Forage Res., **51**(2) : pp. 166-175 (2025) **http://forageresearch.in**

GENETIC VARIABILITY, TRAIT ASSOCIATION, AND MULTIVARIATE ANALYSIS IN FORAGE SORGHUM [*SORGHUM BICOLOR* (L.) MOENCH] GENOTYPES

PUMMY KUMARI[1&5]*, NEERAJ KHAROR[1&5], DIBYENDU SETH[2], S. K. PAHUJA[3], SATPAL[1&5], G. S. DAHIYA[2], VENKATESH BHAT[4] AND R. MADUSUDHANA[4]

[1]Forage Section, Department of Genetics and Plant Breeding, CCS Haryana Agricultural University,
Hisar-125 004 (Haryana), India
[2]Department of Genetics and Plant Breeding, CCS Haryana Agricultural University, Hisar-125 004 (Haryana), India
[3]College of Agriculture, CCS Haryana Agricultural University, Hisar-125 004 (Haryana), India
[4]AICRP on Sorghum and Small Millets, ICAR-Indian Institute of Millets Research [Global Center of excellence on Millets (Shree Anna)], Hyderabad-500030, India
[5]ICAR-AICRP on Sorghum and Small Millets, Forage Section, Department of Genetics and Plant Breeding, CCS Haryana Agricultural University, Hisar-125 004 (Haryana), India
*****(*e-mail : pummy.hau@gmail.com)*

(Received: 16 May 2025; Accepted: 27 September 2025)

SUMMARY

The present study was conducted to evaluate genetic variability, trait associations, and diversity among twenty-six single cut forage sorghum genotypes. Each genotype was grown in uniform plots, and observations were recorded on thirteen quantitative traits: plant height (PH), number of leaves per plant (NL), leaf length (LL), leaf breadth (LB), leaf-stem ratio (LSR), stem girth (SG), plant population (PP), total soluble solids (TSS), crude protein content (CP%), *in vitro* dry matter digestibility (IVDMD), hydrocyanic acid content (HCN), green fodder yield (GFY), and dry fodder yield (DFY). Significant genotypic variability was observed for most traits, notably for GFY, DFY, and PH, indicating potential for selection. Path analysis revealed that GFY had the highest positive and direct effect on DFY, while PH also contributed significantly both directly and indirectly. Cluster analysis grouped the genotypes into eleven clusters, highlighting a broad genetic base and enabling the identification of divergent parents. Principal Component Analysis (PCA) revealed that five components with eigenvalues greater than one explained 76.81% of the total variability. Traits like PH, GFY, and SG were major contributors to the observed diversity. The integration of univariate and multivariate analyses facilitated the identification of superior genotypes like SPV2809, SPV2800, CSH 40F and SPH1984 with desirable traits for biomass yield and forage quality, offering valuable insights for future breeding programs aimed at enhancing sorghum for fodder use.

Key words: Sorghum, Genetic variability, Path analysis, Cluster analysis, PCA, Biomass quality

Sorghum (*Sorghum bicolor* L. Moench) stands as the world's fifth most vital cereal crop after wheat, rice, maize and barley, playing a pivotal role in global food, feed and fodder security (Kumar *et al.*, 2024; Gebre *et al.*, 2025). Renowned for its exceptional drought tolerance and adaptability to high temperatures, sorghum is increasingly recognized for its potential in carbon sequestration and sustainable agriculture, especially in climate-vulnerable regions (Engida, 2023). In India, which supports nearly 20% of the global livestock population, sorghum is integral to livestock nutrition, particularly in arid and semi-arid zones where fodder scarcity is pronounced (Altaf *et al.*, 2024).

Forage sorghum genotypes demonstrate superior *in vitro* dry matter digestibility and greater biomass production than their grain counterparts, rendering them ideal for fulfilling the growing demand for quality fodder in low-input agricultural systems (Chavan *et al.*, 2024). As global livestock product consumption increases, enhancing sorghum's fodder productivity and nutritive value is critical to supporting dairy and livestock-dependent rural economies. Sorghum's broad genetic diversity provides breeders with extensive opportunities to select for favourable alleles associated with yield and nutritional traits (Tirfessa *et al.*, 2025). Genetic variability assessments, including heritability estimates and genetic advance, are essential tools for identifying traits with high selection potential (Engida, 2023). Advanced multivariate methods such as Principal Component Analysis (PCA) and cluster analysis have proven effective in characterizing genotypes and prioritizing traits, thereby enhancing the efficiency of parental line selection in hybrid breeding programs (Altaf et al., 2024). Moreover, correlation and path coefficient analyses provide critical insights into the interrelationships among agronomic traits, guiding targeted breeding strategies for the development of superior forage varieties (Chavan *et al.*, 2024). The incorporation of genes from diverse germplasm has already demonstrated success in producing transgressive segregants and high-performance hybrids with improved fodder characteristics (Rohila *et al.*, 2016 and Rohila *et al.*, 2020). This study investigates the genetic diversity, variability, and trait associations among forage sorghum genotypes, with particular emphasis on morphological and biochemical characteristics. The research aims to identify superior genotypes and key agronomic traits that can reform future breeding initiatives, ultimately contributing to enhanced fodder production, improved livestock nutrition, and greater livelihood security for agricultural communities.

MATERIALS AND METHODS

Twenty-six single cut forage sorghum genotypes including hybrids and varietal genotypes were received under ICAR-AICRP Trials from IIMR, Hyderabad on sorghum which were raised at Forage Research Area, Forage section, Department of Genetics and Plant Breeding, CCS Haryana Agricultural University, Hisar, in a Completely Randomized block design (CRBD) with three replications for one season (*Kharif*, 2020). The experimental material consisted of 26 genotypes, which included the following: SPH1984 (G1), SPH1985 (G2), SPH1986 (G3), SPH1987 (G4), SPH1988 (G5), SPH1989 (G6), SPV2796 (G7), SPV2797 (G8), SPV2798 (G9), SPV2799 (G10), SPV2800 (G11), SPV2801 (G12), SPV2802 (G13), SPV2803 (G14), SPV2804 (G15), SPV2805 (G16), SPV2806 (G17), SPV2807 (G18), SPV2808 (G19), SPV2809 (G20), CSH36F (G21), CSV 30F (G22), CSV 35F (G23), CSH 40F (G24), CSV 21F (G25), and a Local Check (G26). The soil for the experimental plot was sandy loam type with uniform topography and free from waterlogged conditions having a pH of 7.8. Throughout the cropping season, all of the advised agronomic packages of operations, including irrigation, fertilizer applications, and crop protection management, were done. Five plants were randomly selected for each replication, and observations were made on thirteen quantitative traits. The analysis is based on the determination of the thirteen quantitative traits. The traits measured were plant height (PH) in cm, number of leaves per plant (NL), leaf length (LL) in cm, leaf breadth (LB) in cm, Leas-Stem Ratio (LSR), Stem girth (SG) in mm, Plant population (PP) per m/row length, total Soluble Solid (TSS), crude protein (CP %), *in-vitro* dry matter digestibility (IVDMD), HCN, green fodder yield per plot (GFY) in kg and dry fodder yield per plot (DFY) in kg and then yield is converted into q/ha. Various statistical techniques were employed in the study, including analysis of variance (as per Panse and Sukhatme, 1954), phenotypic and genotypic coefficients of variation, heritability estimates (Burton and DeVane, 1953), genetic advance (Johnson et al., 1955), correlation analysis (Singh and Choudhary, 1977), path coefficient analysis (Dewey and Lu, 1959), D^2 statistics (Mahalanobis, 1936), and principal component analysis (PCA). All analyses were conducted in R Studio using R version 4.1.2 (R Core Team, 2021). The 'variability' package (Popat *et al.*, 2020) facilitated the computation of ANOVA, genetic variability metrics and correlation analysis. Visualizations were created using the 'ggplot2' package, version 3.3.4 (Wickham, 2016). For D^2 and cluster analysis, the 'biotools' package (da Silva, 2021) and 'FactoMineR' (Le *et al.*, 2008) were utilized. PCA and biplot analyses were performed using the 'FactoMineR' and 'factoextra' packages to interpret trait contributions and genotype clustering patterns.

