

Multivariate Cluster and Principal Component Analyses on Yield Traits in Foxtail Millet (*Setaria italica* (L.) P. Beauv.) Genotypes

DP Rao¹  and HP Chaturvedi^{1*} 

Received: 09th November 2023 / Accepted: 26th September 2025

ABSTRACT

Purpose: The research aims to employ Multivariate Cluster and Principal Component Analyses to examine yield traits in Foxtail Millet genotypes cultivated in Nagaland's foothills. It addresses the need to identify distinctive groups among millet varieties and discern the principal components influencing yield. This investigation seeks to enhance millet-breeding strategies for improved productivity in the region's specific agro-climatic conditions.

Research Method: The investigation was carried out from July 2022 to May 2023 for four different dates of sowing with a twenty-five days interval. Two environments were maintained under rain fed conditions and the remaining two environments were maintained under irrigated conditions at seven days intervals. The experiment was conducted in a randomized complete block design with three replications in all environments.

Findings and Values: The Elbow Method identified four clusters, with Cluster-1 (16.6%) having five genotypes, Cluster-2 (20.0%) six, Cluster-3 (43.3%) thirteen, and Cluster-4 (20.0%) six. The maximum inter-cluster distance was between Cluster-2 and Cluster-4 (33.09). Principal component analysis revealed five components (eigenvalues >1) explaining 75.8% of total variance, while correlation analysis showed nine significant positive and three negative associations among yield traits. Ward's dendrogram grouped the genotypes into five clusters, with Cluster-1 comprising superior performers are G1 (ERP 26), G5 (ESD 3), G18 (FOX 4336), G22 (FOX 4386), and G25 (FOX 4394), that consistently excelled in grain yield and yield-contributing traits.

Keywords: Foxtail millet, K means, Multivariate analysis, PCA and Ward's method

INTRODUCTION

Foxtail millet (*Setaria italica* (L.) P. Beauv.) is a nutritious cereal belonging to *Poaceae*, a subfamily of *Panicoideae* and genus *Setaria*. It is a valuable crop and belongs to the tribe *Paniceae* (Moharil *et al.* 2019). Foxtail millet is believed to have originated in East Asia, particularly in China and Korea. Its origins can be traced back around 7,400 to 7,900 years ago in the Yellow River Valley of northern China. Foxtail millet is significant in human history and continues to nourish diverse cultures globally (Ali *et al.* 2016). Today, it is cultivated in various tropical and subtropical regions worldwide, including parts of America, Australia, and Asia, although its cultivation

level remains relatively low compared to other staple crops (Taylor 2019). It is a staple food with a rich history, widely grown for nutrition, silage, and hay (Hanna *et al.* 2004).

Foxtail millet is an important crop grown in several parts of the world. According to the Food and Agriculture Organization (FAO) of the United Nations and the International Crops Research Institute for the Semi-Arid Tropics, the global production of foxtail millet was estimated

¹Department of Genetics and Plant Breeding, School of Agricultural Sciences, Nagaland University, Medziphema Campus-797 106, Nagaland, India.

*hpchaturvedi68@gmail.com

to be around 6 million tons in 2023, with India being the largest producer, accounting for more than 50% of the total production (Rao & Chaturvedi 2022). In India, it is cultivated over an area of 87000 ha with a total production of about 0.066 metric tons and with a productivity of 762 kg/ha during 2015-16 (Pallavi *et al.* 2020). The International Year of Millets 2023, a United Nations initiative, aims to raise awareness about the significance of millets as a nutritious and sustainable food source while promoting their cultivation, consumption, and trade. It highlights millets' role in food and nutritional security and their contribution to the Sustainable Development Goals. The year 2023 serves as a platform to share knowledge, best practices, and innovations in millet farming, processing, marketing, and consumption.

Foxtail millet cultivars are selected based on critical factors such as yield and resistance to biotic and abiotic stress (Borrayo *et al.* 2016). However, creating a single cultivar that possesses all of these desirable traits is a complex task, primarily due to the small size of foxtail millet flowers, which makes hybridization challenging (Mokkaraj & Geethanjali 2016). K-means clustering and PCA are indispensable tools in the selection process for trait analysis related to foxtail millet cultivar development. K-means clustering helps group cultivars with similar trait profiles, facilitating parent selection and trait prioritization. At the same time, PCA reduces data complexity, identifies key traits, and aids in visualizing trait relationships, thereby informing more effective cultivar development strategies (Rao & Chaturvedi 2022). These techniques collectively contribute to developing foxtail millet cultivars with enhanced and balanced trait profiles.

Genetic variation can be assessed in two ways: using either univariate or multivariate analysis. Traditionally, the favoured method has been univariate analysis for understanding genetic differences. Multivariate analysis, including K-means clustering and PCA, has become popular in plant breeding due to its ability to handle complex and diverse trait data, aid in parent selection, reduce data dimensionality and provide insights that lead to more efficient and informed cultivar development. These methods

are essential tools for modern plant breeders looking to create improved crop varieties that meet the evolving demands of agriculture and food security (Jarwar *et al.* 2019).

Cluster analysis, including the K-means algorithm, is a statistical tool for studying population variability. K-means is a non-hierarchical approach that aims to maximize variation between K-formed groups and minimize variation within each group (Barth *et al.*, 2022). Determining the predetermined number of groups is a complex task that can lead to imprecise analyses. To mitigate this issue, the Elbow Method is employed to ascertain the optimal cluster count (Syakur *et al.* 2018). PCA is good at putting each genotype into one group and shows which traits contribute the most to the overall diversity on each axis. Because of this, breeders often use PCA to assess genetic diversity and identify promising genotypes for future breeding programs (Sharma 2006).

This study aimed to evaluate various foxtail millet cultivars over multiple environments to access variations in yield and yield traits. This was done to understand better how to improve foxtail millet cultivars in terms of both yield and Yield traits.

MATERIALS AND METHODS

Experiment location

The research took place between July 2022 and May 2023 in four different environments, and it was carried out on four different dates of sowing (Table 1). Each sowing date was chosen to create varying environmental conditions, including different temperatures and moisture levels, throughout the crop growth stages. Two environments were maintained under rain fed condition and other two environments were maintained under irrigated condition at irrigation intervals are once in week. The experiment was conducted at the Research Farm of the Department of Genetics and Plant Breeding School of Agricultural Sciences, Nagaland University, Medziphema, India. The precise coordinates of the research farm are “25°45'35" N and 95°25'45" E” with an altitude of 310 m above mean sea level.

