DOI : 10.58993/ijh/2026.83.1.2

Indian J. Hortic. 83(1), March 2026: 10-15

Genetic diversity analysis of mango genotypes using morphological and quality traits via D² statistics

Amit Kumar[1&2,*], Arvind Kumar[1], Satya Prakash[1] and Vibhu Pandey[1]

[1]Department of Fruit Science, Sardar Vallabhbhai Patel University of Agriculture and Technology, Meerut, U.P. 250 110, India

ABSTRACT

The present investigation was carried out during two consecutive years (2021-22 and 2022-23) on twelve mango (*Mangifera indica* L.) genotypes from an 8–10-year-old orchard to evaluate morphological and quality traits and assess genetic diversity. Fruits were harvested based on maturity indices such as skin color, flesh color and aroma, with undamaged fruits selected for evaluation. A total of 192 fruits representing 48 treatments were analyzed under a Randomized Block Design (RBD). The genotypes studied included Ambika, Pusa Arunima, Dashehari-51, Kesar, Pusa Surya, Mallika, Amrapali, Burma Surakha, Neelum Chausa, Mithua Malda, Rataul and Saurav. Multivariate analysis through Mahalanobis D² statistics and Tocher's method was employed to classify genotypes into distinct clusters. Pooled data over two years revealed four clusters, with Cluster IV recording the highest mean fruit yield per tree (338.31 fruits), followed by Cluster II (268.15) and Cluster III (234.85). The maximum inter-cluster distance was observed between Cluster IV and Cluster II (D=9.789), indicating substantial genetic divergence, while intra-cluster distance ranged from 0.01 (Cluster IV) to 3.114 (Cluster I). Cluster I, comprising six genotypes, exhibited the highest intra-cluster diversity. Based on inter-cluster distances, hybridization between Cluster IV (Amrapali) and Cluster II (Kesar, Burma Surakha and Rataul) is suggested to exploit heterosis and develop superior recombinants. The findings highlight significant genetic diversity among mango genotypes, providing useful information for selecting promising parents in breeding programs aimed at enhancing yield and quality attributes.

Key words: Mango genotypes, genetic diversity, mahalanobis D^2, morphological traits, breeding.

INTRODUCTION

Mango (*Mangifera indica* L.) is one of the most economically significant and widely cultivated fruit crops in the tropical and subtropical regions of the world. Valued for its distinctive flavor, attractive appearance, and rich nutritional profile, mango plays a central role in the fruit industry of many countries, particularly India (Bally & Dillon, 2). Believed to have originated in the Indo-Burma region, the species has undergone substantial genetic diversification over centuries through natural evolution and human-mediated selection (Mukherjee, 12). India, as the global leader in mango production, accounts for approximately 2.35 million hectares of cultivation and an annual output of 20.77 million tonnes (Anon, 1). The fruit is appreciated not only for its commercial value but also for its antioxidant, medicinal, and health-promoting properties (Kumar *et al.,* 6; Kumar *et al.,* 7).

Given the wide-ranging consumer preferences across regions and the diverse environmental challenges affecting production, a comprehensive understanding of genetic variability is critical for crop improvement. However, mango breeding programs often face limitations due to a lack of precise characterization of existing germplasm. While some efforts have been made to evaluate phenotypic traits, these are frequently univariate in nature and fail to capture the complex, multivariate interactions among traits relevant to yield, quality, and adaptability (Vincent and Anushma, 20).

The Mahalanobis D² statistic, a multivariate technique, has proven to be a valuable tool for assessing genetic divergence. It allows for the grouping of genotypes based on multiple quantitative traits and facilitates the identification of genetically diverse parents for use in hybridization programs (Crossa and Franco, 3). The technique offers a more holistic approach to parental selection, maximizing heterosis and the likelihood of obtaining superior recombinants. Despite its broad utility, the application of D² statistics to both morphological and quality-related traits in mango remains relatively underexplored, particularly in the context of Indian germplasm.

