



# Unveiling Genetic Diversity in Grain Amaranth through K-Means Clustering and Principal Component Analysis

Dhanalakshmi. T. N. <sup>a\*</sup>, Madhusudan K. <sup>a</sup>, Ramesh. S. <sup>a</sup>,  
Mohan Rao. A. <sup>a</sup>, Srinivas Reddy. K M. <sup>b</sup> and Anand S. R. <sup>c</sup>

<sup>a</sup> Department of Genetics and Plant Breeding, College of Agriculture, UAS, GKVK, Bengaluru- 560065, India.

<sup>b</sup> Department of Agricultural Entomology, College of Agriculture, UAS, GKVK, Bengaluru- 560065, India.

<sup>c</sup> Department of Agronomy, College of Agriculture, UAS, GKVK, Bengaluru- 560065, India.

## Authors' contributions

This work was carried out in collaboration among all authors. All authors read and approved the final manuscript.

## Article Information

DOI: <https://doi.org/10.9734/ijpss/2024/v36i105111>

## Open Peer Review History:

This journal follows the Advanced Open Peer Review policy. Identity of the Reviewers, Editor(s) and additional Reviewers, peer review comments, different versions of the manuscript, comments of the editors, etc are available here: <https://www.sdiarticle5.com/review-history/125837>

Original Research Article

Received: 26/08/2024  
Accepted: 28/10/2024  
Published: 04/11/2024

## ABSTRACT

Grain amaranth (*Amaranthus spp.*) is a highly versatile and nutrient-dense pseudo-cereal that has gained increasing importance in global food systems. Developing diverse varieties and hybrids of grain amaranth (*Amaranthus spp.*) is essential for harnessing its full potential as a climate-resilient and nutritionally rich crop. Germplasm studies provide a comprehensive understanding of the genetic diversity within amaranth, facilitating the breeding of improved cultivars that combine desirable traits such as higher yield, enhanced nutritional content, and increased tolerance to biotic and abiotic stresses. Varietal development and hybridization can optimize characteristics like seed

\*Corresponding author: E-mail: [ghanugpb@gmail.com](mailto:ghanugpb@gmail.com);

**Cite as:** T. N., Dhanalakshmi., Madhusudan K., Ramesh. S., Mohan Rao. A., Srinivas Reddy. K M., and Anand S. R. 2024. "Unveiling Genetic Diversity in Grain Amaranth through K-Means Clustering and Principal Component Analysis". *International Journal of Plant & Soil Science* 36 (10):611-20. <https://doi.org/10.9734/ijpss/2024/v36i105111>.

size, growth habit, and adaptability, making amaranth more suitable for a range of agroecological zones. In this direction present study conducted by evaluating 180 grain amaranth germplasm accessions for yield and contributing factors during the summer 2024 at the ICAR-IISS Regional Station in Bengaluru adjacent to NSP. Genetic distances among 180-grain amaranth germplasm accessions were determined using K-means clustering, resulting in 06 distinct groupings. Cluster V comprised of more number of accessions (48) and the least number of accessions in cluster VI (15). As per PCA, eight eigen values explains the total variance; among which PC<sub>1</sub> explained maximum variance 82.06% followed by PC<sub>2</sub>, PC<sub>3</sub>, PC<sub>4</sub>, PC<sub>5</sub>, PC<sub>6</sub>, PC<sub>7</sub> and PC<sub>8</sub> explained 8.87%, 5.39%, 1.63%, 1.02%, 0.51%, 0.45% and 0.07 % respectively.

**Keywords:** Grain amaranth; k means; clustering analysis; genetic diversity; principal component analysis.

## 1. INTRODUCTION

Grain amaranth belongs to the *Amaranthaceae* family, which encompasses over 60 species distributed across tropical, subtropical, and temperate regions globally. While several species within the genus possess edible seeds, the most commonly cultivated for grain production include *Amaranthus hypochondriacus*, *A. cruentus*, and *A. caudatus*. These species exhibit considerable genetic diversity, resulting in variations in plant morphology, seed color, size, and nutritional composition [1]. Amaranth is renowned for its adaptability to diverse growing conditions, thriving in regions with high temperatures, limited water availability, and poor soil fertility too. Its ability to withstand environmental stressors, such as drought, heat, and pests, makes it an attractive crop for smallholder farmers in marginal lands. Furthermore, grain amaranth demonstrates a short growing season, typically maturing within 90 to 120 days, thus allowing for multiple harvests in a single year.