Soil sampling and analysis

The top layer of soil, up to a depth of 0-15 cm, was randomly collected from the field in all four environments. To gather the soil samples, a combination of a sharp tool called a cutlass and a small hand trowel was used. Once the samples were collected, they were combined and organized based on their location to create a composite sample after the experiment was completed. This composite sample was then taken to the university laboratory for analysis. To ensure consistency in the distribution of nutrients and to accurately represent the plots, the samples were dried in a shaded area and ground using a glass mortar and pestle. After this process, the sample was sifted, and various tests were conducted to analyze its chemical properties and particle size distribution. These tests included examining the levels of sand, clay, silt, pH, organic carbon (OC),

available nitrogen (N), available potassium (K) and available phosphorus (P). Results are represented at Table 2.

Plant materials

We obtained a collection of one hundred foxtail millet samples, which included reference varieties, from the Indian Institute of Millets Research (IIMR) in Hyderabad. These samples were evaluated during the *Zaid* season in 2022 at same environment. Based on the results of this evaluation, we identified the top 30 genotypes that showed the highest grain yield specifically in this *Nagaland* region. These 30 selected samples were used in our study to assess genetic variability, diversity, and stability across different environments. List of 30 genotypes presented in Table 3.

Table 1. Environmental description of the experimental site.

Code	Sowing date	Season	Latitude	Longitude	Altitude	Av. Temp		Av. Hum(%)		Rainfall (mm)	Year
						Max	Min	Max	Min		
Env-1	01-07-2022	Kharif	25° 45' 15.95" N	93° 51' 44.71 E	310 MSL	31.66	22.30	91.75	69.64	51.92	2022
Env-2	26-07-2022	Kharif(Late)	25° 45' 15.95" N	93° 51' 44.71 E	311 MSL	32.09	22.84	92.10	69.99	55.19	2022
Env-3	01-01-2023	Summer	25° 45' 15.95" N	93° 51' 44.71 E	312 MSL	29.11	17.40	94.48	61.84	15.58	2023
Env-4	26-01-2023	Summer(Late)	25° 45' 15.95" N	93° 51' 44.71 E	313 MSL	28.28	15.97	95.29	60.11	8.46	2023

Env=Environment, Av. Temp= Average temperature, Av. Hum=Average humidity

Table 2. Characterization of soil properties of the experimental region.

Determination	Field-1	Field-2	Field-3	Field-4
Physical analysis	Value			
Sand (%)	42.8	43.4	42.9	45.1
Silt (%)	24.9	26.7	35.1	34.5
Clay (%)	32.2	29.8	21.9	14.2
Textural classes (USDA)	Clay loam	Clay loam	Loam	Sandy Loam
Chemical analysis	Value			
pH	4.68	5.49	6.48	5.74
Organic matter (%)	0.89	0.98	0.94	1.03
Available nitrogen (Kg ha ⁻¹)	193.56	197.94	195.75	207.20
Available phosphorus (Kg ha ⁻¹)	17.08	17.56	16.05	16.85
Available potassium (Kg ha ⁻¹)	124.54	128.36	121.87	120.89

Table 3. List of genotypes used in experiment.

Code	Genotype	Code	Genotype	Code	Genotype
G1	ERP 26	G11	ELS 36	G21	FOX 4384
G2	ERP 57	G12	ELS 40	G22	FOX 4386
G3	ERP 82	G13	FOX 4438	G23	FOX 4390
G4	ERP 90	G14	FOX 4440	G24	FOX 4392
G5	ESD 3	G15	FOX 4478	G25	FOX 4394
G6	ESD 46	G16	FOX 4489	G26	FOX 4396
G7	ESD 75	G17	FOX 4330	G27	FOX 4403
G8	ESD 79	G18	FOX 4336	G28	FOX 4428
G9	ELS 20	G19	FOX 4339	G29	FOX 4420
					SURYANANDI
G10	ELS 34	G20	FOX 4341	G30	(Check)

Experimental design and intercultural practice

The experiment was conducted using a method called randomized complete block design (RCBD) with three replications at all environments. RCBD was chosen because the fertility of the experimental sites varied. Each replication consisted of 30 plots or beds, each measuring 1 meter by 1 meter. There was a gap of 10 centimeters between each bed, and the spacing between individual plants within a bed was 10 centimeters. The distance between rows was 22.5 centimeters. The replications were separated from each other by a distance of 30 centimeters. The total size of the experimental plot was 30 meters by 5 meters, with 30 centimeters of space before the first replication and 30 centimeters of space after the third replication. In total, there were 90 beds across all environments. Throughout the experiment, recommended agricultural practices were followed.

Data collection

To collect data, a total of 14 quantitative characteristics of foxtail millet were considered. These characteristics were chosen based on descriptions and guidelines provided by PPV&FR in 2001 (DUS). For each characteristic, data were gathered from five randomly selected plants within each genotype and replication. The quantitative data collection included recording information on the following traits: 1) days to 50% flowering (DF), 2) days to maturity (DM), 3) plant height (PH), 4) panicle

length (PL), 5) flag leaf length (FL), 6) flag leaf width (FW), 7) peduncle length (PDL), 8) total tiller numbers per plant (NT), 9) panicle width (PW), 10) biological yield (BY), 11) harvest index (HI), 12) Test weight (TW), 13) fodder yield per plant (FY) and 14) grain yield/plant (GY). The measurements for biological yield, harvest index, test weight, fodder yield per plant, and grain yield per plant were verified in the laboratory after the harvest, while the remaining traits were measured directly in the field.

Statistical analysis

The analysis of K-means and PCA was conducted using the OPSTAT open-source software. All diagrams of K-means and PCA generated by R-Studio, using “tidyverse”, “cluster”, “factoMineR” and “factoextra” for K-means and PCA. The clustering analysis using Ward’s minimum variance technique was carried out using the ‘Metan’ package in R-studio (Olivoto *et al.*, 2019).

