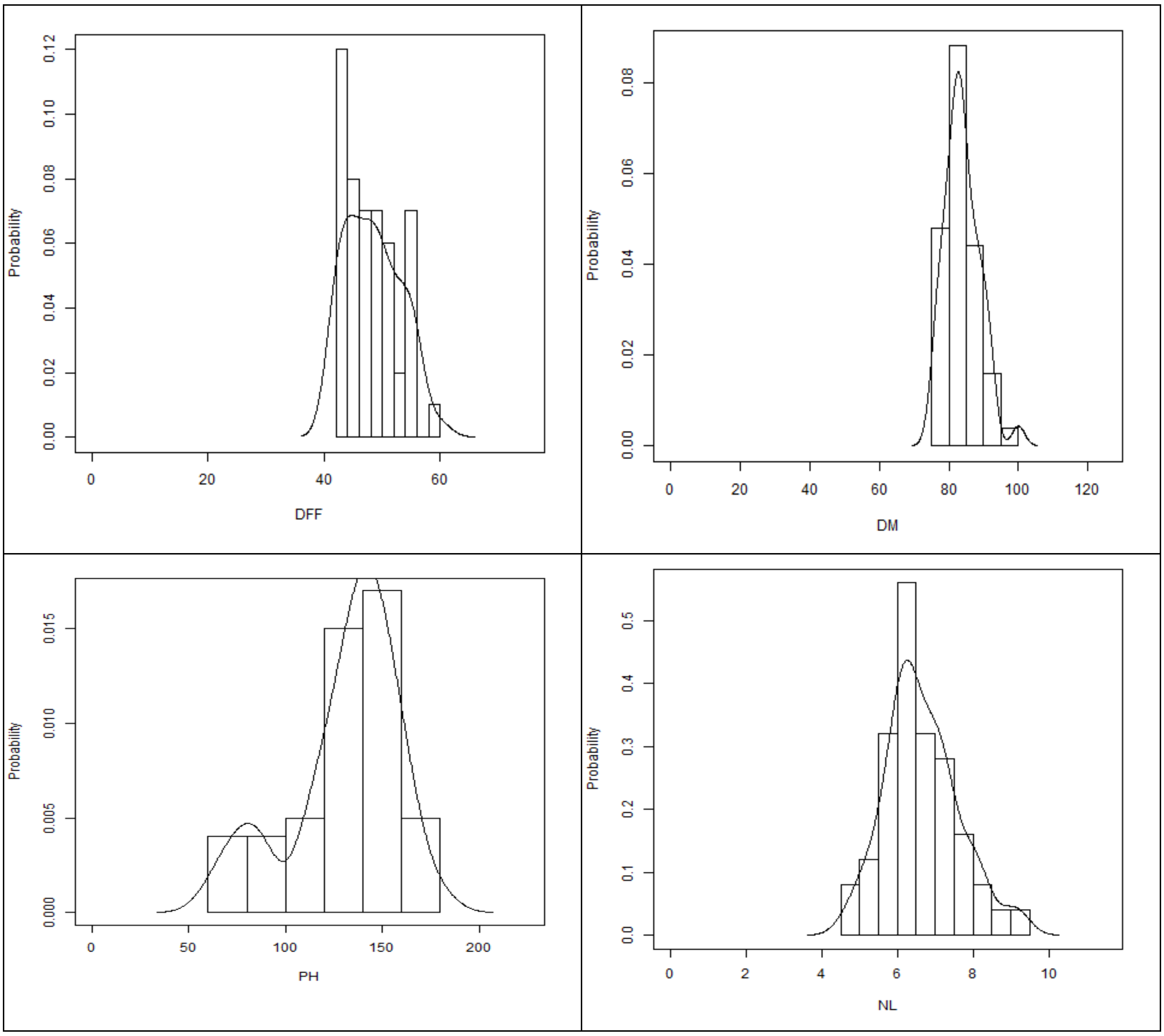
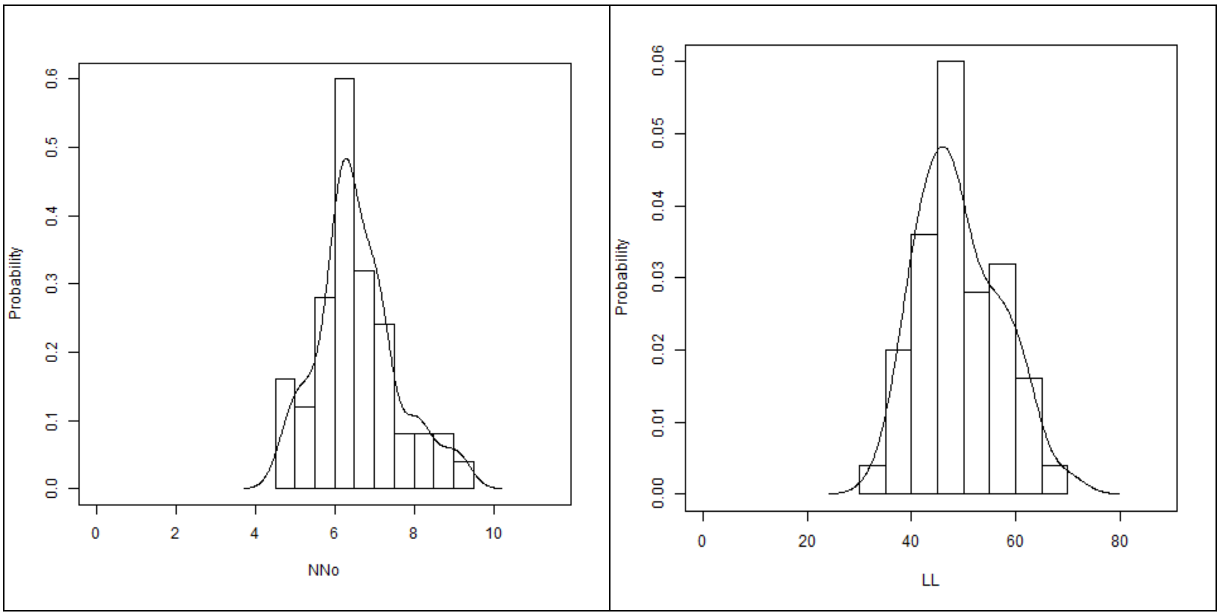
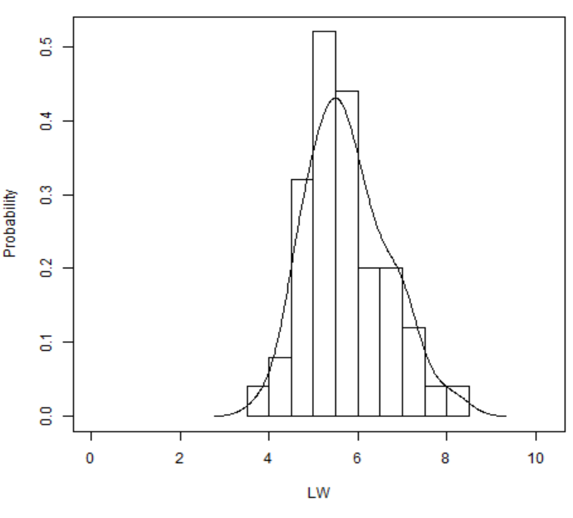
