

# Evaluation of Genetic Diversity in Sorghum Genotypes for Drought Tolerance Using Mahalanobis' $D^2$ Analysis

## ABSTRACT

Genetic diversity is vital in plant breeding and crop improvement, enabling the development of varieties resilient to biotic and abiotic stresses. Mahalanobis'  $D^2$  analysis is a robust multivariate statistical method for assessing genetic diversity, offering precise measures of genetic divergence. This study aimed to evaluate the genetic diversity of twenty sorghum (*Sorghum bicolor* L. Moench) genotypes under rainfed and irrigated conditions using Mahalanobis'  $D^2$  statistics. The experiment was conducted at the Main Agricultural Research Station (MARS), Dharwad, involving twenty genotypes sourced from the AICRP sorghum, Dharwad, India. The genotypes were sown in medium black soil under two moisture levels: rainfed and irrigated. Morphological, physiological, phenological, and yield-related traits were recorded to identify drought-tolerant genotypes. Genetic divergence was estimated using  $D^2$  statistics, and genotypes were grouped into clusters using Tocher's method. Under rainfed conditions, the genotypes were classified into six clusters, while under irrigated conditions, they were grouped into three clusters. The intra-cluster distances ranged from 7.79 to 12.14 under rainfed conditions and from 8.14 to 15.74 under irrigated conditions. The inter-cluster distances revealed significant genetic variation, with the highest divergence observed between Clusters III and V under rainfed conditions and Clusters II and III under irrigated conditions. These findings highlight the presence of substantial genetic diversity among sorghum genotypes, facilitating the selection of diverse parental lines for breeding programs aimed at enhancing drought tolerance. The insights gained from this study contribute to the development of resilient sorghum varieties capable of withstanding water scarcity, ensuring sustainable agriculture, and promoting food security in drought-prone regions.

**Keywords:** Sorghum, Genetic Diversity, Mahalanobis'  $D^2$  Analysis, Drought Tolerance, and Cluster Analysis

## 1. INTRODUCTION

Genetic diversity is fundamental in plant breeding and crop improvement, enabling the development of varieties capable of withstanding various biotic and abiotic stresses. Among the methodologies available for assessing genetic diversity, Mahalanobis'  $D^2$  analysis stands out as a robust multivariate statistical approach [8]. This method provides a precise measure of genetic divergence, aiding in the effective identification and utilization of diverse genetic resources. Sorghum (*Sorghum bicolor* L. Moench) is a resilient cereal crop of paramount importance, especially in drought-prone regions. Renowned for its exceptional drought tolerance, sorghum is a staple food for millions across Africa and Asia. Beyond its role as a food source, sorghum is also used for fodder, biofuel production, and various industrial applications. Enhancing the drought tolerance of sorghum is crucial for securing food production and sustainability in the face of climate change and water scarcity.

Among the various methods developed to study genetic divergence in genotypes, Mahalanobis'  $D^2$  [9] is both reliable and frequently used.  $D^2$  analysis is a useful tool for quantifying the degree of divergence between biological populations at the genotypic level and for assessing the relative contribution of different components to total divergence at both inter- and intra-cluster levels. In this study, genetic divergence was estimated using  $D^2$  statistics as suggested by Mahalanobis (1936), which is based on the multivariate analysis of quantitative traits. This method is highly effective for measuring genetic divergence within a population by employing the concept of statistical distance through multivariate measurements.

It has been noted that plant breeders often use a much less diverse genetic pool than the overall available genetic diversity within a crop [6]. Heterogeneous local populations of the genus form an important source of genetic variation (Zeven, 1998). For selecting parents in hybridization, diversity among parents for the character of interest and estimation of genetic distance is crucial, as diverse plants are expected to exhibit high hybrid vigor [3]. This study focuses on evaluating the genetic diversity among sorghum genotypes with an emphasis on drought tolerance using Mahalanobis'  $D^2$  analysis [5]. By examining the genetic divergence among these genotypes, we aim to identify promising candidates for breeding programs aimed at improving drought tolerance. Understanding the genetic basis of drought tolerance in sorghum will contribute to developing resilient varieties, ensuring food security, and promoting agricultural sustainability in drought-prone areas.