Beyond its agronomic virtues, grain amaranth stands out for its exceptional nutritional composition and potential health benefits. The seeds are rich in high-quality proteins, containing all essential amino acids, particularly lysine, which is deficient in many other cereals. Additionally, grain amaranth is a good source of dietary fiber, vitamins (such as vitamin E, B-complex vitamins, and folate), minerals (including calcium, iron, magnesium, phosphorus, and potassium), and bioactive compounds like polyphenols and flavonoids. Overall, grain amaranth emerges as a remarkable crop with the potential to address persistent global challenges related to food security, nutrition, and sustainable agriculture. Its resilience, nutritional richness, and versatility make it a valuable addition to diversified cropping systems worldwide, offering

opportunities for smallholder farmers to enhance livelihoods and contribute to a more resilient and inclusive food system [2].

Utilizing diverse genetic resources can lead to more sustainable agricultural practices by reducing the reliance on chemical inputs and promoting agro-biodiversity. Breeding programs that utilize diverse genetic resources can produce high-value crops with unique qualities, benefiting farmers and the agricultural economy. By integrating genetic resources into breeding programs, there is a direct contribution to the conservation of plant biodiversity, ensuring the availability of these resources for future generations. To achieve such goals studies related to germplasm collection in crucial; with this aim, the present study has been conducted to analyse the 180-grain amaranth germplasm collection. Isolating the most valuable genotypes from evaluated germplasm pool is a crucial factor when selecting parents in crop breeding. Cluster analysis and principal component analysis assist plant breeders in identifying genetically varied parents to cater the farming community.

## 2. MATERIALS AND METHODS

180 grain amaranth germplasm material inclusive of indigenous and exotic used to conduct the experiment.

### 2.1 Experiment Layout

A total of 180-grain amaranth germplasm materials, comprising both indigenous and exotic accessions, were used in the experiment. These accessions were evaluated alongside two check varieties, Suvarna and KBGA-15, to assess their performance and variability in key agronomic traits in augmented design [3]. The experiment was structured using a systematic online design developed by the Indian Council of Agricultural

Research (ICAR), New Delhi in an augmented design fashion (accessible at <https://drs.icar.gov.in>), to ensure precise and reproducible results. The germplasm entries were organized into 15 blocks, each containing 12 germplasm accessions, which were randomly assigned within the experimental layout.

The crop was cultivated during the summer season of 2024, with a plant spacing configuration of 45 cm between rows and 10 cm between plants within a row. Each accession and check varieties were sown in 1.5 m row length. The field experiment was meticulously managed by following the recommended package of agronomic practices, to support healthy crop development.

Five randomly picked plants were chosen to record the data as per Protection of Plant Varieties and Farmers' Rights Authority (PPVFRA). Eight Quantitative traits such as; days to 50% flowering, leaf length (cm), leaf width (cm), petiole length (cm), inflorescence length (cm), plant height (cm), yield per plant (g) and volume weight of seeds (g/10ml) were recorded for each accession and check varieties as well.

## 2.2 Data Analysis

The following statistical procedure were followed in the present investigation. The average values for each accession were used to compute further statistical analysis.

## 2.3 Clustering Analysis

is a statistical method that groups objects or data points based on their similarities or dissimilarities. Cluster analysis divides a set of objects into clusters; those within one subset are more "similar" to one other, while those inside other subsets are "dissimilar." There must be a means to distinguish between "dissimilar" and "similar" objects. K-means clustering is a key approach for analysing data. K-means is a popular algorithm for clustering data based on known median locations [4].