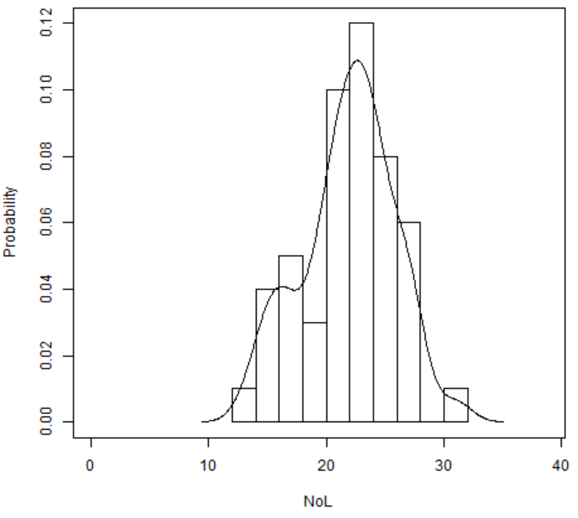
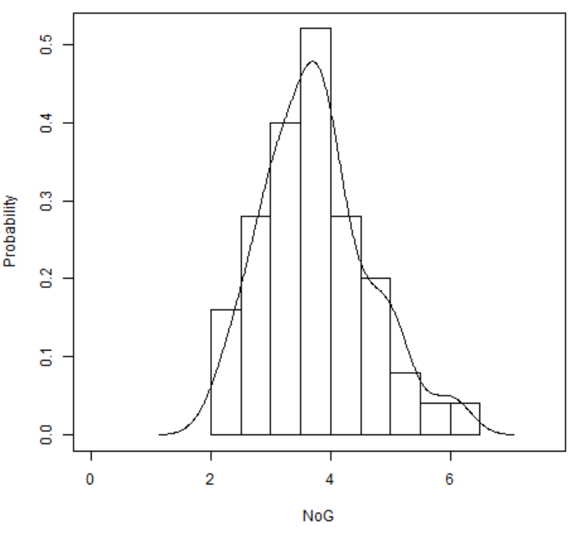