## 2. MATERIALS AND METHODS

### 2.1 Experimental Site

The field experiment took place at the Main Agricultural Research Station (MARS) of the University of Agricultural Sciences in Dharwad. Specifically, it was conducted in plot No. 126 of E-block. MARS is located at a latitude of 15°12' N and a longitude of 76°34' E, sitting 678 meters above sea level.

### 2.2 Experimental Setup

An experiment was conducted on medium black soil classified as verticceptisols, with a depth of 2-3 meters. Twenty rabi sorghum genotypes (listed in Table 1) were sourced from the AICRP sorghum, Dharwad, India. The first season's sowing took place on October 24, 2021, followed by the second season's sowing on November 11, 2022. High-quality seeds were used for row sowing. Each plot consisted of 6 rows with a 45 cm gap between them, and individual plants within a row were spaced 15 centimeters apart. The experiment involved two moisture levels: rainfed and irrigated. In the rainfed condition, no irrigation was provided after sowing. In the irrigated condition, two additional irrigations were conducted: the first 35 days after sowing and the second 65 days after sowing.

**Table 1: List of sorghum genotypes**

Names of sorghum genotypes							
1	SVD-1272R	6	SPV-2217	11	Tandur L	16	M 148-138
2	SVD-1358R	7	CSV-216R	12	Phule Anuradha	17	Basavanmoti
3	SVD-1528R	8	CSV-29R	13	Chitapur – L	18	Phule Vasudha
4	SVD-1403R	9	ICSR-15001	14	DKS- 35	19	BJV-44
5	SPV-486	10	Basavana pada	15	M-35-1	20	ICSR- 13025

### 2.3 Parameters Determined

The observations recorded for 20 sorghum genotypes aimed to identify water stress-tolerant sorghum genotypes by considering morphological, phenological, and yield-related traits. Morphological traits included plant height, leaf area, and leaf area index (LAI) at 30, 60, and 90 days after sowing (DAS). Plant height was measured from the stem base to the tip. Leaf area was determined using the Stickler et al. method, and LAI was computed as the ratio of leaf area to land area per plant.

Physiological traits, including SPAD (chlorophyll meter reading), relative water content, chlorophyll content, and membrane stability index, were observed during different crop growth stages. Phenological traits such as days to 50% flowering and days to physiological maturity were also recorded. Days to 50% flowering represented the time when half of the plants in each treatment bloomed, expressed in days. Days to physiological maturity indicated the duration for a genotype to reach a stage where seeds developed a dark spot, marking the end of the photosynthate supply. Yield parameters were assessed through panicle weight per plant, grain yield per plant, stover yield, harvest index, and grain yield per hectare. Multivariate analysis was conducted using the Mahalanobis D<sup>2</sup> statistic [9], and genotypes were grouped into different clusters following Tocher's method. The inter- and intra-cluster distances were calculated as per the method suggested by Murty and Arunachalam [10] to determine the actual divergence within and between the clusters.

#### **2.4 Mahalanobis' D<sup>2</sup>-statistics**

In the present investigation, genetic divergence was studied using Mahalanobis' generalized distance, as suggested by Rao [11]. Actual variable means were converted to uncorrelated variables using the pivotal condensation method of the inversion matrix. The D<sup>2</sup>-values between the genotypes were calculated as the sum of squares of differences of the values of the equivalent converted variables. For each pair of combinations, the mean deviation was determined. i.e.,  $d_i = Y_i^1 - Y_i^2$ , where  $Y_i$  signifies the converted variables ( $i = 1, 2, 3, 4, 5, \dots, p$ ) were calculated and the D<sup>2</sup> was then calculated as sum of the squares of those deviations, i.e.

$$D^2 = \sum (Y_i^1 - Y_i^2)^2$$

Where, p = Number of characters.

The importance of D<sup>2</sup>-values was tested by treating them as chi-square ( $\chi^2$ ) at p degrees of freedom, where p is the number of characters reflected.