## 2.4 Principal Component Analysis (PCA)

Principal components analysis is a method for identifying hypothetical variables (components) that explain as much of the variance in multidimensional data as possible. The new variables are linear combinations of the prior

ones. [5] described the principle component analysis [6] technique, which was used to condense 07 quantitative features reported in 180 grain amaranth germplasm collections into a small number of uncorrelated variables. In the present investigation, 180 grain amaranth germplasm accessions evaluated were subjected for K-means clustering and PCA by using 'R' software package.

## 3. RESULTS AND DISCUSSION

The analysis revealed significant variation among the germplasm accessions for key agronomic traits in the present study.

### 3.1 K-Means Clustering

K-means clustering is widely used in genetic and agronomic studies for classifying genotypes into distinct groups based on their trait similarities. It is particularly effective because it allows researchers to control the number of clusters, providing flexibility in grouping complex datasets [7]. In the present study, 180-grain amaranth genotypes were classified into 6 distinct clusters using the k-means algorithm, which facilitated the identification of genetic diversity patterns among the accessions. The grouping pattern is detailed in Table 1, revealing variations in the number of accessions assigned to each cluster [8].

**Table 1. Clusters distribution and size formed by k-means of 180-grain amaranth germplasm accessions**

Cluster number	Size
C1	38
C2	30
C3	25
C4	30
C5	42
C6	15
<b>Total</b>	<b>180</b>

Among the clusters, Cluster V emerged as the largest group, containing 48 accessions, indicating a higher level of similarity or common traits among these genotypes. This was followed by Cluster I, which consisted of 38 accessions, representing another diverse yet distinct set of genotypes. Clusters II and IV each comprised 30 accessions, demonstrating comparable diversity levels and trait groupings. Cluster III included 25 accessions, while Cluster VI, the smallest cluster, contained only 15 accessions, suggesting that these genotypes may possess unique traits or

characteristics that differentiate them from the others.

The clustering pattern illustrated that some clusters were more densely populated, potentially indicating genotypic redundancy or close genetic relationships among accessions within those groups. Conversely, smaller clusters like Cluster VI may represent unique or rare genotypes that could be valuable for breeding programs aimed at enhancing specific traits. The proportion of accessions within each cluster, along with the overall cluster sizes, is visually represented in Fig. 1, providing a clear overview of the distribution pattern. The distinct grouping of the genotypes reflects the underlying genetic structure and diversity present in the germplasm, which can be crucial for guiding future breeding strategies. The variability in cluster sizes also suggests that while some clusters might have a broad genetic base, others are likely more specialized, requiring

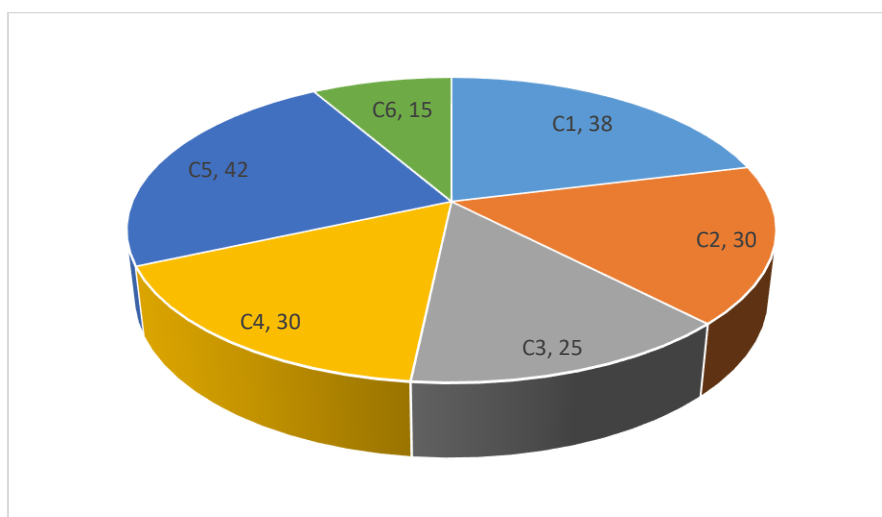
targeted efforts for genetic conservation or improvement.