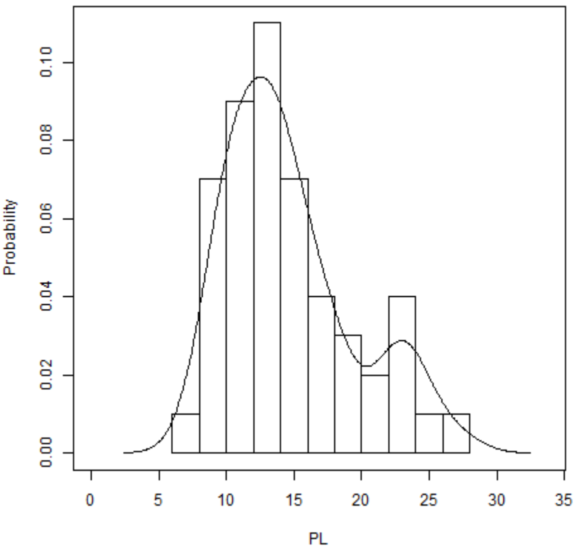
Fig 1: Frequency distribution of twelve characters studied in 50 sorghum genotypes

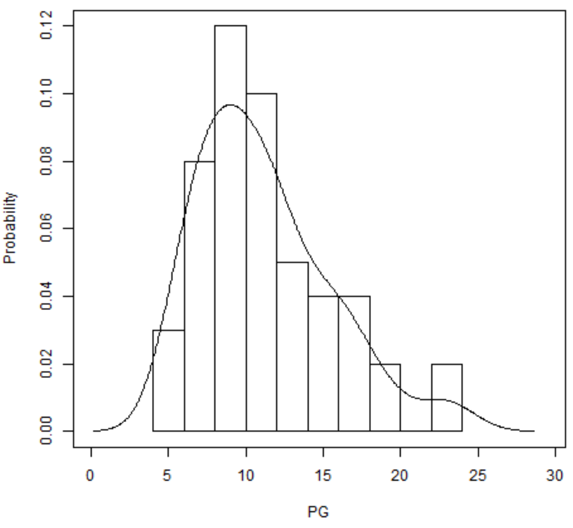


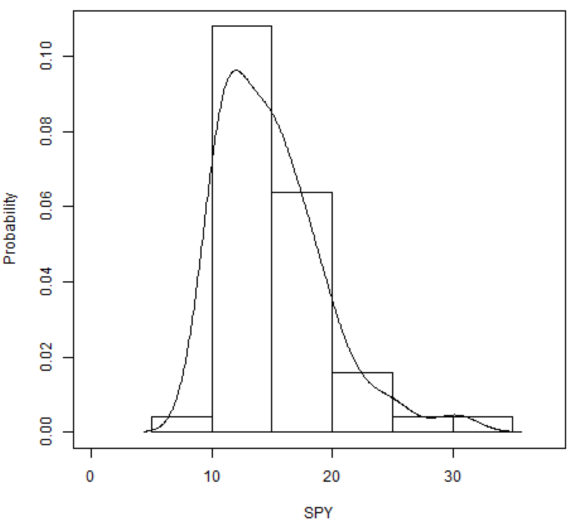












DFF-Days to 50% flowering	LW-Leaf width
DM-Days to maturity	NoL-Node length
PH-Plant height	NG-Node girth
NL-Number of leaves	PL-Panicle length
NNo-Number of nodes	PG-Panicle girth
LL-Leaf length	SPY-Single Plant Yield