##### **2.4.1 Grouping of genotypes by Tocher's method**

After organizing the D<sup>2</sup> values of all combinations of one genotype with the others in ascending order of magnitude, the genotypes were grouped into several clusters using Tocher's method, as explained by Rao [11]. The criterion used in this method was that any two varieties belonging to the same cluster should, on average, show a smaller D<sup>2</sup> value than those belonging to different clusters. Subsequently, inter- and intra-cluster distances were calculated, and their relationships were diagrammatically represented.

## 2.5 Cluster of D<sup>2</sup> values

All  $n(n-1)/2$  D<sup>2</sup> values were clustered using Tocher's method described by Rao [11].

### 2.5.1 Intra cluster distance

$$\text{Square of the intra cluster distance} = \frac{\sum D^2_i}{n}$$

Where,  $\sum D^2_i$  is the sum of distance between all possible combinations of the entries included in a cluster and  $n$  = Number of all possible combinations

### 2.5.2. Inter cluster distance

$$\text{Square of the intra cluster distance} = \frac{\sum D^2_i}{n_i n_j}$$

Where,  $\sum D^2_i$  is the sum of distances between all possible combinations ( $n_i n_j$ ) of the entries included in the clusters study;  $n_i$  = Number of entries in cluster I;  $n_j$  = Number of entries in cluster j

## 3. RESULT AND DISCUSSION

### 3.1 Genetic Diversity in Sorghum Genotypes Under Rainfed Conditions

D<sup>2</sup> statistics is a valuable tool for evaluating the genetic diversity of sorghum genotypes under drought stress [7]. Understanding genetic diversity in relation to drought stress is crucial for developing resilient sorghum cultivars capable of withstanding water scarcity. D<sup>2</sup> statistics enables the assessment of diversity among sorghum genotypes under both irrigated and rainfed conditions. The D<sup>2</sup> cluster diagram (Fig. 1) visually illustrates the genetic relationships and diversity among different sorghum genotypes [4]. Based on D<sup>2</sup> values calculated using Tocher's method, the genotypes under rainfed conditions were classified into six clusters: Cluster I with six genotypes, Cluster II with two genotypes, Clusters III and IV with four genotypes each, and Clusters V and VI with two genotypes each (Table 2). The criterion for grouping was that the average D<sup>2</sup> value within a cluster should be smaller than the D<sup>2</sup> value between clusters. Cluster I contained the highest number of cultivars, with six out of the total 20 genotypes (Table 3). Clusters II, V, and VI had only two cultivars each under stressed conditions. The cluster diagram and supplementary table provided the average inter-cluster and intra-cluster distances using D<sup>2</sup> values, representing the genetic diversity in a clear and comprehensible manner.

The D<sup>2</sup> value of intra-cluster distances ranged from 7.79 (Cluster II) to 12.14 (Cluster VI), indicating that clusters II and VI were highly distinct from clusters I, III, and V, as well as from each other. Cluster II appeared to be less distinct from Cluster IV and Cluster I compared to its

distinctiveness from Cluster III and V. The average inter-cluster distance varied from 13.23 (Clusters II and V) to 24.87 (Clusters III and V). Analyzing the  $D^2$  inter-cluster distances revealed that cultivars in Clusters III and V (with a distance of 618.52) and Clusters II and III (with a distance of 490.35) exhibited greater divergence and wider genetic variation compared to the other clusters. This distribution of genotypes across clusters clearly indicates the presence of genetic diversity among sorghum genotypes under rainfed conditions (Table 3).

### **3.2 Genetic Diversity in Sorghum Genotypes Under Irrigated Conditions**

Under irrigated conditions, the twenty sorghum genotypes were categorized into three clusters: Cluster I with sixteen genotypes, and Clusters II and III with two genotypes each (Fig. 2). Based on data collected under irrigated conditions, the genotypes were categorized into three clusters using Tocher's method of cluster formation, considering the relative magnitude of  $D^2$  values (Table 4). Cluster I contained the highest number of cultivars, with sixteen out of the total twenty genotypes, while Clusters II and III had only two genotypes each. The average inter-cluster and intra-cluster distances provided valuable insights into the genetic diversity and relationships among the different clusters. The  $D^2$  value of intra-cluster distances ranged from 8.14 (Cluster II) to 15.74 (Cluster III), indicating high divergence between Clusters II and III, as well as within each cluster. Analyzing the  $D^2$  inter-cluster distances revealed that cultivars in Clusters I and III (with a distance of 335.69) and Clusters II and III (with a distance of 391.19) exhibited greater divergence and wider genetic variation compared to the other clusters (Table 5).