Overall, k-means clustering not only facilitated a robust classification of the genotypes but also highlighted the diversity and complexity within the grain amaranth collection, providing valuable insights for germplasm management and the selection of promising parental lines for hybridization [9] and [10].

Table 2 displays the accessions categorized into six distinct clusters, illustrating the genetic diversity within the grain amaranth germplasm collection. This classification is essential for understanding the relationships and variations among the different genotypes. By grouping the accessions based on their phenotypic traits and genetic similarities, the table reveals underlying patterns of diversity that are critical for effective breeding strategies.

**Table 2. List of genotypes belonging to different clusters formed by k-mean clustering from 180 grain amaranth genotypes**

Cluster no.	Size	Genotypes
C1	38	IC0095325, IC38301, EC0519542, IC00382665, NIC33418, IC38408, IC0554125, IC0038416, EC322050, EC223651, IC0554058, IC35643, IC26266, IC0038255-5, IC5596, IC35457, EC223659, IC0554033, IC35713, IC35399, NIC22559, IC0042282-7, IC37514, EC277969, EC519509, IC5564, IC0038621, IC19585, IC326899, EC322993, IC120702, IC003851, IC195016, IC0021958, IC195822, IC13899, IC313253 and IC0035577
C2	30	EC34183, IC95255, IC0042311-7, IC0615024, IC17316, IC031100, IC0554102, IC0035701, NIC33487, EC10238, IC120617, NIC130270, IC35588, IC120671, IC95446, IC35680, IC0553999, IC0035461, IC35755, EC170310, IC0616907, IC95262, IC0448736, IC0024262, IC32886, IC0038539, IC341496, IC36824, IC415592 and IC0553987
C3	25	IC42344, IC95454, IC81701, IC95372, EC038702, IC0038134, EC359410, IC107607, IC95630, IC38356, EC223644, EC359408, IC95299, IC81694, EC289388, IC37150, IC0554010, IC395399, IC540817, IC19588, IC38014, IC038385, IC37145, NIC11318 and EC519511
C4	30	IC0554102, IC21981, IC0038265, IC415268, IC0024262, IC0095564, EC387000, IC1384528, IC35777, IC95284, IC325933, IC0108432, IC35699, NIC22536, IC35659, IC35426, EC130848, IC42337, IC120610, EC170326, IC395391, IC95634, IC33861, IC0281936, IC007933, EC323001, EC389899, IC0554124, IC422836 and IC031100
C5	42	IC0278926, IC396961, EC328874, NIC16168, IC328945, IC93945, IC0035556, IC195810, IC0042315-3, IC0035436, IC0120653, IC0035506, IC42309, IC120690, IC120668, IC120671, IC120748, IC0696, NIC21584, IC0528306, IC0035586, IC0120573, IC0281933, EC38261, IC0038573, IC0038206, IC35590, EC386997, IC0033758, IC0198131, IC35638, IC0035547, IC94662, IC0448663, IC0554124, IC95499, IC0037320, IC396972, IC95634, IC0038065, IC333749 and IC95372
C6	15	IC447680, IC95624, IC120664, IC249602, IC120671, IC0024262, IC0448639, IC0616907, IC0035470, IC95690, IC35755, IC0038554, IC108435, EC277970 and IC38259



**Fig. 1. Pie diagram depicting different clusters and size in 180 grain amaranth germplasm accessions**

The estimates of various quantitative traits mean across distinct clusters are summarized in Table 3, providing valuable insights into the phenotypic characteristics of the grain amaranth accessions. Notably, Cluster 2 exhibited the highest mean values for key traits such as leaf length, plant height, and volume weight of seeds. This indicates that the accessions within this cluster may possess superior growth attributes and seed quality, making them potential candidates for breeding programs aimed at enhancing these desirable characteristics.

In terms of leaf width and petiole length, Cluster 2 again showcased the best performance, suggesting that accessions in this cluster have well-developed foliage, which can contribute positively to photosynthetic efficiency and overall plant vigor. Conversely, Cluster 4 was distinguished by its significantly longer inflorescence length, which is an important trait for both aesthetic appeal and reproductive success. This characteristic could be advantageous for breeding initiatives focusing on ornamental or high-yielding varieties.