RESULTS AND DISCUSSION

1.1. Descriptive Statistics

The morphological heterogeneity among the forage sorghum genotypes was revealed by the descriptive statistics for thirteen quantitative features which opens the door for improvement through hybridization and selection. The First order statistical measures viz., maximum, minimum, mean, variance, standard error of mean (SEm), coefficient of variation (CV), broad sense heritability and genetic advance for the measured traits are shown in Table 1.

The mean value for GFY 50.27 kg/plot with minimum and maximum value of 35.6 kg/plot and 65.5 kg/plot were recorded by the genotypes SPV2797 and

TABLE 1

Descriptive statistics for various morphological and biochemical traits

Traits	DFY	GFY	PH	NLP	LL	LB	LSR	PP	SG	TSS	TP	IVDMD	HCN
Maximum (Replication)	18.60	71.20	330.56	28.33	98.67	9.83	0.34	15.07	25.27	8.67	1.96	8.94	190.86
Minimum (Replication)	7.30	30.15	103.56	12.67	51.00	4.20	0.24	2.00	11.24	4.00	0.73	2.53	11.01
Grand Mean	12.42	50.27	224.92	21.14	75.68	6.52	0.29	6.93	17.52	7.03	1.26	5.39	106.67
Standard Error of Mean (SEm)	1.55	6.28	24.75	2.09	4.23	0.67	0.01	1.32	1.43	0.40	0.13	0.56	17.41
Environmental Variance	7.20	118.38	1837.05	13.08	53.77	1.36	0.00	5.22	6.11	0.47	0.05	0.94	909.60
Genotypic Variance	1.66	33.15	863.41	0.95	15.82	0.03	0.00	0.71	0.89	0.08	0.03	0.71	1094.27
Phenotypic Variance	8.86	151.53	2700.46	14.03	69.58	1.39	0.00	5.92	6.99	0.55	0.08	1.65	2003.87
Environmental Coefficient of Variance	21.61	21.65	19.06	17.11	9.69	17.90	6.44	32.94	14.11	9.77	18.32	17.96	28.27
Genotypic Coefficient of Variance	10.39	11.45	13.06	4.61	5.25	2.56	3.50	12.12	5.37	4.01	13.70	15.62	31.01
Phenotypic Coefficient of Variance	23.97	24.49	23.10	17.72	11.02	18.08	7.01	35.10	15.10	10.56	22.87	23.80	41.97
Heritability (Broad Sense)	0.19	0.22	0.32	0.07	0.23	0.02	0.25	0.12	0.13	0.14	0.36	0.43	0.55
Genetic Advance	1.15	5.55	34.23	0.52	3.91	0.05	0.01	0.60	0.69	0.22	0.21	1.14	50.36
Genetic Advance as percentage of mean	9.27	11.04	15.22	2.47	5.16	0.75	3.61	8.62	3.94	3.14	16.90	21.13	47.21

SPV2809, respectively. The genotype SPV 2809 had the maximum DFY whereas SPV 2807 had the minimum DFY. The overall mean for DFY was 12.42 kg/plot. Shortest height of 131.8 cm was observed in SPV2807 and tall growth was shown by SPV2809 (286.8 cm) and average total plant height was 224.92 cm. SPV2796 has the maximum number of leaves per plant whereas; SPV2797 has the minimum (15.3). SPV2808 had the minimum leaf length of 64.3 cm whereas maximum leaf length of 91.6 cm was recorded in SPV2805. The average leaf length was 75.68cm whereas the average leaf width was 6.52 cm. SPH1985 recorded minimum leaf breadth (5.3 cm) whereas maximum leaf breadth of 8.5 cm was recorded in SPV2796. The stem diameter ranged variation from 1.38 cm to 2.16 cm. The lowest stem thickness was measured in HJ 541 (LC) and highest stem diameter was recorded in SPH1986. The mean value of 1.752 cm was observed for stem diameter. The genotype SPV 2807 had the lowest plant population and SPV2802 had the maximum plant population of 10.1. Among the traits studied, the average of TSS was 7.03 with maximum TSS content in SPV2809 and minimum TSS content in SPV2803. The total crude protein percent trait recorded the overall mean of 1.26 g. The genotype SPV2807 was recorded the lowest of 0.8 g and genotype SPV2809 had the highest of 1.68 g (in % value). Minimum IVDMD% was recorded for SPV2806 (3.7) whereas maximum was recorded for CSH36F (8). Lastly, the overall average HCN content was 106.67 µg/g with genotype SPV2800 and SPH1988 recording the minimum and maximum of 45.6 µg/g and 187.7 µg/g, respectively.

The largest variation was observed for HCN with genotypic CV of 31.01% followed by IVDMD i.e., 15.62%, total crude protein (13.70%), plant height (13.06%), Green fodder yield (11.45%) and Dry fodder yield (10.39%), were all rather high to moderate, indicating vulnerability to environmental fluctuation impacting their expression to some extent. The minimum level of variation was observed in leaf breadth followed by leaf stem ratio and EV (2.56%, 3.50% and 3.87% respectively). Highest heritability was observed for HCN followed by IVDMD and total crude protein. Leaf breadth followed by EV showed least broad sense heritability.

1.1. ANOVA

The analysis of variance (ANOVA) conducted for thirteen agro-morphological and quality traits in forage sorghum genotypes revealed considerable variability among treatments, highlighting the genetic diversity present in the evaluated material. Significant differences ($p < 0.01$) were observed for green fodder yield, indicating substantial genetic variation among the genotypes (Table 2). This suggests potential for effective selection and genetic improvement. Plant height also varied significantly at the 1% level, with high mean sum of squares observed for treatments compared to error, suggesting strong genetic control. Stem girth, a trait linked with plant robustness and feed quality, also exhibited significant treatment effects. Genotypes G3 (SPH1986) and G23 (CSV 35F) showed higher stem girth which contribute to better standability and biomass quality.