RESULTS AND DISCUSSION

K-Means

The optimal K value for the population was four, according to the Elbow Method. Cluster -1 (16.6%) have 5 genotypes (G1, G5, G18, G22, and G25), cluster-2 (20.0%) have 6 genotypes (G2, G17, G24, G28, G29, and G30), cluster-3 (43.3%) have 13 genotypes (G3, G4, G7, G8, G9, G11, G12, G19, G20, G21, G23, G26, G27) and cluster-4 (20.0%)

had 6 genotypes (G6, G10, G13, G14, G15, and G16). Cluster-1 was better than other clusters, according to the 5% confidence interval, for yield and yield-related traits of grain yield, fodder yield, panicle length, flag leaf area and tillers per plant. Classification of cluster based on k means represented in Table 4 and Fig. 1.

Cluster mean represents the average values of all the variables for the data points belonging to a particular cluster. For the trait “DF” (Days to Flowering), Cluster 1 exhibits the highest value at 74.104, while Cluster 2 shows the lowest at 69.905. When considering “DM” (Days to Maturity), Cluster 1 also displays the maximum value of 114.826, and Cluster 2 records the minimum at 108.14. In terms of “PH” (Plant Height), the highest value of 125.332 is observed in Cluster 2, with the lowest value of 92.837 in Cluster 4. In terms of “PL” (Panicle Length), Cluster 1 demonstrates the highest value at 15.5, while Cluster 4 exhibits the lowest value at 13.128. For “FL” (Flag Leaf Length), Cluster 1 again presents the highest value at 23.042, and Cluster 4 shows the lowest at 20.535. “FW” (Flag Leaf Width) exhibits the maximum value in Cluster 2 at 1.92 and the minimum in Cluster 1 at 1.788. Regarding “PDL” (Peduncle length), Cluster 1 records the highest value at 22.654 and Cluster 3 displays the lowest at 19.234. Lastly, the trait “NT” (Number of Tillers) reveals Cluster 1 with the highest value at 3.772 and Cluster 3 with the lowest at 3.545. When examining “PW” (Panicle Weight), Cluster 1 demonstrates the highest value at 1.954, while Cluster 4 exhibits the lowest value at 1.688. In terms of “BY” (Biological Yield), Cluster 1 shows the maximum value at 40.296, whereas Cluster 2 records the lowest at 28.553. “HI” (Harvest Index) presents the highest value in Cluster 1 at 45.823 and the lowest in Cluster 2 at 15.45. For “FY” (Fodder Yield), Cluster 1 exhibits the highest value at 22.276, while Cluster 2 displays the lowest at 15.45. “TW” (Test Weight) exhibits the maximum value in Cluster 1 at 2.805, with the lowest in Cluster 4 at 2.768. Finally, “GY” (Grain Yield) reveals Cluster 1 with the highest value at 18.036 and Cluster 2 with the lowest at 13.112. Cluster means is presented in Table 5.

The analysis of various foxtail millet genotypes

across different clusters has provided valuable insights into the diversity of traits and their potential implications for future breeding programs. When assessing the level of divergence among clusters, it becomes evident that Cluster 1 stands out as remarkably divergent in several key traits. In the context of high-yield breeding, Cluster 1 stands out as it demonstrates higher values in various yield-related traits, including “BY” (Biological Yield), “HI” (Harvest Index), “FY” (Fodder Yield), “TW” (Test Weight), and “GY” (Grain Yield). This suggests that Cluster 1 is a promising source for developing foxtail millet cultivars focusing on maximizing yield potential. Cluster 2 showcases the lowest values in “DF” (Days to Flowering), “DM” (Days to Maturity), and “PH” (Plant Height), indicating a shorter growth cycle and reduced plant height compared to other clusters. Similar results reported by Borrayo *et al.* (2016) in foxtail millet; Nandini *et al.* (2020) in barnyard millet and Nandini *et al.* (2018) in foxtail millet.

Future breeding efforts could involve crossing genotypes from Cluster 1 with other high-yielding cultivars or those from Cluster 2 to create a balanced set of traits encompassing high yield and desirable characteristics like early maturation and compact growth. By selecting genotypes that offer a combination of these attributes, plant breeders can work towards developing foxtail millet varieties that not only exhibit high yields but also are also well suited for specific environmental conditions and cropping systems. Ultimately, the goal is to strike a balance between yield and other desirable traits to create cultivars that meet the diverse needs of agriculture and food security.

Inter-cluster distances were observed to vary among foxtail millet genotypes, providing insights into their genetic relationships. The inter-cluster distances ranged from 16.33 to 33.09. Clusters-2 and cluster-4 showed the maximum inter-cluster distance (33.09), followed by cluster-1 and cluster-4 (31.85), cluster-3 and cluster-4 (17.37), cluster-2 and cluster-3 (17.08), cluster-1 and cluster-2 (17.00), and cluster-1 and cluster-3 (16.33). Inter-cluster distances are presented in Table 6.

Variety development aims to create new cultivars with specific traits. Selection of these clusters has heights inter cluster distance respective environmental, suggesting a considerable genetic variation within this cluster (Chaudhary *et al.* 2015). Selecting genotypes from this cluster for breeding programs can help in developing diverse and distinct varieties with a broader range of traits and Utilizing parents from this cluster in hybridization can lead to increased heterozygosity and potentially improved performance in the hybrid offspring (Jain & Diwan 2021).

The inter-cluster distances provide insights into the genetic relationships and relatedness among clusters. Larger inter-cluster distances suggest greater dissimilarity and differentiation between clusters, indicating distinct and genetically diverse groups (Singh *et al.* 2017). On the other hand, smaller inter-cluster distances suggest closer genetic relationships and similarities between clusters, possibly sharing common traits or ancestry (Everitt *et al.* 2011). In this study, Clusters-2 and cluster-4 showed

the maximum inter-cluster distance (33.09), followed by cluster-1 and cluster-4 (31.85).

Principal component analysis

In analyzing 30 foxtail millet genotypes' yield data using principal component analysis, we found that five of the 14 components had values greater than 1 (Table 7 and Fig. 2). These five key components accounted for around 75.8% of the overall variability in the data. The remaining nine components, however, contributed to only 24.2% of the variation observed among the foxtail millet genotypes. The results allowed us to identify specific productivity elements for each principal component, giving insights into the variation in population characteristics. PC1 explains 25.9%, PC2 explains 44.2%, PC3 explains 58.0%, PC4 explains 68.6%, and PC5 explains 75.8% of the total variance among different yield traits.