The  $D^2$  cluster diagram [1] provides a clear and understandable representation of the clustering pattern and genetic distances between genotypes, allowing for the identification of groups or clusters of genotypes with similar genetic characteristics [2]. This information enhances our understanding of the genetic diversity and relatedness of sorghum populations under both irrigated and rainfed conditions. In summary,  $D^2$  statistics plays a crucial role in drought stress studies as it enables the assessment of genetic diversity in sorghum genotypes [12]. It aids in the identification of drought-tolerant genotypes, selection of suitable parental lines, understanding of genetic relationships, targeting of trait improvement, and conservation of genetic resources. The insights gained from  $D^2$  statistics contribute to the development of resilient sorghum varieties capable of withstanding water scarcity, ensuring sustainable agriculture.

**Table 2 : Average intra and inter cluster distance D<sup>2</sup> and D values among six clusters of twenty genotypes of sorghum by Tocher method under rainfed condition**

	<b>Cluster I</b>	<b>Cluster II</b>	<b>Cluster III</b>	<b>Cluster IV</b>	<b>Cluster V</b>	<b>Cluster VI</b>
<b>Cluster I</b>	131.50 (11.47)	175.83 (13.26)	381.80 (19.54)	199.38 (14.12)	273.87 (16.55)	221.12 (14.87)
<b>Cluster II</b>	175.83 (13.26)	60.62 (7.79)	490.35 (22.14)	309.54 (17.59)	175.11 (13.23)	301.10 (17.35)
<b>Cluster III</b>	381.80 (19.54)	490.35 (22.14)	122.61 (11.07)	265.09 (16.28)	618.52 (24.87)	252.57 (15.89)
<b>Cluster IV</b>	199.38 (14.12)	309.54 (17.59)	265.09 (16.28)	119.51 (10.93)	406.39 (20.16)	220.08 (14.84)
<b>Cluster V</b>	273.87 (16.55)	175.11 (13.23)	618.52 (24.87)	406.39 (20.16)	139.53 (11.81)	436.73 (20.90)
<b>Cluster VI</b>	221.12 (14.87)	301.10 (17.35)	252.57 (15.89)	220.08 (14.84)	436.73 (20.90)	147.35 (12.14)

**Table 3 :Grouping of 20 sorghum cultivars into 6 clusters under rainfed condition**

<b>Cluster</b>	<b>No. of genotypes included</b>	<b>Name of the genotypes</b>
Cluster I	6	ICSR-15001, Tandur L, ICSR-13025, SVD-1358R, CSV-216R, SVD-1403R
Cluster II	2	SVD-1528R, SPV-2217
Cluster III	4	Phule Anuradha, M-35-1, BJV-44, DKS- 35
Cluster IV	4	Basavan moti, Phule Vasudha, CSV-29R, SVD-1272R
Cluster V	2	Chitapur – L, M 148-138
Cluster VI	2	SPV-486, Basavana pada

**Table 4: Average intra and inter Cluster distance  $D^2$  and D values among six clusters of twenty genotypes of sorghum by Tocher Method under irrigated condition**

	<b>Cluster I</b>	<b>Cluster II</b>	<b>Cluster III</b>
<b>Cluster I</b>	187.52 (13.69)	288.49 (16.98)	335.69 (18.32)
<b>Cluster II</b>	288.49 (16.98)	66.26 (8.14)	391.19 (19.78)
<b>Cluster III</b>	335.69 (18.32)	391.19 (19.78)	247.70 (15.74)