On the other hand, some traits such as number of leaves per plant, plant population at harvest, total soluble solids (TSS), protein content, *in vitro* dry matter digestibility (IVDMD), and hydrocyanic acid (HCN) content did not show statistically significant differences among genotypes. This could be attributed either to a narrow genetic base for these traits in the studied material or to greater environmental influence, as reflected by relatively higher error mean squares in some cases. Particularly, traits like protein, IVDMD, and HCN exhibited higher replication mean squares, indicating environmental variability and the need for more controlled evaluation or increased replication for precise estimation.

Overall, traits that displayed significant differences-especially green fodder yield, plant height, and leaf morphological characteristics-demonstrated the presence of exploitable genetic variability. These traits can serve as selection indices in breeding programs aimed at developing high-yielding and nutritionally superior forage sorghum genotypes. The identification of high-performing genotypes also lays a strong foundation for further multi-location evaluations and potential varietal development.

TABLE 2

Analysis of variance for various traits

Trait/Character	Replication MSS	Treatment MSS	Error MSS
Green Fodder Yield	112.86	246.58**	103.99
Dry Fodder Yield	7.28	13.82*	6.39
Plant Height	748.71	4608.72*	1746.30
Number of Leaves/Plant	56.32	17.09	12.50
Leaf Length	298.36	99.87*	54.44
Leaf Breadth	4.04	1.97*	1.10
Leaf Stem Ratio	0.0007	0.0007*	0.0003
Plant Population at Harvest	69.37	5.71	6.03
Stem Girth	102.60	10.36*	5.31
TSS	2.47	0.69	0.48
Protein	899.64	0.31	0.33
IVDMD	16991.46	24.11	28.78
HCN	95804.78	2135.94	1438.02

1.1. Correlation

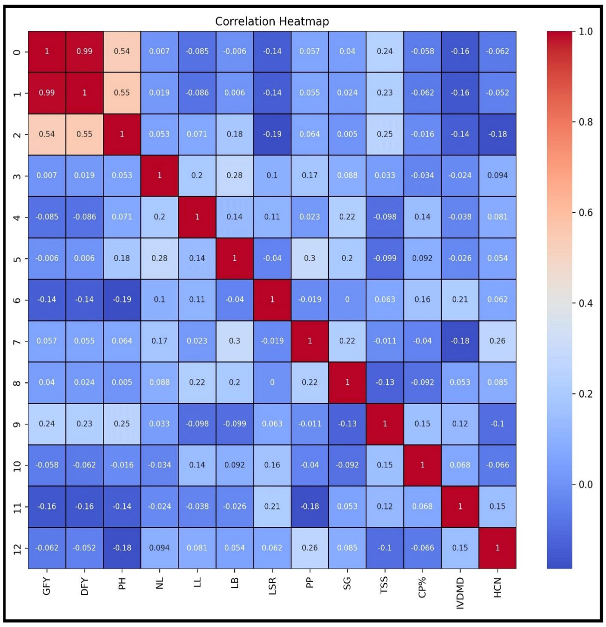
The phenotypic correlation analysis revealed significant interrelationships among key agronomic and quality traits of forage sorghum genotypes, providing insights for trait-based selection in breeding programs (Fig. 1). Green fodder yield (GFY) showed a highly significant and positive correlation with dry fodder yield (DFY) (0.994**), indicating that improvements in one are likely to enhance the other. GFY also had a significant positive correlation with plant height (PH) (0.5358**) and total soluble solids (TSS) (0.2362*), suggesting that taller genotypes with higher sugar content tend to produce more green biomass similar results were reported by Kumari *et al*., 2024. However, GFY displayed weak and non-significant positive associations with traits like plant population (PP), stem girth (SG), and number of leaves (NL), and a negative, non-significant relationship with *in vitro* dry matter digestibility (IVDMD), leaf stem ratio (LSR), leaf length (LL), hydrocyanic acid (HCN), crude protein (CP%), and leaf breadth (LB).

Similarly, DFY exhibited significant positive correlations with PH (0.5482**) and TSS (0.2249*), mirroring trends observed in GFY and confirming the shared physiological basis of yield-related traits. Non-significant yet positive correlations were noted with PP, SG, NL, and LB, while negative associations were observed with IVDMD, LSR, LL, CP%, and HCN. Plant height (PH) also correlated positively and significantly with TSS (0.2502*) and DFY, and had weak positive associations with LB, LL, PP, NL, and SG. However, its relationships with LSR, HCN, IVDMD, and CP% were negative and non-significant, suggesting limited influence of these traits on plant stature. Among foliar traits, number of leaves (NL) showed a significant positive correlation with LB (0.278*), indicating a relationship between leaf count and width. NL also had weak positive correlations with LL, PP, LSR, HCN, SG, PH, TSS, and DFY, though none were statistically significant. Negative associations were recorded with CP% and IVDMD.

Leaf length (LL) had a general trend of positive, non-significant correlations with most traits, including SG, NL, CP%, LB, LSR, HCN, PH, and PP. It was negatively correlated with TSS, DFY, and IVDMD, though these were not significant. Leaf breadth (LB) correlated positively and significantly with PP (0.2981**) and NL (0.278*), suggesting these traits can serve as indirect selection indices for denser foliage. LB also showed positive non-significant correlations with SG, PH, LL, CP%, HCN, and DFY, while negative correlations were seen with TSS, LSR, and IVDMD. Traits like leaf stem ratio (LSR) and stem girth (SG) generally exhibited weak correlations with other traits, mostly non-significant. LSR was positively associated with IVDMD, CP%, and LL, but negatively related to PH, DFY, LB, and PP. SG showed only weak, non-significant positive correlations with most traits and weak negative correlations with TSS and CP%. TSS emerged as a trait of moderate interest due to its significant positive correlation with PH and DFY.

The crude protein content (CP%) showed uniformly weak, non-significant correlations with all traits, both positive and negative, suggesting its independence from major agronomic traits under current environmental and genetic conditions. IVDMD was positively, though non-significantly, associated with LSR, HCN, TSS, CP%, and SG. Its negative correlation with GFY, DFY, PH, LL, LB, and NL suggests that improvements in yield might slightly compromise digestibility, although not significantly. Finally, HCN exhibited a significant positive association with PP (0.2601*), highlighting a potential trade-off between plant density and toxicity. Other associations were non-significant, with both positive (e.g., IVDMD, SG, NL, LL) and negative trends (e.g., PH, TSS, CP%, DFY).