The analysis of flowering times revealed interesting trends as well. Clusters 4 and 6 had a lower number of days to 50% flowering, indicating that the accessions within these clusters are likely to mature earlier. Early flowering can be beneficial for cultivation in regions with shorter growing seasons or for farmers aiming to achieve multiple cropping cycles within a single year. In contrast, Cluster 2 exhibited a longer duration before reaching 50%

flowering, which may be attributed to its robust growth and developmental traits. Lastly, Cluster 3 stood out with the highest grain yield per plant, emphasizing its potential for high productivity. This finding highlights the importance of this cluster in breeding programs focused on increasing grain yield, which is a primary objective for enhancing food security and economic viability in grain amaranth production.

Overall, the data presented in Table 3 illustrates significant variability among the clusters regarding various quantitative traits, underscoring the genetic diversity within the grain amaranth germplasm collection. Understanding these trait means is crucial for guiding future breeding strategies, allowing researchers to select accessions that exhibit desirable combinations of traits for improved crop performance and resilience.

The estimates of variance among the clusters provided further insights into the genetic diversity within the germplasm collection (Table 4). For leaf length, the highest variance was observed in Cluster 3, reflecting a broad genetic range that could facilitate the selection of individuals with desirable leaf characteristics. Cluster 1 exhibited the greatest variance for both leaf width and plant height, indicating a diversity of morphological traits that may enhance adaptability to different environmental conditions. Additionally, Cluster 3 displayed significant variance in petiole length, which plays a crucial role in leaf arrangement and support. In terms of inflorescence length, Cluster 2 showed the

**Table 3. Estimate of quantitative trait means of 180-grain amaranth germplasm accessions belonging to different clusters**

Sl. No.	Traits	Mean of clusters						'F' statistic	Probability
		C1	C2	C3	C4	C5	C6		
1	Leaf length (cm)	15.76	19.42	14.47	9.93	15.2	10.24	25.88	6.2 × 10 <sup>-3</sup>
2	Leaf width (cm)	9.88	9.36	7.99	6.11	7.64	8.52	15.25	0.9 × 10 <sup>-3</sup>
3	Petiole length (cm)	12.38	13.33	9.78	7.82	11.94	10.98	13.11	0.2 × 10 <sup>-3</sup>
4	Inflorescence length (cm)	35.14	35.53	34.25	38.65	33.49	31.28	2.21	1 × 10 <sup>-3</sup>
5	Plant height (cm)	128.69	135.28	125.77	126.26	124.27	115.28	11.89	1 × 10 <sup>-3</sup>
6	Days to 50% flowering	35.62	38.28	39.14	31.23	32.43	31.26	11.03	1 × 10 <sup>-4</sup>
7	Grain yield plant-1 (g)	21.77	25.61	26.85	22.07	25.13	20.89	41.99	1 × 10 <sup>-3</sup>
8	Volume weight of seeds (g/10ml)	6.94	7.06	6.89	6.54	6.48	6.99	7.56	1 × 10 <sup>-3</sup>

**Table 4. Estimate of quantitative trait variance of 180-grain amaranth germplasm accessions belonging to different clusters**

Sl. No.	Traits	Variance of clusters						'F' statistic	Probability
		C1	C2	C3	C4	C5	C6		
1	Leaf length (cm)	20.82	10.31	17.03	0.34	19.58	10.28	11.29	1 × 10 <sup>-3</sup>
2	Leaf width (cm)	7.28	6.58	5.72	0.16	4.85	2.36	5.98	1 × 10 <sup>-3</sup>
3	Petiole length (cm)	5.27	7.66	27.72	1.62	19.95	8.56	10.98	1 × 10 <sup>-2</sup>
4	Inflorescence length (cm)	78.91	89.46	43.93	9.56	57.97	35.29	1.96	0
5	Plant height (cm)	423.83	399.01	276.36	164.71	487.98	400.28	19.7	1 × 10 <sup>-4</sup>
6	Days to 50% flowering	6.95	2.38	1.02	1.15	35.1	28.96	10.95	1 × 10 <sup>-4</sup>
7	Grain yield plant-1 (g)	3.24	0.92	5.84	2.83	7.23	6.28	20.36	1 × 10 <sup>-3</sup>
8	Volume weight of seeds (g/10ml)	0.28	0.2	0.31	0.14	0.61	0.58	0.54	2 × 10 <sup>-3</sup>