**Table 5: Grouping of 20 sorghum cultivars into 3 clusters under irrigated condition**

<b>Cluster</b>	<b>No. of genotypes included</b>	<b>Name of the genotypes</b>
<b>Cluster I</b>	16	SVD-1358R, SVD-1528R, SVD-1403R, SPV-2217, CSV-216R, CSV-29R, ICSR-15001, Basavana pada, Phule Anuradha. Chitapur – L, M-35-1, M 148-138, Basavan moti, Phule Vasudha, BJV-44, ICSR- 13025
<b>Cluster II</b>	2	SVD-1272R, Tandur L
<b>Cluster III</b>	2	SPV-486, DKS- 35

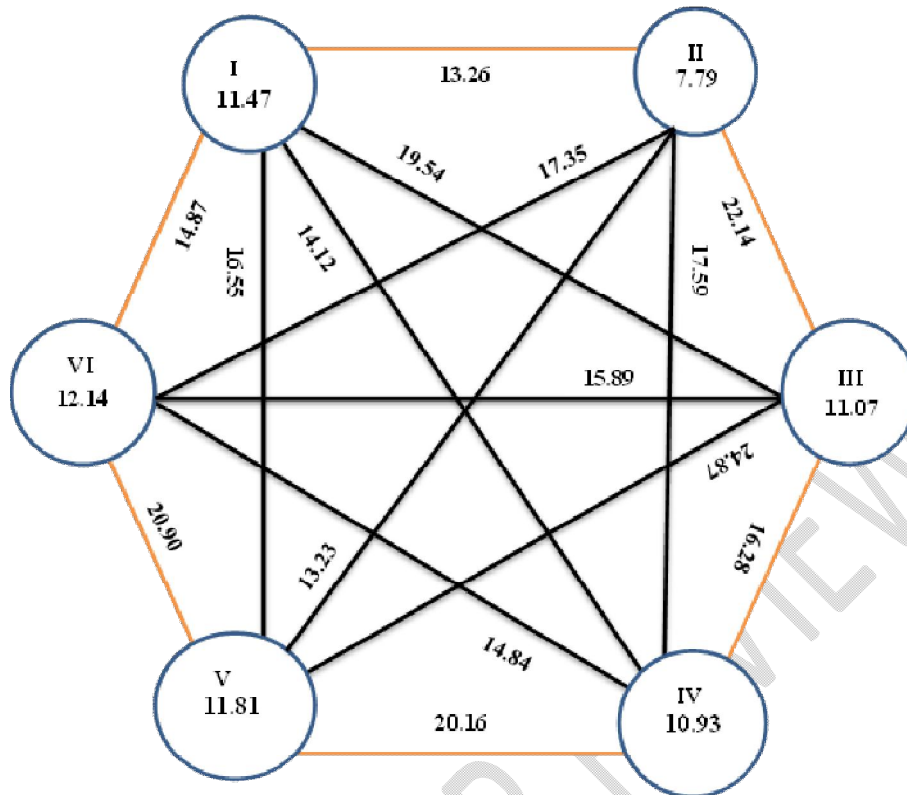


Fig 1: Cluster diagram showing inter-cluster and intra-cluster distance D values for sorghum genotypes under rainfed conditions

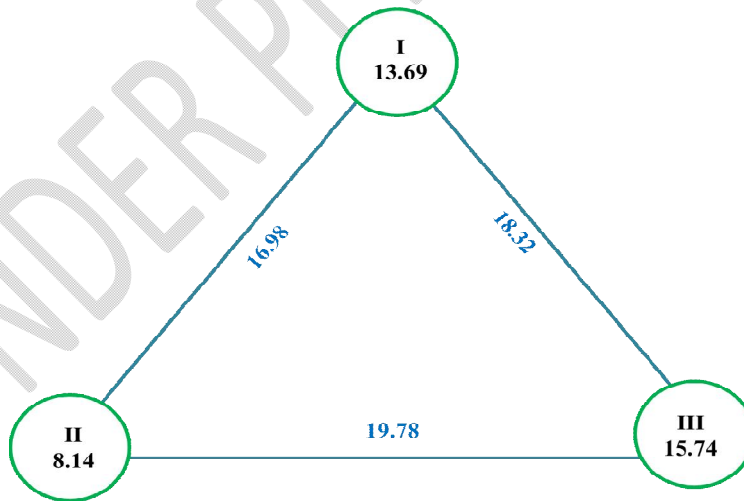


Fig 2: Cluster diagram showing inter-cluster and intra-cluster distance D values for sorghum genotypes under irrigated conditions