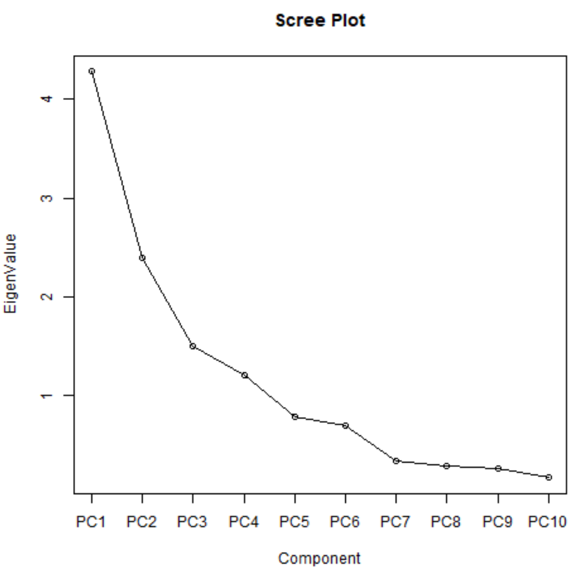


Fig 2: Scree plot diagram of eign values contructed on 12 traits recorded in 50 sorghum germplasm

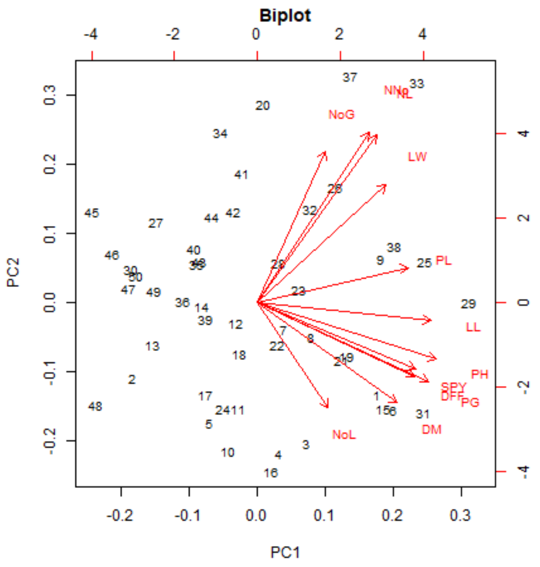


Fig 3: Biplot of trait relationship of 50 sorghum germplasm

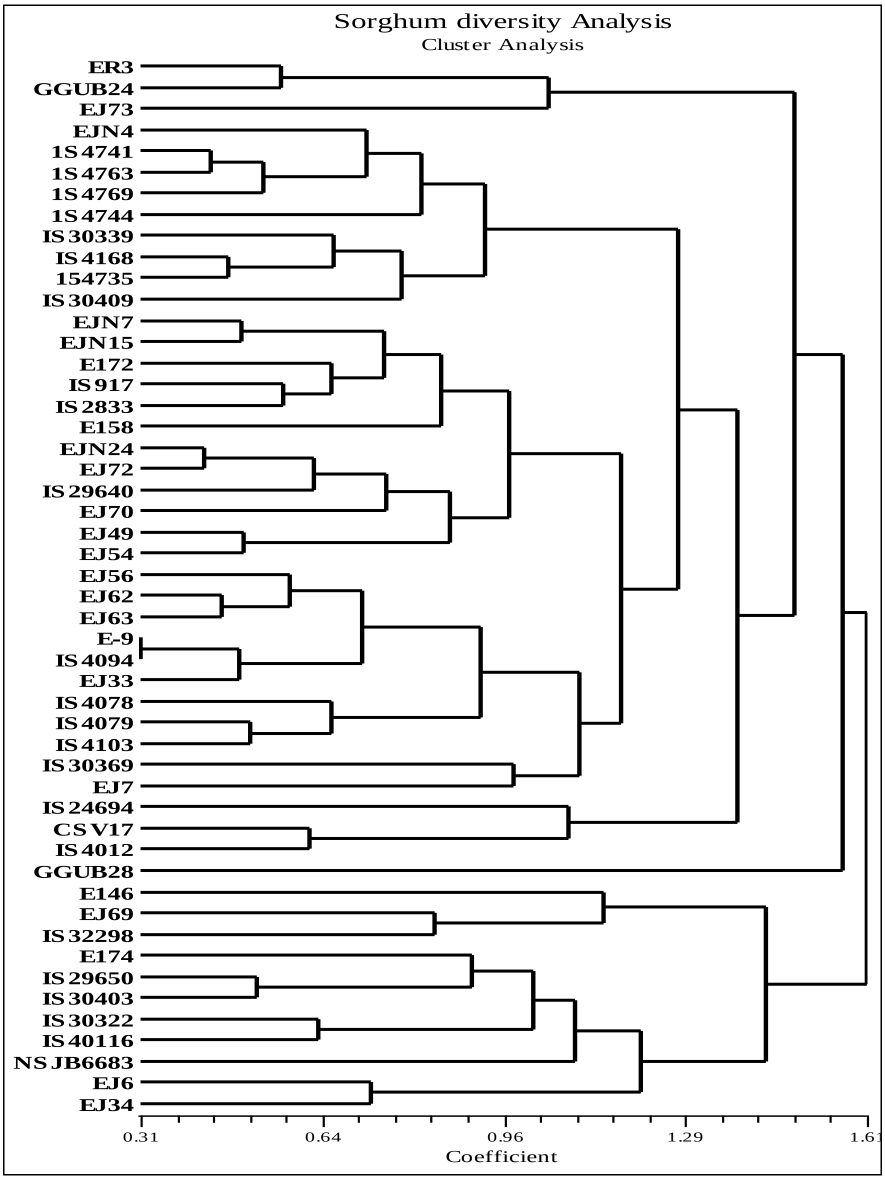


Fig 4: Cluster analysis depicting genetic relationships among 50 sorghum genotypes

Conclusion

Principal component 1 encompassed contributions from various traits, including plant height, leaf length, panicle girth, days to 50% flowering, single plant yield, and panicle length, which collectively accounted for 35.68% of the total variability observed. These findings will be utilized to identify suitable parent plants for the enhancement of the morphological traits examined in this research. The cluster analysis indicated that hybridization among Sorghum genotypes from distant clusters could result in increased levels of heterosis. Specifically, crossing genotypes with high single plant yield, such as ER3, GGUB24, and EJ73, with those exhibiting high panicle girths, like EJ6 and EJ34, is expected to yield significant heterosis for hybrid development and provide better segregants in the F2 generation for varietal development.

Reference

1. Ajay BC, Byregowda M, Veerakumar GN, Ganapathy KN, Murlidhar Meena, Prashanth Babu H.
2. Ali MA, Jabran, K, Awan, SI Abbas, A. Ehsanullah, ZM, Acet, T. Farooq, J and Rehman, A. Morphophysiological diversity and its implications for improving drought tolerance in grain sorghum at different growth stages. Australian J. Crop Sci.,2011:5(3):311-320.
3. Haywood Ondiko J, Recha WC. Global Sorghum Production Constraints: A Review. Annals of Arid Zone,2022:61(1):1-12.
4. Jain SK, Patel PR. Principal component and cluster analysis in sorghum (sorghum bicolor (l.) moench). Forage Res.,2016:42(2):90-95