In this study, the average data was analyzed using principal component analysis, which shows the most significant factor influencing the overall variation on each axis. The

Table 4. Total Clusters and the accessions falling in each one of them.

Cluster No	Accessions	No. of genotypes
1	G1, G5, G18, G22, G25,	5
2	G2, G17, G24, G28, G29, G30,	6
	G3, G4, G7, G8, G9, G11, G12, G19, G20, G21,	
3	G23, G26, G27,	13
4	G6, G10, G13, G14, G15, G16,	6

Table 5. Cluster means.

	DF	DM	PH	PL	FL	FW	PDL	NT	PW	BY	HI	FY	TW	GY
Cluster 1	74.10	114.83	121.53	15.50	23.04	1.79	22.65	3.77	1.95	40.30	44.61	22.28	2.77	18.04
Cluster 2	69.91	108.14	125.33	14.90	21.80	1.92	22.21	3.59	1.87	28.55	45.82	15.45	2.78	13.11
Cluster 3	72.09	110.58	109.81	13.14	21.00	1.89	19.23	3.55	1.77	32.73	44.97	18.12	2.81	14.67
Cluster 4	71.54	112.14	92.84	13.13	20.54	1.91	19.30	3.73	1.69	30.16	44.11	16.83	2.77	13.34

Table 6. Inter Cluster Distances.

	Cluster 1	Cluster 2	Cluster 3	Cluster 4
Cluster 1	0	17.005	16.337	31.859
Cluster 2		0	17.088	33.096
Cluster 3			0	17.376
Cluster 4				0

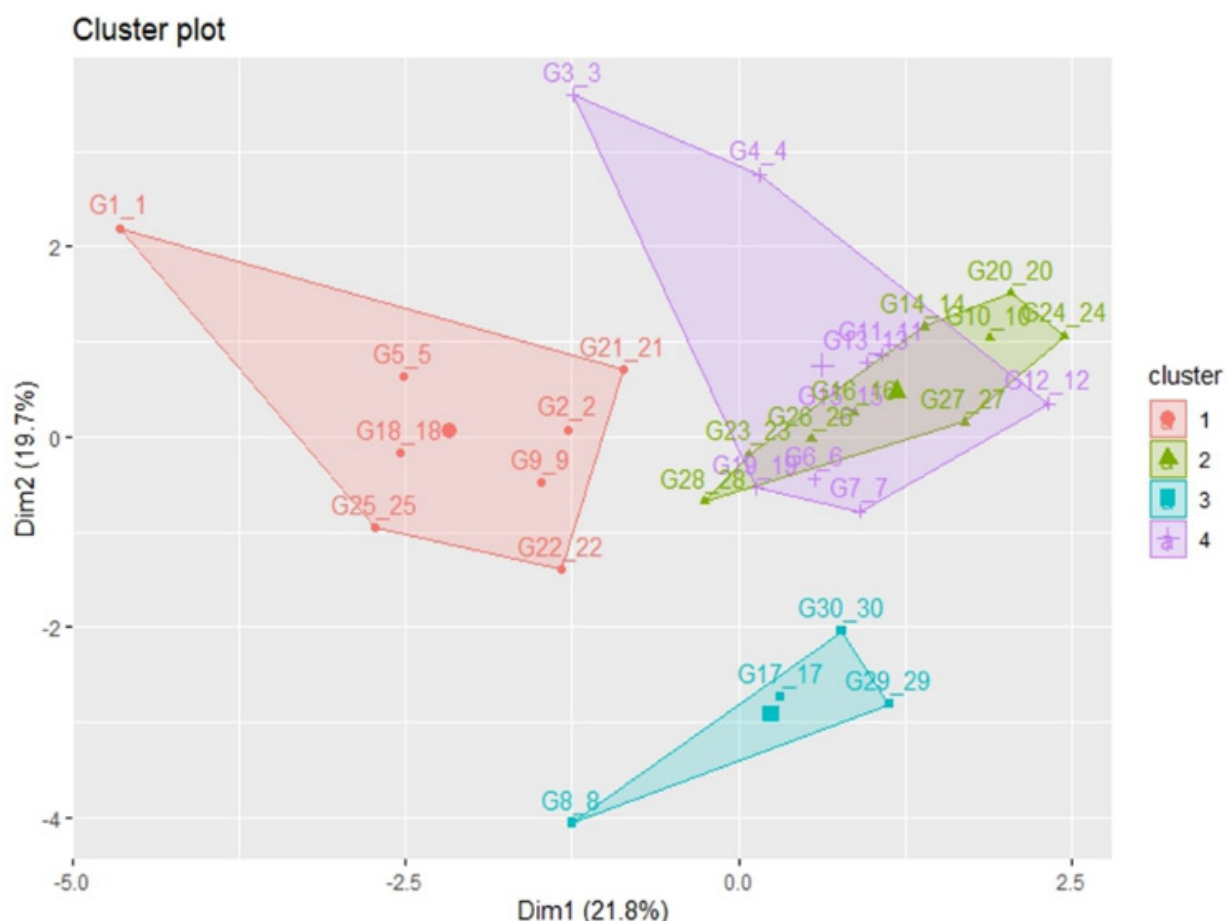


Figure 1. K-means cluster analysis for yield and yield traits evaluated in 30 foxtail millet genotypes during 2022-2023.

eigenvalues obtained from this analysis are commonly utilized to decide on the number of factors to keep. Typically, the sum of these eigenvalues equals the total number of variables (Hailegiorgis *et al.* 2011). In the analysis of 30 foxtail millet genotypes, 5 out of 14 principal components had values exceeding eigenvalues more than 1, representing around 75.8% of the data variability. The remaining nine components contributed to only 24.2% of the variation observed.