The study highlights GFY and DFY as central traits with strong positive correlations to PH and TSS, suggesting these as key contributors to biomass production. The significant correlations among certain morphological and quality traits offer potential for simultaneous trait improvement through indirect selection, supporting the formulation of ideotypes in forage sorghum breeding



	GFY	DFY	PH	NL	LL	LB	LSR	PP	SG	TSS	CP%	IVDMD	HCN
0	1	0.99	0.54	0.007	-0.085	-0.006	-0.14	0.057	0.04	0.24	-0.058	-0.16	-0.062
1	0.99	1	0.55	0.019	-0.086	0.006	-0.14	0.055	0.024	0.23	-0.062	-0.16	-0.052
2	0.54	0.55	1	0.053	0.071	0.18	-0.19	0.064	0.005	0.25	-0.016	-0.14	-0.18
3	0.007	0.019	0.053	1	0.2	0.28	0.1	0.17	0.088	0.033	-0.034	-0.024	0.094
4	-0.085	-0.086	0.071	0.2	1	0.14	0.11	0.023	0.22	-0.098	0.14	-0.038	0.081
5	-0.006	0.006	0.18	0.28	0.14	1	-0.04	0.3	0.2	-0.099	0.092	-0.026	0.054
6	-0.14	-0.14	-0.19	0.1	0.11	-0.04	1	-0.019	0	0.063	0.16	0.21	0.062
7	0.057	0.055	0.064	0.17	0.023	0.3	-0.019	1	0.22	-0.011	-0.04	-0.18	0.26
8	0.04	0.024	0.005	0.088	0.22	0.2	0	0.22	1	-0.13	-0.092	0.053	0.085
9	0.24	0.23	0.25	0.033	-0.098	-0.099	0.063	-0.011	-0.13	1	0.15	0.12	-0.1
10	-0.058	-0.062	-0.016	-0.034	0.14	0.092	0.16	-0.04	-0.092	0.15	1	0.068	-0.066
11	-0.16	-0.16	-0.14	-0.024	-0.038	-0.026	0.21	-0.18	0.053	0.12	0.068	1	0.15
12	-0.062	-0.052	-0.18	0.094	0.081	0.054	0.062	0.26	0.085	-0.1	-0.066	0.15	1

Fig. 1. Correlation plot for various agro-morphological and biochemical traits.

1.1. Path Coefficient

The phenotypic path coefficient analysis aimed to understand the direct and indirect effects of various traits on dry fodder yield (DFY) as the dependent variable. The analysis revealed a very low residual value of 1.067%, indicating that the independent variables collectively explained 98.933% of the variation in DFY (Table 3). This highlights the robustness of the model and the relevance of the included traits in influencing fodder yield. Among the traits studied, green fodder yield (GFY) exhibited the highest and most significant positive correlation with DFY (r = 0.994**). It also recorded a strong direct effect (0.98601), indicating that GFY is a key determinant of DFY. Additionally, GFY showed small positive indirect effects via plant height (PH), leaf length (LL), crude protein (CP%), and number of leaves per plant (NL), and minor negative indirect effects through traits like IVDMD, leaf breadth (LB), plant population (PP), stem girth (SG), hydrocyanic acid (HCN), leaf stem ratio (LSR), and total soluble solids (TSS). Similar findings were reported by Deep *et al.* (2019) and Ramdevsinh (2010).

Plant height (PH) also showed a significant positive correlation with DFY (r = 0.5482**) and had a small but positive direct effect (0.02593). Its indirect effects were mostly positive via GFY, LB, NL, and CP%, while small negative effects were observed through IVDMD, SG, LL, PP, LSR, HCN, and TSS. This implies that PH influences DFY more through its positive association with other contributing traits. Traits such as number of leaves (NL) and leaf breadth (LB) showed positive, though non-significant, correlations with DFY. NL had a small positive direct effect (0.01117) and exhibited favourable indirect effects through GFY, LB, PH, HCN, LSR, LL, PP, SG and TSS%. Similarly, LB had a direct effect of 0.0095 and contributed positively through PH, NL, TSS, and HCN, with minimal negative indirect contributions via other traits. Similar observations were also reported by Reddy (2023).

Interestingly, leaf length (LL) exhibited a negative direct effect on DFY (-0.00555) and an overall negative correlation, mainly due to a large negative indirect effect via GFY. Although some indirect effects through traits like PH, TSS, and LB were positive, they were insufficient to offset the dominant negative pathways. Leaf stem ratio (LSR) had a low positive direct effect (0.00626) but a negative correlation with DFY due to major negative indirect effects through

GFY and PH. Its positive influences via traits like NL, HCN, PP, and IVDMD were relatively minor. The direct effect of plant population (PP) was negative (-0.00753), though the trait showed a weak positive correlation with DFY. Its positive indirect effects via GFY, HCN, and LB highlight complex interplays, while negative pathways through SG, LL, and LSR contributed to the overall diminished impact. These findings are in accordance with Bairwa (2018) and Uttam (2021).

Stem girth (SG) had a negative direct effect on DFY (-0.01912), despite showing a weak positive correlation. However, its indirect effects via GFY and TSS were notable, pointing to its limited independent contribution. Notably, TSS had a significant positive correlation (r = 0.2249*) but a negative direct effect (-0.01546), suggesting its contribution to DFY is primarily indirect-especially through GFY, PH, and SG. The crude protein content (CP%) had a negative direct effect (-0.00388) and overall weak, negative correlation with DFY, primarily driven by negative indirect effects *via* GFY, HCN, and TSS. Its few positive pathways through SG, LSR, and LB was relatively minor. IVDMD showed a very small positive direct effect (0.00023) and a negative correlation with DFY, primarily due to its strong negative indirect effect through GFY. However, it showed marginal positive effects through HCN, PP, and LSR. Finally, HCN had a positive direct effect (0.01341) but an overall weak and negative correlation with DFY, driven by negative indirect effects *via* GFY and PH, although some minor positive effects were noted through TSS, NL, and LB.

GFY emerged as the most influential trait for DFY due to its dominant direct and indirect effects. Traits like PH, TSS, and LB contributed indirectly, while traits such as LL, PP, SG, and CP% exerted negative influences. The analysis clearly identifies GFY and PH as strong candidates for selection when targeting higher dry fodder yield in forage sorghum breeding programs. Similar findings were also reported by Deep *et al.* (2019), Uttam (2021), Kumari *et al.* (2023) and Kumari *et al.* (2024).