5. Kumar AA. “Botany, taxonomy and breeding,” in The Sorghum Genome: Compendium of Plant Genomes, eds. S. Rakshit and Y. H. Wang (Cham: Springer), 2016, 27–45.
6. Neladri Sekhar Sarkar, Kalaimagal T, Kavithamani D, Chandirakala R, Manonmani S, Raveendran M. Assessing Genetic Relationships, Trait Associations and Diversity Patterns in Sorghum Germplasm Through Correlation, Cluster and Principal Component Analysis. Agricultural Science Digest,2024:44(3):445-452
7. Premkumar V, Patil SG, Djanaguiraman M and Sridevy S. “Assessment of Genetic Diversity in Sorghum (*Sorghum Bicolor (L.).* Moench) Genotypes Using Principal Components and Cluster Analysis”. International Journal of Plant & Soil Science,2022:34(21):577-585.
8. Shamini Selvi. Assessment of Frequency Distribution in F3 generation of Sorghum (*Sorghum bicolor* L. Moench.) For Grain Yield And Its Attributed Traits. *Madras Agric. J.*, 2023:109(4-6):24-28.
9. Shushay Welderufael, Fetein Abay, Amsalu Ayana, and Tilahun Amede. Genotype by Trait (GT) and Genotype by YieldTraits (GYT) Analysis of Sorghum Landraces in Tigray, Northern Ethiopia. Crop Breed Genet Genom, volume 5 issue 2 2023, pages1-18.
10. Sonal Chavan, T Shashikala, Ephrem Habyarimana, S Vanisri, T Ramesh and N Sandhya Kishore. Principal component analysis for agronomic traits in multi-cut forage sorghum. International Journal of Research in Agronomy,2024:7(11):38-42.
11. Trikoesoemaningtyas, AR Fadilah, PW Burnama, Rahayu F, Rachman F, *et al*. Genetic Variations In Sorghum Segregating Populations Based On Yield And Amylose Content. SABRAO Journal of Breeding and Genetics,2024:56(4):1357-1366.
12. Weikai Yan, Judith Frégeau-Reid. Genotype by Yield Trait (GYT) Biplot: a Novel Approach for Genotype Selection based on Multiple Traits. Scientific Reports,2018:8:1-18.