The loadings of principal components for a set of yield variables represent how each variable contributes to the variance captured by each principal component. In this study (Table 8), PC1, variables like BY (0.495), FY (0.466) and GY (0.483) have high positive loadings, indicating that they strongly influence the first principal component. In PC2, variables like DF (0.497), DM (0.457) and TW (0.27) have high positive loadings, indicating a strong positive relationship with this principal component. Conversely, variables such as PH (-0.138),

PL (-0.31), and PW (-0.351) have negative loadings, suggesting an inverse relationship with PC2. In PC3, variables like PH and HI have high absolute loadings, with PH (-0.425) having a negative loading and HI (-0.33) having a strongly negative loading. This indicates a significant inverse relationship between these variables and PC3. Conversely, variables like FY (0.301) have a positive loading, suggesting a positive association with PC3. In PC4, variables like FL (0.479), PDL (0.361), TW (0.494) and NT (0.374) have relatively high positive loadings, indicating a strong positive association with this principal component. Conversely, variables such as PW (-0.4) have negative loadings, suggesting an inverse relationship with PC4. In PC5, variables like PH (0.335) and TW (0.496) have relatively high positive loadings, suggesting a strong positive relationship with this principal component. On the other hand, variables such as DM (-0.407), FL (-0.363), NT (-0.254) and FW (-0.419) have negative loadings, indicating an inverse association with PC5.

The positive and negative effects of factors show how components relate to different varieties (Adilova *et al.* 2020). Hence, the positive and negative productivity factors influenced the clustering process. In the principal component analysis, biological yield, days to flowering, and fodder yield per plant were chosen for the first, second, and third groups, respectively. At the same time, test weight was selected for the fourth and fifth groups. These three major components had a more significant impact on genotype clustering compared to the other components.

The t-test revealed nine significant positive correlations and three significant negative correlations among the 14 pairs of yield traits assessed, with p-values of < 0.05 , < 0.01 , and < 0.001 . High phenotypic correlations ($r > 0.66$) were measured between BY and GY ($r = 0.94$), FY and GY ($r = 0.82$), FY and BY ($r = 0.97$), DF and DM ($r = 0.82$). Medium correlations ($0.33 < r < 0.67$) were measured between PH and PL ($r = 0.40$), GY and PL ($r = 0.45$), FL and PDL ($r = 0.46$), PL and PDL ($r = 0.42$) and PDL and PH ($r = 0.39$). The correlation is presented in Fig. 3.

Identification of similarly responded genotypes

To examine how genotypes interacted with the environment, we used IPCA I and IPCA II scores within the GGE framework. To simplify our findings, we performed cluster analysis on these scores for the early *kharif*- late *kharif* of 2022 and the early summer- late summer of 2023, focusing on grain yield per plant. This helped group genotypes with similar responses, making it easier to understand than interpret the complex biplot. We established an 80% similarity threshold for clustering genotypes using Ward's minimum variance technique. The results, presented in Table 9 and Fig. 4, showed that we categorized 30 genotypes into five distinct responsive clusters.

Cluster-1 had six genotypes with a 45.70% similarity in their responses. In Cluster 2, seven genotypes exhibited a 45.02% similarity in their response to environmental conditions for grain yield per plant. Similarly, Cluster-3

contained five genotypes with an 18.52% resemblance in this aspect. Cluster-4 included eight genotypes with a negative similarity index (-1.208), indicating a highly diverse response pattern confirmed by the graphical representation. Cluster-5 has four genotypes (G22, G5, G1, G25) similarity index (-122.50) in a high negative direction, indicating its highly differential response nature. This highlights the usefulness of calculating similarity index values to identify genotypes that respond similarly to environmental factors for specific traits.

Cluster 5 consisted of four genotypes (G22, G5, G1, G25) with a very high negative similarity index (-122.50), which makes this cluster highly distinctive compared to the others. Such extreme divergence indicates that these genotypes exhibit unique and unstable responses to environmental variation. In multi-environment trials, clusters with negative similarity often capture genotypes that are specifically adapted, rather than broadly adapted, to particular environments (Olivoto *et al.* 2019). This makes them valuable for targeted breeding programs, especially where local adaptation is more important than general stability.

Interestingly, three of these genotypes (G1, G5, G22) were also highlighted as superior for grain yield and yield-related traits in the K-means clustering analysis, which focused on mean performance. Their reallocation into Cluster 5 under the GGE/IPCA-based Ward's clustering underscores that while they possess high yield potential, their stability and adaptability vary substantially. This contrast between high productivity and high environmental sensitivity makes them ideal candidates for targeted breeding programs aimed at developing varieties for specific agro-ecological niches rather than broad adaptation (Nandini *et al.* 2018, Rao & Chaturvedi 2022).

G18 may have a yield profile similar to G1, G5, G22, and G25 (hence grouped with them in K-means), but its interaction with the environment differs significantly, which caused it to separate under Ward's clustering. Traits like days to flowering or harvest index, which load differently in PCA and IPCA analyses, may have influenced G18's placement. If G18

Table 7. Eigen values of Correlation Matrix.

	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8	PC9	PC10	PC11	PC12	PC13	PC14
Eigenvalues	3.621	2.563	1.929	1.493	1	0.947	0.725	0.537	0.439	0.351	0.277	0.099	0.019	0
Proportion	0.259	0.183	0.138	0.107	0.071	0.068	0.052	0.038	0.031	0.025	0.02	0.007	0.001	0
Cumulative Proportion	0.259	0.442	0.58	0.686	0.758	0.825	0.877	0.915	0.947	0.972	0.992	0.999	1	1

Table 8. Loadings (Eigenvectors) of Correlation Matrix.