In the present investigation, emphasis is placed on the use of Mahalanobis D² analysis to quantify

*Corresponding author email: amitworld701@gmail.com

[2]Department of Fruit Science, Chandra Shekhar Azad University of Agriculture and Technology, Kanpur-208 002, U.P., India

genetic divergence based on second-order statistics, which, while having certain limitations in precision (Ma *et al.,* 10), provide meaningful insights into the relative contribution of each trait to total variability. By identifying inter-cluster distances and clustering genotypes based on multivariate performance, this approach aids in the rational selection of parents for breeding programs targeting quality improvement, stress adaptation, and yield stability (Kumar *et al.,* 8).

Understanding genetic distance is also essential for leveraging transgressive segregation, where offspring outperform both parents. Furthermore, standardizing variables before analysis ensures that each trait contributes equally to the measurement of genetic divergence, thereby improving the accuracy of selection decisions (Sheidai *et al.,* 17).

Therefore, the current study aims to assess the extent of genetic diversity among mango genotypes cultivated across India using Mahalanobis D^2 statistics based on key morphological and fruit quality traits. The ultimate objective is to identify genetically diverse and agronomically promising genotypes that can be utilized as parents in future mango breeding and hybridization programs.

MATERIALS AND METHODS

The present investigation was conducted during 2021 to 2023 at Horticultural Research Centre, Sardar Vallabhbhai Patel University of Agriculture and Technology, Meerut, Uttar Pradesh. The experimental material consisted of twelve genetically diverse mango genotypes collected from different agro-climatic regions of India, ensuring a broad genetic base. These included 'Ambika' (Amrapali × Janardhan Pasand) and 'Dashehari 51' (a clonal selection of Dashehari) obtained from the Central Institute for Subtropical Horticulture (CISH), Lucknow; 'Pusa Arunima' (Amrapali × Sensation), 'Pusa Surya' (a selection from Eldon), 'Mallika' (Neelum × Dashehari) and 'Amrapali' (Dashehari × Neelum) sourced from Indian Agricultural Research Institute (IARI), New Delhi. Additionally, 'Burma Surakha' and 'Saurav' were collected from Saharanpur district, while 'Rataul', an open pollinated seedling selection, originated from Baghpat district. Local cultivars such as 'Neelum Chausa' and 'Mithua Malda' were also included. The widely known cultivar 'Kesar', recognized for its Geographical Indication (GI No. 185), was obtained from Junagadh Agricultural University (JAU), Gujarat, where it is predominantly cultivated in the Gir region of Junagadh and Amreli districts.

All experimental trees were maintained at a uniform spacing of 6 × 6 meters and were 8 to 10 years old at the time of observation, representing the bearing stage suitable for reliable yield and quality assessment. The trial was laid out in a Randomized Block Design (RBD) with four replications, comprising a total of 48 trees. Each genotype was represented by four plants per replication. Standard agronomic practices including nutrient and pest management were uniformly followed across all treatments to ensure accurate genetic evaluation. The morphological and qualitative data collected used for multivariate analysis utilizing Mahalanobis (11) generalized distance, D^2 statistics which was used for computing genetic divergence. The original measurements were transformed into the standardized uncorrelated variables by pivotal condensation (Rao, 15). The divergence between any two genotypes was obtained as the sum of squares of the difference in the values of the corresponding transformed values.

For analysis, the average of each parameter's value or character was employed. Statistical Package for Agricultural Research (SPAR 2.0) and Indostat were used to perform the statistical analysis of the data. The Mahalanobis D^2 statistical methods were applied to the diversity analysis.

$$D^2=\lambda^{Ij}\delta j\delta_J$$

Where,

λ_{Ij} = Reciprocal matrix to the common dispersion matrix

λ_1 = Difference between the mean values of the two populations for Ith character

λ_J = Difference between mean values of the two populations for Jth character.

$D^2 = \Sigma\Sigma\ \delta_{IJ}d_{ij}$ = Sample estimated of X_{IJ}

Where,

δ_{IJ} - Sample estimated of X_{IJ}

d_i - Sample estimated of δ_I

Square of intra cluster distance = ΣDi2/N

Where, ΣDi2 = sum of distance between all possible combinations of entries included in a cluster

N = number of all possible combinations

The inter cluster distances were calculated by the formula described by Singh and Chaudhary (18).