1.2. Cluster Analysis

Cluster analysis using Tocher's method based on Mahalanobis D^2 statistics effectively grouped the sorghum genotypes into distinct clusters, reflecting the extent of genetic diversity among them. The genotypes were distributed into eleven clusters (Tables 4-6; Fig. 2). Cluster 1 had the largest number of genotypes (10), indicating a high degree of genetic similarity among the genotypes in this group. It was followed by Cluster 4 and Cluster 3, which included a moderate number of genotypes (4 and 3, respectively). In contrast, Clusters 6, 7, 8, 9, 10, and 11 each contained only one genotype, suggesting that these genotypes are highly unique and genetically divergent from the others. Trait-wise contribution to genetic diversity was evaluated. Among all the traits, plant height (PH) contributed the most to the total genetic divergence, accounting for a substantial 4180.17% of the diversity, followed by green fodder yield (GFY) with 2555.43%. This highlights that traits like PH and GFY are the most influential in differentiating the genotypes and should be prioritized in selection for genetic improvement. Cluster mean values provided further insight into trait-specific performance across the groups. For green fodder yield (GFY), the highest cluster mean was observed in Cluster 1 (58.10), suggesting this cluster holds potential for selecting high biomass genotypes, while the lowest was found in Cluster 8 (35.63). Regarding dry fodder yield (DFY), values ranged from 14.63 in Cluster 7 to 8.8 in Cluster 3, indicating variability in total biomass accumulation and dry matter content.

Plant height (PH) ranged dramatically across clusters, with the tallest genotypes grouped in Cluster 2 (274.04) and the shortest in Cluster 3 (131.82). Number of leaves per plant (NL) was highest in Cluster 7 (24.0), indicating superior foliage density, and lowest in Cluster 8 (15.33). For leaf length (LL), Cluster 2 again recorded the highest mean (91.56), while Cluster 11 showed the lowest (64.33). Leaf breadth (LB) was greatest in Cluster 4 (8.04) and least in Cluster 3 (5.5), indicating that broader-leaf genotypes were grouped in Cluster 4. The leaf-stem ratio (LSR), important for fodder quality, was highest in Cluster 5 (0.318) and lowest in Cluster 10 (0.267). Plant population at harvest (PP) ranged from 9.17 in Cluster 11 to 4.5 in Cluster 3, reflecting variation in stand establishment and survivability. Stem girth (SG) was maximum in Cluster 4 (21.20) and minimum in Cluster 6 (14.63), suggesting structural differences that may relate to lodging resistance. For total soluble solids (TSS), an indicator of sugar content, the maximum was recorded in Cluster 7 (7.56) and minimum in Cluster 9 (6.41). Crude protein (CP%) ranged from 7.35 in Cluster 2 to 6.33 in Cluster 3, showing moderate variability across clusters. *In vitro* dry matter digestibility (IVDMD) was highest in Cluster 3 (33.2), marking it as promising for forage quality, while the lowest value was observed in Cluster 6 (25.73). Notably, hydrocyanic acid (HCN) content varied widely, with the maximum in Cluster 9 (108.76) and the minimum in Cluster 6 (35.28). This suggests that Cluster 6 genotypes are potentially safer for feeding due to lower toxicity risk. Similar results were reported by Laxmi *et al.* (2019a).

The clustering revealed substantial genetic variation among the sorghum genotypes. Cluster 1 and 2 appear promising for yield traits, while Clusters 3 and 6 are notable for quality traits like digestibility and low HCN content. These findings support the potential for selecting diverse parents from distant clusters to maximize heterosis and develop superior forage sorghum varieties.

1.1. Principal component analysis

Principal Component Analysis (PCA) was carried out using 13 morpho-agronomic and quality

TABLE 3

Path coefficient analysis for morphological and quality parameters based on DFY

Trait	GFY	PH	NL	LL	LB	LSR	PP	SG	TSS	CP%	IVDMD	HCN	DFY
GFY	0.9860	0.0139	0.0001	0.0005	-0.0001	-0.0009	-0.0004	-0.0008	-0.0037	0.0002	0.0000	-0.0008	0.994**
PH	0.5283	0.0259	0.0006	-0.0004	0.0018	-0.0012	-0.0005	-0.0001	-0.0039	0.0001	0.0000	-0.0024	0.5482**
NL	0.0066	0.0014	0.0112	-0.0011	0.0026	0.0006	-0.0013	-0.0017	-0.0005	0.0001	0.0000	0.0013	0.0192 NS
LL	-0.0842	0.0018	0.0022	-0.0055	0.0013	0.0007	-0.0002	-0.0043	0.0015	-0.0006	0.0000	0.0011	-0.0861 NS
LB	-0.0060	0.0048	0.0031	-0.0008	0.0095	-0.0002	-0.0022	-0.0037	0.0015	-0.0004	0.0000	0.0007	0.0063 NS
LSR	-0.1423	-0.0049	0.0011	-0.0006	-0.0004	0.0063	0.0001	0.0000	-0.0010	-0.0006	0.0000	0.0008	-0.1413 NS
PP	0.0564	0.0017	0.0020	-0.0001	0.0028	-0.0001	-0.0075	-0.0042	0.0002	0.0002	0.0000	0.0035	0.0547 NS
SG	0.0394	0.0001	0.0010	-0.0012	0.0019	0.0000	-0.0016	-0.0191	0.0020	0.0004	0.0000	0.0011	0.0239 NS
TSS	0.2329	0.0065	0.0004	0.0005	-0.0009	0.0004	0.0001	0.0025	-0.0155	-0.0006	0.0000	-0.0013	0.2249*
CP%	-0.0571	-0.0004	-0.0004	-0.0008	0.0009	0.0010	0.0003	0.0018	-0.0024	-0.0039	0.0000	-0.0009	-0.0619 NS
IVDMD	-0.1596	-0.0037	-0.0003	0.0002	-0.0002	0.0013	0.0013	-0.0010	-0.0019	-0.0003	0.0002	0.0020	-0.1619 NS
HCN	-0.0608	-0.0046	0.0010	-0.0005	0.0005	0.0004	-0.0020	-0.0016	0.0016	0.0003	0.0000	0.0134	-0.0523 NS

TABLE 4

Estimates of Intra- and Inter-Cluster Distances

Cluster Number	Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5	Cluster 6	Cluster 7	Cluster 8	Cluster 9	Cluster 10	Cluster 11
Cluster 1	8.75	28.24	19.95	19.14	16.60	15.27	21.37	13.90	16.44	11.15	33.98
Cluster 2	28.24	5.51	19.13	33.26	24.04	17.38	16.41	21.90	14.05	34.51	25.13
Cluster 3	19.95	19.13	8.89	18.13	16.90	25.57	11.19	11.12	20.25	25.91	18.86
Cluster 4	19.14	33.26	18.13	8.47	17.02	29.78	18.99	25.08	33.78	13.38	28.22
Cluster 5	16.60	24.04	16.90	17.02	7.61	26.81	15.92	22.48	20.02	14.90	12.45
Cluster 6	15.27	17.38	25.57	29.78	26.81	0.00	26.22	21.37	16.18	18.36	46.29
Cluster 7	21.37	16.41	11.19	18.99	15.92	26.22	0.00	17.04	18.80	27.73	17.87
Cluster 8	13.90	21.90	11.12	25.08	22.48	21.37	17.04	0.00	13.06	26.95	30.08
Cluster 9	16.44	14.05	20.25	33.78	20.02	16.18	18.80	13.06	0.00	29.06	35.73
Cluster 10	11.15	34.51	25.91	13.38	14.90	18.36	27.73	26.95	29.06	0.00	31.91
cluster 11	33.98	25.13	18.86	28.22	12.45	46.29	17.87	30.08	35.73	31.91	0.00