	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8	PC9	PC10	PC11	PC12	PC13	PC14
DF	0.178	0.497	-0.083	-0.126	-0.15	-0.216	-0.233	-0.185	0.305	0.03	-0.19	0.638	0.061	0
DM	0.173	0.457	-0.125	-0.146	-0.407	-0.153	-0.194	0.03	0.001	0.283	0.222	-0.598	-0.096	-0.003
PH	0.215	-0.138	-0.425	-0.144	0.335	-0.17	-0.167	-0.429	-0.482	0.237	-0.295	-0.07	0.004	0
PL	0.244	-0.31	-0.33	-0.128	-0.11	0.214	0.025	0.321	0.496	0.169	-0.511	-0.139	-0.054	0.002
FL	0.18	-0.04	-0.289	0.479	-0.363	-0.309	0.124	-0.058	-0.104	-0.596	-0.175	-0.097	-0.012	0.003
FW	-0.079	-0.371	0.171	0.08	-0.419	-0.415	0.373	-0.213	0.033	0.515	0.019	0.148	0.007	0.001
PDL	0.246	-0.223	-0.228	0.361	0.163	-0.227	-0.377	0.435	0.006	0.219	0.449	0.214	0.044	-0.004
NT	0.069	-0.204	0.223	0.374	-0.254	0.451	-0.519	-0.456	0.089	0.082	-0.022	-0.035	-0.041	0.002
PW	0.122	-0.351	0.111	-0.4	0.148	-0.366	-0.196	-0.293	0.423	-0.351	0.271	-0.169	-0.076	-0.001
BY	0.495	0.018	0.179	-0.035	0.021	0.118	0.197	-0.013	-0.077	-0.012	0.081	0.042	-0.021	0.808
HI	-0.134	-0.03	-0.582	-0.091	-0.13	0.341	0.253	-0.236	0.085	-0.022	0.445	0.203	-0.364	0.039
FY	0.466	0.027	0.301	-0.006	0.057	0.028	0.127	0.075	-0.146	-0.01	-0.038	0.102	-0.675	-0.421
TW	0.012	0.27	-0.021	0.494	0.496	-0.094	0.273	-0.261	0.442	0.186	-0.008	-0.226	-0.066	-0.001
GY	0.483	-0.001	-0.011	-0.071	-0.034	0.242	0.286	-0.128	0.006	-0.021	0.23	-0.001	0.618	-0.41

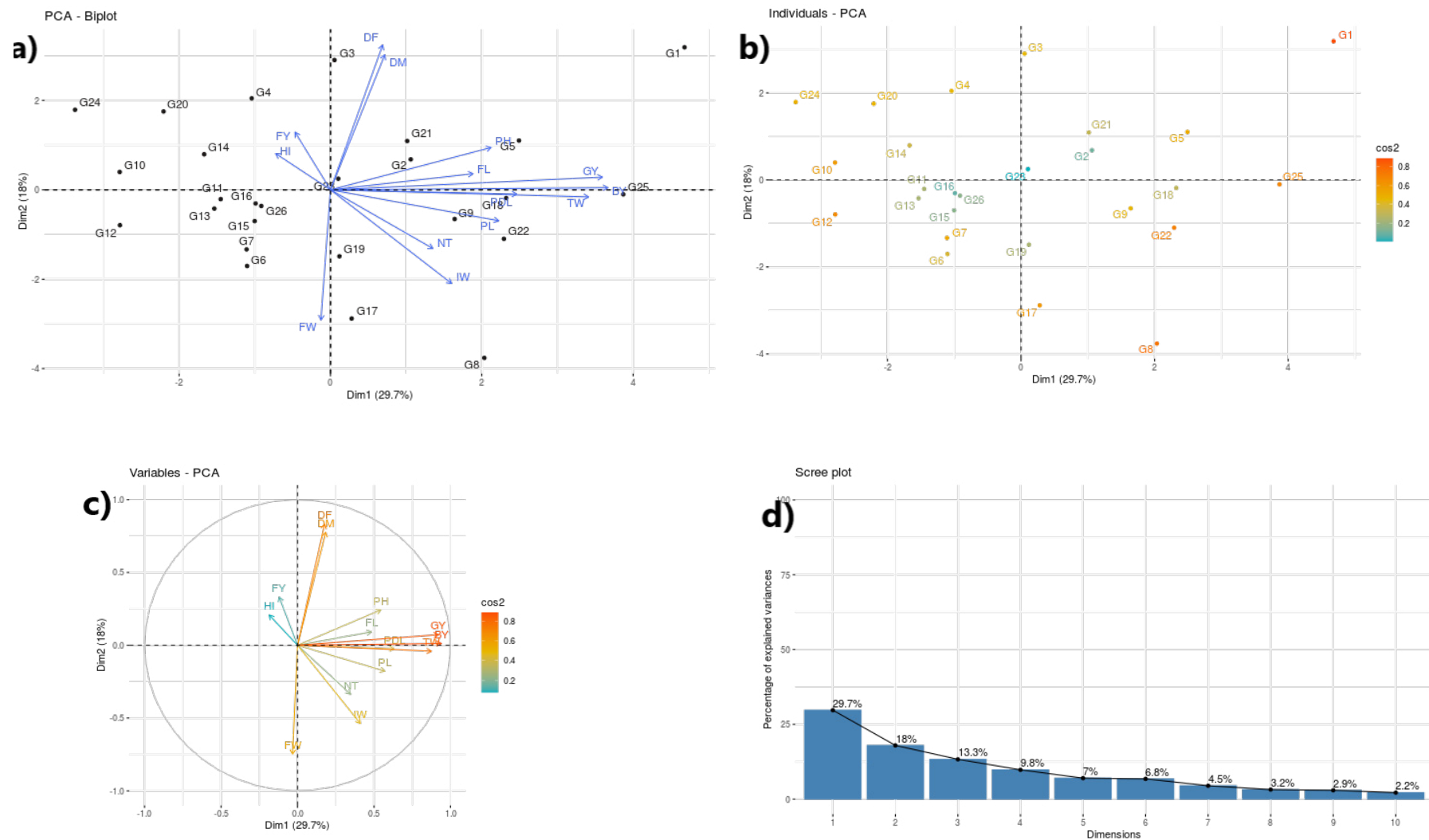


Figure 2. A, B). Distribution and grouping of germplasm lines across first two principal components, C). PCA graph of different traits for first two principal components, D). screen plot and threshold Eigen values for quantitative traits in foxtail millet.