Square of inter cluster distance = ------------- ∑ D2 /ninj

Where, ∑ D2 = sum of distances between all possible combinations of the (ninj) of the entries included in the cluster study ni = number of entries in cluster i nj = number of entries in cluster j

RESULTS AND DISCUSSION

The 12 inter-specific derivatives were grouped four clusters using the Tocher's method (Table 1). The distribution of 12 derivatives into four clusters for the pooled analysis. The mean performance for different clusters of genotypes for yield and its component was appreciably maximum for number of fruit per tree cluster IV (338.31) followed by cluster II

Table 1: Cluster mean Mahalanobis D^2 pooled analysis in different genotypes of mango.

Cluster		NF	FWT	FL	FW	FPW	SWT	SL	KWT	KL	KW	TSS	ACD	RS	NRS	TS	TC	AA	PC	TA	FY
I	Mean	228.39	244.00	11.15	6.74	158.30	32.88	10.52	17.69	6.61	2.67	18.66	0.25	5.51	9.49	15.53	4.27	32.04	64.28	0.59	57.92
	± SE	53.63	29.24	0.94	0.73	18.70	4.12	1.05	1.51	0.69	0.14	1.00	0.01	0.89	1.44	1.99	2.58	5.79	12.07	0.12	18.24
II	Mean	268.15	188.03	9.28	5.77	115.90	27.88	8.45	15.30	5.33	2.60	17.69	0.23	4.58	9.44	14.49	5.31	23.91	52.38	0.49	47.76
	± SE	83.47	48.27	0.96	0.48	42.66	5.16	1.14	1.69	0.54	0.22	0.18	0.03	0.28	0.74	0.96	3.12	3.00	11.84	0.12	3.33
III	Mean	234.85	328.94	12.43	7.31	233.70	37.40	10.61	20.74	6.57	3.35	19.79	0.21	5.70	11.65	17.97	4.33	38.53	55.65	0.80	77.34
	± SE	14.98	3.38	0.28	0.03	20.87	1.70	0.18	2.67	0.18	0.08	0.94	0.02	0.64	0.45	0.17	0.05	9.59	0.87	0.02	5.75
IV	Mean	338.31	321.45	12.90	7.08	243.30	36.21	10.18	19.34	6.83	2.99	22.03	0.28	4.91	14.62	20.31	7.73	37.94	119.73	1.28	108.82
	± SE	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

Where, NF:Number of fruit per tree; FWT:Fruit weight (g); FL:Fruit length (cm); FW:Fruit width (cm); FPW:Fruit pulp weight (g); SWT:Stone weight (g); SL:Stone length (cm); KWT:Kernel weight (g); KL:Kernel length (cm); KW:Kernel width (cm); TSS:TSS (Brix); ACD:Acidity (%); RS:Reducing sugar (%); NRS:Non-reducing sugar (%); TS:Total sugar (%); TC:Total carotenoid (mg/100g); AA:Ascorbic acid (mg/100 g); PC:Phenol content (mg GAE/100 g); TA:Total antioxidants (μmol Trolox 100g); FY:Fruit yield (kg)/tree