highest variance, highlighting the diversity in reproductive traits among the accessions within this group. Cluster 6 exhibited the greatest variance for days to 50% flowering, grain yield, and volume weight, suggesting a rich source of traits that can be utilized for breeding programs focused on improving both productivity and seed quality.

In summary, the combination of mean estimates and variance analysis across clusters underscores the substantial genetic diversity present within the grain amaranth germplasm collection. This diversity is essential for informing breeding strategies and selection practices, ultimately contributing to the development of improved varieties with enhanced agronomic performance, adaptability, and consumers' choice. Understanding both average trait performance and the extent of variability among accessions will significantly influence the success of breeding programs aimed at optimizing grain amaranth cultivars [11-14].

### 3.2 Principal Component Analysis

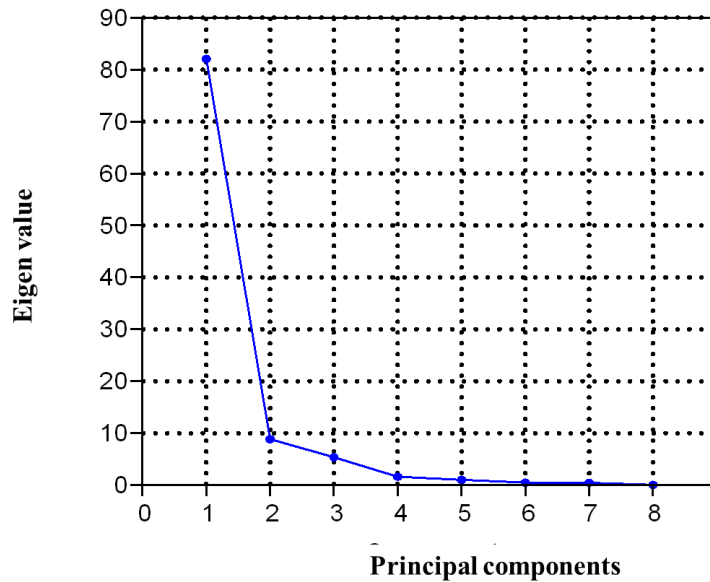
Principal Component Analysis (PCA) is a powerful statistical technique used to identify hypothetical variables, known as components, that account for the maximum variance within multidimensional datasets. These new variables are derived as linear combinations of the original variables, allowing researchers to simplify complex data structures while preserving essential information. The foundational work of Hotelling (1933) built upon the earlier contributions of Pearson (1901) to develop PCA, establishing it as a robust tool for data reduction and analysis. In this study, PCA was applied to condense nine quantitative traits collected from a diverse set of 180 grain amaranth germplasm accessions into a smaller number of uncorrelated variables. This approach enables a more

straightforward interpretation of the relationships among accessions based on their phenotypic traits. To ensure that each of the nine numerical characteristics contributed equally to the analysis, all traits were standardized. Standardization is crucial in PCA, as it eliminates biases that could arise from differences in measurement scales among the traits. The correlation matrix was constructed to represent the covariance between these standardized variables. This matrix serves as the foundation for calculating the Eigenvalues and standardized corresponding Eigenvectors, which are essential components of PCA. Eigenvectors provide the coefficients that define the principal components, while the Eigenvalues indicate the amount of variance accounted for by each principal component.