#### 4. CONCLUSION

This study underscores the significance of genetic diversity in sorghum (*Sorghum bicolor* L. Moench) for enhancing drought tolerance, a crucial trait for maintaining crop productivity in drought-prone regions. By employing Mahalanobis'  $D^2$  analysis, we effectively quantified the genetic divergence among 20 sorghum genotypes under both rainfed and irrigated conditions. The results revealed substantial genetic diversity, with genotypes grouped into distinct clusters based on their  $D^2$  values. Under rainfed conditions, six clusters were identified, highlighting significant genetic variation among the genotypes. Clusters III and V exhibited the greatest inter-cluster distances, indicating wide genetic variation and potential for breeding programs aimed at improving drought tolerance. Similarly, under irrigated conditions, three clusters were formed, with Clusters II and III showing the highest intra-cluster distances, reflecting considerable genetic divergence. The findings emphasize the importance of utilizing genetically diverse parental lines for hybridization to achieve high hybrid vigor and enhance drought tolerance. The  $D^2$  analysis provided a clear understanding of the genetic relationships and diversity among the sorghum genotypes, aiding in the identification of promising candidates for breeding programs.

#### DISCLAIMER (ARTIFICIAL INTELLIGENCE)

Author(s) hereby declare that NO generative AI technologies such as Large Language Models (ChatGPT, COPILOT, etc) and text-to-image generators have been used during writing or editing of manuscripts.

#### REFERENCES

1. Deep H, Chakraborty I, Arya S, Lamba P, Pahuja SK, Pantnagar J. Genetic diversity analysis by  $D^2$  clustering of fodder yield and its related traits in forage sorghum. Pantnagar Journal of Research. 2020;18(3):203-209.
2. Doijad SP, Bagade AB, More AW. Evaluation of sorghum germplasm for genetic diversity using  $D^2$  statistics. Electronic Journal of Plant Breeding, 2016;7(4):934-938.
3. Harrington JB. Yielding capacity of wheat crosses as indicated by bulk hybrid tests. Canadian J. Res. 1940;18:578-84.
4. Indian G, Sankaranarayanan R, Murugesan S, Rajangam J. Assessment of genetic divergence using Mahalanobis  $D^2$  analysis in mango. Journal of Agriculture and Ecology. 2019;7(7): 38–46.

5. Jakhar DH, Ray A, Singh R. Genetic Diversity Analysis in Maize (*Zea mays* L.) by  $D^2$  Clustering of Yield and Yield Accredited Traits. International Journal of Agriculture, Environment and Biotechnology; 2018; 3:833-836.
6. Joshi B K, Randy G. Gardner, Dilip RP. Diversity Analysis of Tomato Cultivars Based on Coefficient of Parentage and RAPD Molecular Markers, J.Crop Imp.2012;26:177–196.
7. Kumar S, Rattan P, Sharma JP, Gupta RK.  $D^2$  analysis for fruit yield and quality components in tomato (*Lycopersicon esculentum* Mill.). Indian J Plant Gen Res, 2010;23(3):318-320.
8. Lahane PS, Ravindrababu Y, Gami RA, Khule AA, Acharya S. Genetic diversity analysis of land races of maize (*Zea mays* L.) in Gujarat. Advances in Life Sciences. 2016; 5(7):2903-2907.
9. Mahalanobis PC. On the generalized distance in statistics. Proc. Nat. Inst. Sci. India. 1936;2:49-55.
10. Murty BR, Arunachalam V. Computer programmes for some problems in biometrical genetics of Mahalanobis  $D^2$  in classificatory problems. Indian J Gen. 1967;27:60-69.
11. Rao CR. Advanced Statistical Methods in Biometric Research. John Wiley Sons, New York, 1952; 390-398
12. Zeven AC. Landraces: A review of definitions and classifications Euphytica, 1998;104: 127-139.