(268.15), cluster III (234.85). On the other hand, this cluster also possessed least mean value for cluster I (228.39). Fruit weight (g) reported maximum cluster III (328.94) followed by cluster IV (321.45), cluster I (244.00) whereas least mean value resulted for cluster II (188.03). Similarly, fruit length (cm) also resulted maximum for cluster IV (12.90) following cluster III (12.43), cluster I (11.15) and least mean value for cluster for cluster II (9.28). Furthermore, fruit width (cm) also had maximum range cluster III (7.31) followed by cluster IV (7.08), cluster I (6.74) and least was noticed for cluster II (5.77). Moreover, fruit pulp weight (g) also reported maximum under cluster IV (243.30) followed by cluster III (233.70), cluster I (158.30) and least mean value reported by cluster II (115.90). In further the stone weight (g) also reported maximum weight for cluster III (37.40) followed the similar pattern by cluster IV (36.21), cluster I (32.88) and least mean value for cluster II (27.88).However, stone length (cm) also tends to report maximum cluster I (10.61), followed by cluster III (10.61), cluster IV (10.18) and least mean value for cluster II (8.45). Kernel Weight (g) also found maximum for cluster III (20.74), cluster IV (19.34), cluster I (17.52) and least mean value for cluster II (15.30). Moreover, maximum kernel length (cm) was recorded under cluster IV (6.83) followed by cluster I (6.61), cluster III (6.57) and least mean value for cluster II (5.33). The similar pattern was observed for kernel width (cm) which found maximum for cluster III (3.35) followed by cluster IV (2.99), cluster I (2.67) and least mean value for cluster II (2.60). In further for TSS (Brix) maximum was drawn for cluster IV (22.03) in line with cluster III (19.79), cluster I (18.66) and least mean value for cluster II (17.69). Acidity (%) cluster IV (0.28), cluster I (0.25), cluster II (0.23) and least mean value for cluster II (0.23). Reducing Sugar (%) also reported maximum cluster III (5.70) further following the cluster I (5.51), cluster IV (4.91) and least mean value for cluster II (4.58). Non-reducing sugar (%) had maximum cluster IV (14.62) with (11.65), (9.49) and least mean value for cluster I (9.44). Total Sugar (%) also had similar trend with maximum cluster IV (20.31), cluster III (17.97), cluster I (15.53) and least mean value for cluster II (14.49). Relatively, the total carotenoid (mg/100g), also had maximum cluster IV (7.73) with in line with cluster II (5.31), cluster III (4.33) and least mean value for cluster I (4.27). Ascorbic acid (mg/100 g) also identified the maximum cluster IV (37.94) followed by cluster III (38.53), cluster I (32.04) and least mean value for cluster II (23.91). However, Phenol content (mg GAE/100 g) revealed maximum cluster IV (119.73) followed by cluster I (64.28), cluster III (55.65) and least mean value for cluster II (52.38). Total antioxidants (μmol Trolox 100g) cluster

IV (1.28), cluster III (0.80), cluster I (0.59) and least mean value for cluster II (0.49). The fruit yield (kg)/ Tree further stated the maximum cluster pattern under IV (108.82), followed by cluster III (77.34), cluster I (57.92) and least mean value for cluster II (47.76).

From the present experiment the detailed analysis revealed to be worth noted here that the analytical observations which concluded that cluster III appeared to be the most promising cluster to get the high yielding genotypes with maximum in terms of quantitative and quality traits *viz.,* (number of fruit per tree, fruit length (cm), fruit pulp weight (g), stone weight (g), kernel weight (g), kernel length (cm), kernel width (cm), TSS (Brix), acidity (%), non-reducing sugar (%), total sugar (%), total carotenoid (mg/100g), phenol content (mg GAE/100 g), total antioxidants (µmol Trolox 100g) and fruit yield (kg)/Tree). However, cluster IV prominently proved significantly next better in terms of quantitative and quality traits *viz.,* fruit weight (g), fruit length (cm), fruit width (cm), fruit pulp weight (g), stone weight (g), stone length (cm), kernel weight (g), kernel length (cm), kernel width (cm), TSS (Brix), reducing sugar (%), non-reducing sugar (%), total sugar (%), ascorbic acid (mg/100 g), total antioxidants (µmol Trolox 100g). Such character mean-based clustering has also been reported by Kumar *et al.* (8), Rajpoot *et al.* (14), and Umesh *et al.* (19).

When targeting genotypes for a breeding program, cluster-based mean estimations are highly helpful since they eliminate the laborious process of excluding out inferior germplasm lines as stated by Lal *et al.* (9). Therefore, in advanced breeding trials, genotypes from desired clusters could be employed straight for final field evaluation.

For pooled analysis the average inter and intra cluster distances among the five clusters for the pooled year analysis are presented (Table 2). The perusal of inter and intra cluster distance revealed that inter-cluster values were greater than intra cluster values. The maximum inter-cluster distance was observed (D=9.789) between cluster IV and II followed by cluster III and II (D=7.508) and cluster IV and I (D=7.359). However, the intra-cluster distance (D) ranged from 0.01 (cluster IV) to 3.114 (cluster I). It is very cleared from the data that minimum inter-cluster distance reveals that genotypes are closely related and this close proximity is widely used in backcrossing programs. However, maximum intra-cluster distance indicates that the genotypes are genetically divergent. Earlier studies have been well corroborated to the present findings with that of Sankaran *et al.* (16) in mango, Plathia *et al.* (13) in jamun.