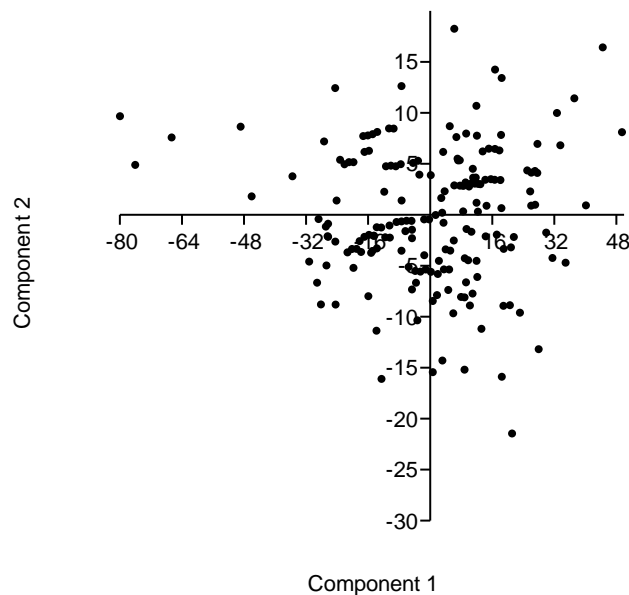
The analysis (Table 5) yielded eight Eigenvalues, each corresponding to a principal component. The Eigenvalue of a principal component reveals how much variance it explains relative to the total variance among the nine quantitative traits. By focusing on the first two principal components, which capture the largest portions of variance, researchers can gain significant insights into the underlying structure of the data. To visualize the results of the PCA, a scatter plot was created using the first two principal components for each germplasm accession. This two-dimensional graph (Fig. 2), allows for a clear visual representation of the relationships among the accessions. In the scatter plot (Fig. 3), closely related accessions will appear grouped together, indicating similarity in their trait expressions, while those that are distantly located will suggest greater dissimilarity. This visualization is instrumental in identifying clusters of accessions that share common characteristics, aiding in the selection of diverse parental lines for breeding programs.

**Table 5. Number of principal components and their Eigen values and contribution to total variability for 08 quantitative traits in 180 grain amaranth germplasm accessions**

Principal Component	Eigen value	% contribution to total variation
1	406.59	82.06
2	43.96	8.87
3	26.70	5.39
4	8.07	1.63
5	5.05	1.02
6	2.54	0.51
7	2.23	0.45
8	0.35	0.07



**Fig. 2. Contribution of each principal component to total explained variance in the phenotypic diversity of *Amaranthus spp***



**Fig. 3. Scatter plot depicting first two principal components explains maximum variance in the phenotypic diversity of *Amaranthus spp***

Overall, PCA serves as a valuable tool in this study by condensing the complexity of the germplasm data into comprehensible patterns, facilitating further exploration of genetic diversity and trait relationships within the grain amaranth collection. By employing PCA, researchers can make informed decisions regarding breeding strategies and the utilization of germplasm resources, ultimately contributing to the improvement of grain amaranth cultivars [15-17].

#### 4. CONCLUSION

K-means clustering was selected as the methodology to systematically organize the genetic variability present within a diverse set of 180 genotypes of grain amaranth. This clustering technique is particularly effective for analysing large datasets, as it allows for the classification of genotypes into distinct groups based on their phenotypic or genotypic characteristics. Genetic



distances among the 180 grain amaranth germplasm accessions were determined through the application of K-means clustering, which successfully identified six distinct groupings within the dataset. This classification allows researchers to understand the genetic variability present in the collection, facilitating targeted breeding strategies. The substantial variance explained by the first two principal components suggests that a relatively small number of traits drive the overall genetic diversity within the grain amaranth accessions.

By integrating K-means clustering and PCA in genetic studies enables researchers to map out genetic resources with greater precision, making it possible to build more efficient, targeted, and sustainable breeding programs. By systematically harnessing the diversity within the germplasm, these findings contribute to the development of new cultivars that meet the evolving challenges of food security and agricultural resilience. By identifying distinct clusters and understanding the principal components contributing to variance, researchers can make informed decisions regarding the selection of accessions for breeding programs. This knowledge not only aids in the enhancement of desirable traits but also contributes to the conservation and utilization of genetic resources in grain amaranth.

#### 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 this manuscript.

#### COMPETING INTERESTS

Authors have declared that no competing interests exist.