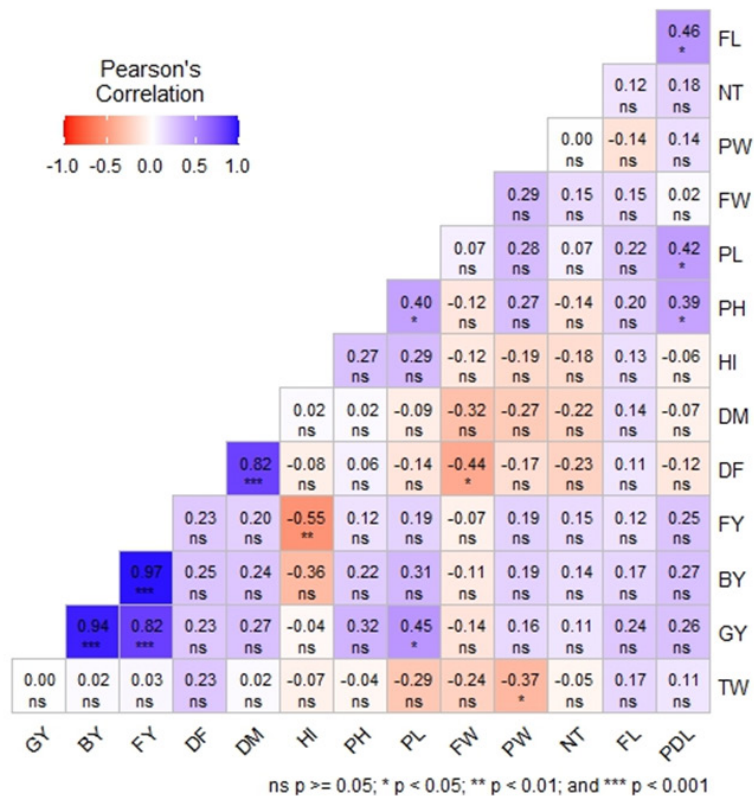


Figure 3. Pearson's phenotypic correlations among yield and yield traits evaluated in 30 foxtail millet genotypes.

responded differently in terms of flowering time or biomass allocation under stress, this would shift its similarity index away from G1, G5, and G22. This indicates G18 could be a niche genotype, with value for targeted breeding programs focused on specific environments (e.g., drought-prone or short-season areas).

The uniqueness of Cluster 5 can be further explained by its high inter-cluster distance (16.75) compared to the other groups, suggesting that these genotypes harbor rare allelic combinations and unique response traits (Chaudhary *et al.* 2015). In foxtail millet and related crops, such divergent genotypes are frequently recommended as donor parents for crossing, since greater genetic distance increases the probability of heterosis (Borrayo *et al.* 2016 and Jain & Diwan 2021). Thus, while Cluster 5 genotypes are less stable across environments, they may be strategically valuable for exploiting hybrid vigor or introducing resilience traits into breeding pools.

Moreover, genotypes that fall into divergent clusters often show environment-specific tolerance to stresses such as drought, low

soil fertility, or heat (Nandini *et al.* 2020). It is possible that G22, G5, G1, and G25 carry adaptive mechanisms not visible in average yield data but expressed under particular stress-prone environments. This interpretation aligns with findings in multi-environment trials of small millets and cereals, where unstable but divergent genotypes provided novel adaptation strategies (Everitt *et al.* 2011).

Table 9. Clusters Similarity Distance.

Clusters	No. of genotypes	Similarity	Distance
1	6	45.701	4.09
2	7	45.028	4.14
3	5	18.523	6.13
4	8	-1.208	7.62
5	4	-122.505	16.75

CONCLUSION

The present study employed multivariate approaches, including GGE biplot-based IPCA scores and Ward's clustering, to evaluate foxtail millet genotypes across environments. Thirty genotypes were grouped into five

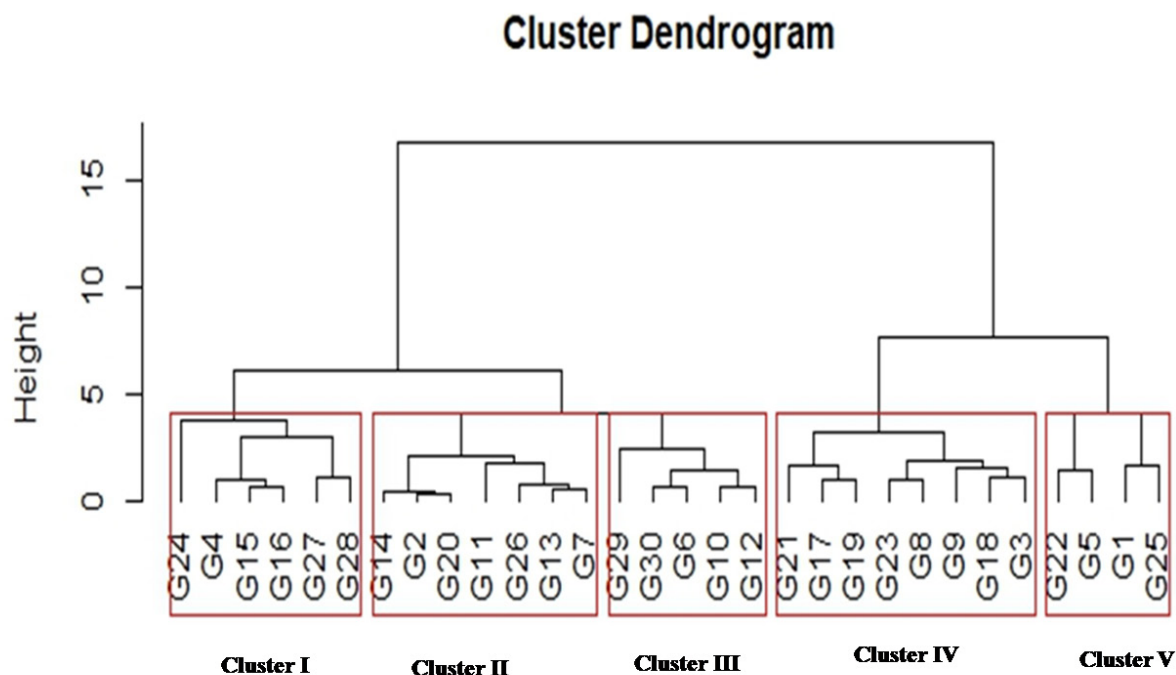


Figure 4. Clustering the similarly responding genotypes into different similarity groups based on Ward's minimum variance technique in 30 foxtail millet genotypes.

clusters, revealing both stable and divergent response patterns. Clusters 1 and 2 showed broader adaptability, whereas Cluster 5 (G22, G5, G1, G25) exhibited a strongly negative similarity index, reflecting high environmental sensitivity. Despite their instability, these

genotypes represent valuable sources of unique traits and genetic diversity. Integrating both performance and stability analyses is essential for breeding strategies aimed at balancing yield stability and genetic diversification to develop resilient foxtail millet cultivars.