TABLE 5

Cluster wise mean for various traits under study

Trait	Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5	Cluster 6	Cluster 7	Cluster 8	Cluster 9	Cluster 10	Cluster 11
GFY	58.10	57.23	36.27	45.57	40.33	44.40	57.67	35.63	47.08	40.07	56.07
DFY	14.25	13.93	8.80	11.31	10.10	10.90	14.63	9.17	11.72	10.10	13.50
PH	249.78	274.04	131.82	220.85	218.56	249.52	271.18	142.78	175.74	227.61	198.63
NL	21.38	23.56	21.67	23.26	20.63	20.33	24.00	15.33	20.59	18.22	21.67
LL	74.21	91.56	80.00	76.59	84.28	75.78	79.00	70.78	73.59	72.56	64.33
LB	6.55	6.22	5.50	8.04	6.44	5.74	7.20	5.87	6.35	5.66	6.11
LSR	0.28	0.29	0.30	0.28	0.32	0.29	0.28	0.28	0.30	0.27	0.30
PP	6.27	7.73	4.50	8.25	6.97	6.39	7.07	6.89	8.12	6.39	9.17
SG	16.74	18.27	18.71	21.20	17.42	14.63	18.22	16.46	17.50	17.46	15.89
TSS	7.22	7.44	6.78	6.74	7.17	6.45	7.56	7.33	6.41	7.00	7.22
CP	6.77	7.35	6.33	6.65	6.98	6.47	7.00	7.00	6.77	6.78	6.76
IVDMD	28.97	28.40	33.20	31.11	26.07	25.73	31.20	29.07	30.67	30.33	27.07
HCN	59.22	81.16	63.42	77.01	78.10	35.28	45.47	74.02	108.76	68.41	75.24

TABLE 6
List of genotypes along with Cluster sequence

Cluster	No. of Genotypes	Genotypes
Cluster 1	10	SPH1984, SPH1986, SPH1987, SPH1988, SPV2796, SPV2802, SPV2803, SPV2806, SPV2808, CSH36F
Cluster 2	2	SPH1985, SPV2798
Cluster 3	3	SPH1989, CSV 30F, CSV 35F
Cluster 4	3	SPV2805, SPV2809, CSH 40F
Cluster 5	2	SPV2801, HJ 541 (Local Check)
Cluster 6	1	SPV2797
Cluster 7	1	SPV2799
Cluster 8	1	SPV2800
Cluster 9	1	SPV2804
Cluster 10	1	SPV2807
Cluster 11	1	CSV 21F

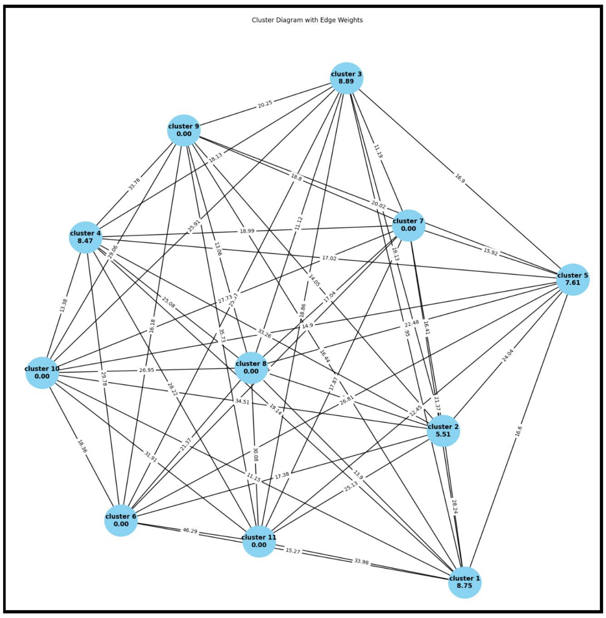


Fig. 2. Clustering of forage sorghum genotypes based on modified Tocher method.

TABLE 7
PCA results showing component-wise variance and cumulative contribution

Principal component	S. D.	Proportion of variance	Cumulative proportion
PC1	1.8363	25.938	25.938
PC2	1.5968	19.614	45.552
PC3	1.2648	12.306	57.858
PC4	1.1447	10.08	67.938
PC5	1.0735	8.865	76.803
PC6	0.9118	6.395	83.197
PC7	0.7804	4.684	87.882
PC8	0.7691	4.55	92.432
PC9	0.5936	2.71	95.143
PC10	0.5466	2.298	97.44
PC11	0.4589	1.62	99.06
PC12	0.3455	0.918	99.979
PC13	0.0525	0.021	100

traits to explore the underlying structure of variability among the genotypes. The analysis employed the singular value decomposition approach and extracted 13 principal components (PCs). According to Kaiser's criterion, only those PCs with eigenvalues greater than 1 were considered significant for further interpretation. In this case, five components met the criterion, cumulatively explaining a substantial portion of the total variance (Table 7; Figure. 3 & 4).

The first principal component (PC1) explained the highest proportion of variation, accounting for 25.94% of the total variability with an eigenvalue of 1.84 (Figure 5 & 6). This component was positively associated with HCN (0.295), LSR (0.2421), stem girth (SG) (0.1805), plant population (PP) (0.1655), IVDMD (0.1373), leaf length (LL) (0.09), and leaf breadth (LB) (0.0329). However, it showed strong negative correlations with crude protein (CP%) (-0.017), number of leaves (NL) (-0.0577), TSS (-0.34), plant height (PH) (-0.4518), green fodder yield (GFY) (-0.4688), and dry fodder yield (DFY) (-0.4724), indicating that this component captured a contrast between yield-related traits and traits associated with plant structure and HCN toxicity. Similar results were also reported by Kumari *et al.*, 2024.

PC2 accounted for 19.61% of the total variance and had weak positive correlations with TSS (0.0259) and LSR (0.0053). On the other hand, it showed moderate to strong negative correlations with several traits, most notably NL (-0.5226), LB (-0.5179), PP (-0.4184), LL (-0.2747), SG (-0.2575), and HCN (-0.2434). These associations suggest that PC2 primarily captured variation related to foliage and structural dimensions, with a negative association with plant population and leaf traits.