#### REFERENCES

1. Aderibigbe, OR, Ezekiel, OO, Owolade, SO, Korese, JK., Sturm, B, Hensel, O. Exploring the potentials of underutilized grain amaranth (*Amaranthus spp.*) along the value chain for food and nutrition security: A review. *Critical Reviews in Food Science and Nutrition*. 2022;62(3):656-669.
2. Dhanraha VK, Mandal J, Bhat JS, Heritable variation and predicted selection response of green yield and its component traits in vegetable amaranth. *International Journal of Bio-resource, Environment and Agricultural Sciences*. 2015;1(4):146-152
3. Federer, WT, Augmented designs with one-way elimination of heterogeneity. *Biometrics*. 1961;17:447-473.
4. Forgy, EW, Cluster analysis of multivariate data efficiency vs. interpretability of classifications. *Biometrics*. 1965;21:768-769.
5. Hotelling, H, Analysis of a complex of statistical variables into principal components. *Journal of Educational Psychology*. 1933;24:417-441.
6. Pearson K. On lines and planes of closest fit to system of points in space. *Philosophical Magazine*. 1901;2:557-572.
7. Macqueen J, Some Methods for classification and analysis of multivariate observations, *Proceedings of 5th Berkeley Symposium on Mathematical Statistics and Probability*, Berkeley, University of California Press. 1967;1:281-297.
8. Fang-Xiang Wu, Genetic weighted k-means algorithm for clustering large-scale gene expression data. *BMC Bioinformatics*. 2008;9(6):S12 DOI: 10.1186/1471-2105-9-S6-S12
9. Venkatesh L, Niranjana Murthy, SD Nehru, Analysis of genetic diversity in grain amaranth (*Amaranthus spp.*). *Indian Journal of Genetics and Plant Breeding*. 2014;74(4):522-525.
10. Vijayalakshmi, A. Lokeshwar Reddy, S. Rama Devi, M. Mohammed Imran, Genetic diversity study through K-Means clustering in germplasm accessions of chickpea (*Cicer arietinum L.*). *Electronic Journal of Plant Breeding*. 2023;13(4):1402-1407 Available: <https://doi.org/10.37992/2022.1304.172>
11. Pandey RM, Singh R. Genetic divergence in grain amaranth (*Amaranthus hypochondriacus L.*). *Genetika*. 2011;43(1):41-49.
12. Bhoya RC, Jadhao AG, Sivasubbu S, Singh AR, Sabharwal A, Palande NV. Genetic divergence analysis of amaranth (*Amaranthus hypochondriacus L.*). *International Journal of Developmental Neuroscience*. 2017;60:26-33.
13. Yogendra Kumar, Rameshwar Prasad, Chameli. Genetic divergence analysis in amaranthus (*Amaranthus spp.*) genotypes for yield and its component characters, *The Pharma Innovation*. 2019;8(12):41-45.
14. Sonam Singh Thakur, Bhagwat Saran Asati, Shri MK Chandrakar, Genetic

- diversity study in red amaranths (*Amaranthus tricolor* L.), The Pharma Innovation. 2021;10(11):660-663
15. Divyaramakrishnan CK, Savithramma DL. Tailoring genetic diversity of mungbean [*Vigna radiata* (L). Wilczek] germplasm through principal component and cluster analysis for yield and yield related traits. International Journal of Agronomy and Agricultural Research. 2014;5(2): 94-102.
  16. Solomon Mengistu, Maize Germplasm Characterization Using Principal Component and Cluster Analysis, Asian Journal of Plant Science and Research. 2021;11(6):199-204
  17. Mohd. Khurram Niyazi, Gideon Synrem, Sharad Pandey, Studies on Genetic Diversity and Principal Component Analysis in Wheat Germplasm (*Triticum aestivum* L.), International Journal of Plant & Soil Science. 2023;35(15):234-243.

**Disclaimer/Publisher's Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of the publisher and/or the editor(s). This publisher and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.

© Copyright (2024): Author(s). The licensee is the journal publisher. This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

*Peer-review history:*

*The peer review history for this paper can be accessed here:*  
<https://www.sdiarticle5.com/review-history/125837>