REFERENCES

- Adilova SS, Qulmamatova DE, Baboev SK, Bozorov TA, Morgunov AI (2020) Multivariate cluster and principal component analyses of selected yield traits in Uzbek bread wheat cultivars. *American Journal of Plant Sciences*, 11(06), 903. <https://doi.org/10.4236/ajps.2020.116066>.
- Ali A, Choi YM, Do YH, Lee S, Oh S, Park HJ, Lee MC (2016) EST-SSR based genetic diversity and population structure among Korean landraces of foxtail millet (*Setaria italica* L.). *Korean Journal of Plant Resources*, 29(3), 322–330. <https://doi.org/10.7732/kjpr.2016.29.3.322>.
- Barth E, de Resende JTV, Marigulele KH, de Resende MDV, da Silva ALBR, Ru S (2022) Multivariate analysis methods improve the selection of strawberry genotypes with low cold requirement. *Scientific Reports*, 12(1), 11458. <https://doi.org/10.1038/s41598-022-15688-4>.
- Borrayo E, Machida-Hirano R, Takeya M, Kawase M, Watanabe K (2016) Principal components analysis-K-means transposon element-based foxtail millet core collection selection method. *BMC Genetics*, 17(1), 1–10. <https://doi.org/10.1186/s12863-016-0343-z>.
- Chaudhary S, Sagar P, Hooda BK, Arya RK (2015) Multivariate analysis of pearl millet data to delineate genetic variation. *Forage Research*, 40(4), 201–208. <https://doi.org/10.59512/aaz.2023.62.2.10>.

- Everitt BS, Landau S, Leese M, Stahl D (2011) Cluster Analysis. *Wiley Series in Probability and Statistics*. <https://doi.org/10.1002/9780470977811>.
- Hailegiorgis D, Mesfin M, Genet T (2011) Genetic divergence analysis on some bread wheat genotypes grown in Ethiopia. *Journal of Central European Agriculture*, 12(2), 0–0. <https://doi.org/10.5513/jcea01/12.2.922>.
- Hanna WW, Baltensperger DD, Seetharam A (2004) Pearl millet and other millets. Warm-season (C4) grasses, 45, 537–560. <https://doi.org/10.5513/jcea01/12.2.922>.
- Jain SK, Diwan D (2021) Principal component and cluster analysis for quantitative traits to identify high yielding genotypes of pearl millet [*Pennisetum glaucum* (L.) R. Br.]. *Forage Research*, 46(4), 308–314.
- Jarwar AH, Wang X, Iqbal MS, Sarfraz Z, Wang L, Ma Q, Shuli F (2019) Genetic divergence on the basis of principal component, correlation and cluster analysis of yield and quality traits in cotton cultivars. *Pakistan Journal of Botany*, 51(3), 1143–1148.
- Moharil MP, Ingle KP, Jadhav PV, Gawai DC, Khelurkar VC, Suprasanna P (2019) Foxtail millet (*Setaria italica* L.): potential of smaller millet for future breeding. *Advances in Plant Breeding Strategies: Cereals*, Volume 5, 133–163. [https://doi.org/10.30848/pjb2019-3\(38\)](https://doi.org/10.30848/pjb2019-3(38)).
- Mokkaraj J, Geethanjali S (2016) Genetic diversity and variability in foxtail millet (*Setaria italica* L.) germplasm based on morphological traits. *Electronic Journal of Plant Breeding*, 7(2), 303–316. <https://doi.org/10.5958/0975-928x.2016.00038.7>.
- Nandini C, Sujata B, Tippeswamy V (2018) Characterization of foxtail millet (*Setaria italica* (L.) Beauv.) germplasm for qualitative and quantitative traits to enhance its utilization. *Academia Journal of Agricultural Research*, 6(5), 121–129.
- Nandini C, Bhat S, Saritha HS, Pandey CD, Pandey SP, Bai L, Gowda J (2020) Characterization of barnyard millet (*Echinochloa frumentacea* (Roxb.) Link) germplasm for quantitative traits to enhance its utilization. *Electronic Journal of Plant Breeding*, 11(04), 1066–1072. <https://doi.org/10.37992/2020.1104.173>.
- Olivoto T, Lúcio AD, da Silva JA, Marchioro VS, de Souza VQ, Jost E (2019) Mean performance and stability in multi-environment trials I: combining features of AMMI and BLUP techniques. *Agronomy Journal*, 111(6), 2949–2960. <https://doi.org/10.2134/agronj2019.03.0220>.
- Pallavi NL, Venkatesh R, Ram BJ, Suresh BG (2020) Studies on correlation and path coefficient analysis in foxtail millet [*Setaria italica* (L.) Beauv]. *International Journal of Chemical Studies*, 8(6), 1941–1946. <https://doi.org/10.22271/chemi.2020.v8.i6ab.11048>.
- Rao DP, Chaturvedi HP (2022) Genetic diversity analysis in foxtail millet (*Setaria italica* (L.) P. Beauv.) genotypes. *Journal of Plant Science Research*, 38(2), 907–918. <https://doi.org/10.32381/jpsr.2022.38.02.44>.
- Sharma JR (2006) Statistical and biometrical techniques in plant breeding. New Age International, New Delhi, India, 432 p.
- Singh S, Singh SK, Korada M, Khaire A, Singh DK (2017) Morpho-molecular diversity analysis in rice (*Oryza sativa* L.) genotypes using microsatellite markers. *Indian Journal of Agricultural Research*, 1, 8. <https://doi.org/10.18805/ijare.a-5885>.
- Syakur MA, Khotimah BK, Rochman EMS, Satoto BD (2018) Integration k-means clustering method and elbow method for identification of the best customer profile cluster. *IOP Conference*

Series: *Materials Science and Engineering*, 336, 012017.<https://doi.org/10.1088/1757-899x/336/1/012017>.

Taylor JR (2019) Sorghum and millets: Taxonomy, history, distribution, and production. In Sorghum and millets (pp. 1–21). *AACC International Press*. <https://doi.org/10.1016/b978-0-12-811527-5.00001-0>.