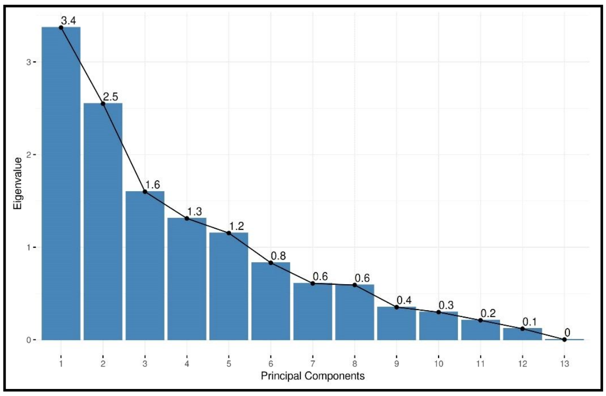
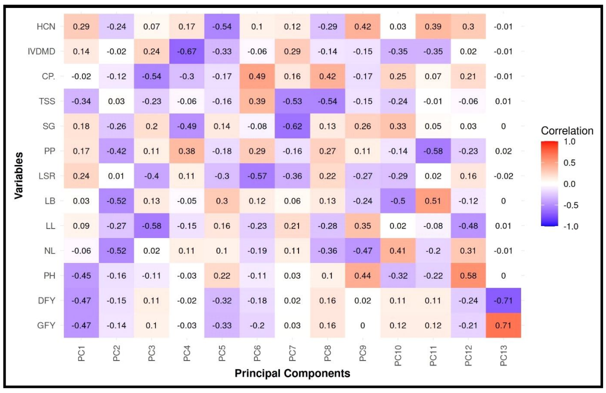


Fig. 3. Scree plot illustrating the contribution of principal components.



	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8	PC9	PC10	PC11	PC12	PC13
HCN	0.29	-0.24	0.07	0.17	-0.54	0.1	0.12	-0.29	0.42	0.03	0.39	0.3	-0.01
IVDMD	0.14	-0.02	0.24	-0.67	-0.33	-0.06	0.29	-0.14	-0.15	-0.35	-0.35	0.02	-0.01
CP.	-0.02	-0.12	-0.54	-0.3	-0.17	0.49	0.16	0.42	-0.17	0.25	0.07	0.21	-0.01
TSS	-0.34	0.03	-0.23	-0.06	-0.16	0.39	-0.53	-0.54	-0.15	-0.24	-0.01	-0.06	0.01
SG	0.18	-0.26	0.2	-0.49	0.14	-0.08	-0.62	0.13	0.26	0.33	0.05	0.03	0
PP	0.17	-0.42	0.11	0.38	-0.18	0.29	-0.16	0.27	0.11	-0.14	-0.58	-0.23	0.02
LSR	0.24	0.01	-0.4	0.11	-0.3	-0.57	-0.36	0.22	-0.27	-0.29	0.02	0.16	-0.02
LB	0.03	-0.52	0.13	-0.05	0.3	0.12	0.06	0.13	-0.24	-0.5	0.51	-0.12	0
LL	0.09	-0.27	-0.58	-0.15	0.16	-0.23	0.21	-0.28	0.35	0.02	-0.08	-0.48	0.01
NL	-0.06	-0.52	0.02	0.11	0.1	-0.19	0.11	-0.36	-0.47	0.41	-0.2	0.31	-0.01
PH	-0.45	-0.16	-0.11	-0.03	0.22	-0.11	0.03	0.1	0.44	-0.32	-0.22	0.58	0
DFY	-0.47	-0.15	0.11	-0.02	-0.32	-0.18	0.02	0.16	0.02	0.11	0.11	-0.24	-0.71
GFY	-0.47	-0.14	0.1	-0.03	-0.33	-0.2	0.03	0.16	0	0.12	0.12	-0.21	0.71

Fig. 4. Heat map depicting correlation between principal components and various variables.

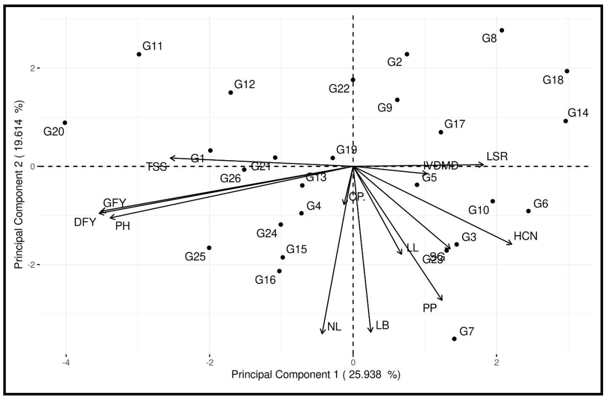


Fig. 5. Principal Component Biplot Illustrating Trait Vectors and Genotypic Dispersion.

PC3, which explained 12.31% of the variability, was positively correlated with IVDMD (0.2447), SG (0.2048), LB (0.1332), PP (0.11), DFY (0.1055), GFY (0.102), and HCN (0.0655). This component highlighted traits relevant to yield quality and plant robustness. However, it was negatively correlated with PH (-0.1093), TSS (-0.2254), LSR (-0.3993), CP% (-0.5362), and particularly LL (-0.5764), showing a contrast between digestibility traits and some structural leaf and biochemical attributes.

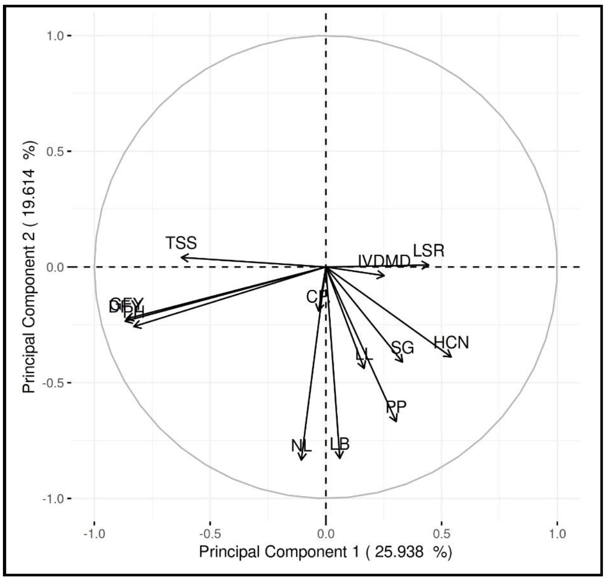


Fig. 6. Trait contribution to principal components in PCA.

The fourth component (PC4) accounted for 10.08% of the variance and exhibited positive correlations with PP (0.3768), HCN (0.1651), NL (0.113), and LSR (0.1068), suggesting its importance in capturing variation related to plant density and toxicity. Negative correlations were observed with most other traits, including SG (-0.4945) and IVDMD (-0.6654), indicating an inverse relationship between forage digestibility and plant structural or toxicological parameters.

Lastly, PC5 contributed 8.87% of the total variation and showed positive correlations with LB (0.3027), PH (0.2246), LL (0.1601), SG (0.143), and NL (0.0999). Conversely, it was negatively associated with GFY (-0.3342), DFY (-0.3212), IVDMD (-0.3294), HCN (-0.545), and LSR (-0.3029), among others. This component reflects a contrast between leaf and stem dimensions and yield-quality traits.

In conclusion, the PCA successfully grouped the traits into components that capture major axes of variation within the sorghum genotypes. The first few components explain a significant proportion of the variability, with PC1 and PC2 in particular highlighting trade-offs between biomass yield and structural or quality traits. These insights are valuable for multivariate selection and trait prioritization in sorghum breeding programs and in line as reported by Laxmi *et al.*, 2019b.

CONCLUSION

The considerable morphological heterogeneity observed among genotypes for key agronomic and quality traits-such as green and dry fodder yield, plant height, leaf dimensions, and digestibility-indicates substantial scope for improvement through targeted hybridization and selection for the desired traits. Trait association studies further highlighted the central role of green fodder yield (GFY) and dry fodder yield (DFY) as primary contributors to biomass productivity. Multivariate analyses further support the identification of genetically divergent, high-performing genotypes like SPH1984, SPH1986 and SPV2799 suitable for use as parents in breeding programs. The study thus lays a strong foundation for developing improved forage sorghum cultivars that integrate high productivity, nutritional quality, and adaptability using improved cultivars like SPV2809, SPV2800 and SPH1984.

REFERENCES

Altaf, M. T., W. Liaqat, A. Jamil, M. F. Jan, and F. S. Baloch, 2024 : A bibliometric analysis of genome-wide association study (GWAS) and sorghum (*Sorghum bicolor* L) based on Web of Science using VOS viewer. *Physiology and Molecular Biology of Plants,* https://doi.org/10.1007/s42729-024-01888-6

Bairwa, L. L., 2018 : *Variability, correlation and path analysis in multicut forage sorghum* [Master's thesis, Rajasthan College of Agriculture, MPUAT].

https://krishikosh.egranth.ac.in/handle/1/5810051525

Burton, G. W., and D. E. DeVane, 1953 : Estimating heritability in tall fescue (*Festuca arundinacea*) from replicated clonal material. *Agronomy Journal,* **45**: 478-481.

Chavan, S., B. V. V. Prasad, K. Kasanaboina and C. V. S. Kumar, 2024 : Identification of traits influencing forage yield through correlation and path analysis in sorghum. *Heliyon*, 10: e15838.

da Silva, A. R., 2021: *biotools: Tools for biostatistics* [R package]. GitHub https://github.com/arsilva87/biotools

Deep, H., S., Arya, P., Kumari, S. K. Pahuja and J. Tokas, 2019 : Genetic parameters, correlation and path coefficient analysis for fodder yield and quality in forage sorghum. *Green Farming*.

Dewey, D. R., and K. Lu, 1959 : A correlation and path coefficient analysis of components of crested wheatgrass seed production. *Agronomy Journal,* **51**(9): 515-518.

Engida, C. E., 2023 : Genetic diversity in yield traits and kernel composition of selected Ethiopian sorghum landraces. *University of the Free State Repository*. https://scholar.ufs.ac.za/handle/11660/12184.

Gebre, W., F. Mekbib, A. Tirfessa, and A. Bekele, 2025 : Phenotypic Diversity of Sorghum Genotypes based on Qualitative and Quantitative Traits. *Heliyon*, 11, e00758.

Johnson, H. W., H. F. Robinson, and R. E. Comstock, 1955 : Estimates of genetic and environmental variability in soybeans. *Agronomy Journal,* **47**(7): 314–318.

Kumar, K., S. Chavan, B. V. V. Prasad, and C. V. S. Kumar, 2024 : Identification of stable restorers and high-yielding hybrids using diverse sorghum male sterile cytoplasmic sources and established pollen parents. *Heliyon*, 10, e15838.

Kumari, P., D. S. Phogat, Satpal, N. Kharor, and S.K. Pahuja, 2023 : Breeding climate smart forage sorghum – A review. *Forage Research,* **49**(1): 1-12.

Kumari, P., N. Kharor, S. K. Pahuja, and D. S. Phogat, 2024 : Morphological and biochemical traits attribution for fodder yield and quality in forage sorghum using principal component analysis. *Vegetos*.

https://doi.org/10.1007/s42535-024-01017-9

Laxmi, V., S.K. Pahuja and P. Kumari, 2019b : Identification of new sources for good quality high biomass yield and other promising traits in mini core collection of forage Sorghum. *Indian J. Plant Gen. Resources*, **32**(2): 150-157.

Laxmi, V., S.K. Pahuja, P. Kumari and U.N. Joshi, 2019a : Genetic divergence studies for agro-morphological, insect pest and quality parameters in mini core collection of forage sorghum. *Forage Research*, **44**: 237-241.

Le, S., J. Josse, and F. Husson, 2008 : FactoMineR: An R package for multivariate analysis. *Journal of Statistical Software*, **25**(1): 1-18. *https://doi.org/10.18637/jss.v025.i01*.

Mahalanobis, P. C., 1936 : On the generalized distance in statistics. *Proceedings of the National Institute of Sciences of India*, **2**(1): 49-55. https://www.jstor.org/stable/25047632.

Panse, V. G., and Sukhatme, P. V. 1954: *Statistical methods for agricultural workers*.

Popat, R., A. Tiwari, and D. Chauhan, 2020 : *Variability: Genetic variability analysis in R* [R package version]. Comprehensive R Archive Network (CRAN). Available at: https://cran.r-project.org/package=variability

Ramdevsinh, R. H., 2010 : *Study of variability, correlation and path analysis in forage sorghum* [Master's thesis, Junagadh Agricultural University]. KrishiKosh Repository.

Reddy, M. S., 2023 : *Biochemical analysis of forage sorghum under different fertilizer levels* [M.Sc. thesis, PJTSAU].

Rohila, N., 2016 : *Morphological and biochemical characterization of various forage sorghum genotypes* (*Sorghum bicolor* L. Moench). CCSHAU, Hisar.

Rohila, N., S. Arya, and P. Kumari, 2020 : Correlation and path coefficient analysis of various sorghum genotypes (*Sorghum bicolor* L. Moench) for quality traits. *Electronic Journal of Plant Breeding,* **11**(1): 306-309.

Singh, K. B., and B. D. Chaudhary, 1977 : *Biometrical methods in quantitative genetic analysis* (1st ed.). Kalyani Publishers.

Tirfessa, A., *et al.,* 2025 : Use of cluster and principal component analyses to characterize sorghum genotypes under drought-prone environments. *Heliyon*, 11: e00758. https://doi.org/10.1016/j.heliyon.2025.e00758.

Uttam, M. B. R., 2021 : *Variability and path coefficient analysis in forage sorghum* [M.Sc. thesis, Dr. PDKV, Akola].

https://krishikosh.egranth.ac.in/handle/1/7010126384

Wickham, H., 2016 : *ggplot2: Elegant graphics for data analysis*. Springer. https://doi.org/10.1007/978-3-319-24277-4.