The analysis of variance for the pooled year revealed significant differences among the mango genotypes for all the different traits suggesting considerable genetic variability in the population. Using the D^2 value as squares of generalized distance, all the mango genotypes were grouped into 4 cluster (Table 3). Maxiumum genotypes were grouped into cluster I (Dashehari-5, Pusa Surya, Neelum Chausa, Mithua Malda and Saurav) which included six genotypes. The cluster II (Kesar, Burma Surakha and Rataul) which included three genotypes. Moreover, cluster III (Ambika and Pusa Arunima) also two genotypes and cluster IV (Mallika) which was one genotype. The present finding highlights that D^2 statistics has been identified as a useful method for estimating genetic divergence as reported by Das *et al.* (4). Additionally, because it is a numerical estimate, it offers the benefit of allowing comparisons across all potential population pairs within a given group as stated by Ma *et al.* (10).

The genotypes that belong to the same cluster most likely don't differ all that much from one another. It is not anticipated that crossing types within the same cluster will produce desired characteristics. Crosses between cluster members that are separated by intercluster distances are therefore probably advantageous for future breeding programmes as referred by Kar *et al.* (5).

The present investigation highlights the pivotal role of genetic divergence in the selection of parents for hybridization in mango breeding programs. The results clearly indicate that crossing

Table 2: Average inter and intra (bold) distances values for genotypes of mango.

Cluster	I	II	III	IV
I	3.114			
II	4.203	**2.794**		
III	4.596	7.508	**1.652**	
IV	7.359	9.789	6.047	**0.001**

Table 3: Distribution of mango genotypes in clusters pattern of 12 mango genotypes by Tocher's method.

Cluster	No. of genotypes	Genotypes
I	6	Dashehari-51, Pusa Surya, Mallika, Neelum Chausa, Mithua Malda, Saurav
II	3	Kesar, Burma Surakha, Rataul
III	2	Ambika, Pusa Arunima
IV	1	Amrapali

parents from highly divergent clusters leads to the development of promising hybrids, likely due to the complementary gene action arising from their diverse genetic backgrounds. Among the studied clusters, the maximum inter-cluster distance was recorded between Cluster IV and Cluster II, signifying a wide genetic gap. This suggests that genotypes from Cluster IV (Amrapali) and Cluster II (including Kesar, Burma Surakha and Rataul) possess considerable potential for generating superior segregates when used as parental lines. The findings of this study, therefore, offer valuable insights into the genetic architecture of mango and provide a robust basis for selecting genetically diverse parents to enhance the efficiency of breeding programs. Thus, the study contributes significantly to understanding the extent of genetic diversity and its practical implications in the genetic improvement of mango.

AUTHOR'S CONTRIBUTIONS

Conceptualization of research (AK, AK[2], SP); Designing of the experiments (AK[2], SP, AK); Contribution of experimental materials (AK, AK[2]); Execution of experiments and data collection (AK); Analysis of data and interpretation (AK); Preparation of manuscript and literature review (AK, VP); Supervision and critical review (AK[2], SP, AK).

DECLARATION

The authors declare no competing interests.

ACKNOWLEDGMENT

The authors acknowledge the support rendered by Department of Fruit Science, College of Horticulture, SVPUA&T, Meerut for providing necessary facilities to carry out the research work.

REFERENCES

1. Anonymous, 2023-24. Indian horticulture database. Ministry of Agriculture & Farmers Welfare, Government of India. www.agricoop.nic.in

2. Bally, I. S. and Dillon, N. L. 2018. Mango (*Mangifera indica* L.) breeding. *Advances in Plant Breeding Strategies: Fruits:* **3**: 811-896.

3. Crossa, J. and Franco, J. 2004. Statistical methods for classifying genotypes. *Euphytica*, **137**: 19-37.

4. Das, S. S., Kishore, K., Lenka, D., Dash, D. K., Samal, K. C., Samant, D. and Dash, S. N. 2021. Assessment of genetic diversity in morphological traits of mango genotypes using D^2 statistics. *Int. J. Plant Soil Sci.* **33**(17): 31-42.

5. Kar, R. K., Mishra, T. K., Pradhan, B., Behera, D., Das, S., Swain, D. K. and Hossain, A. 2024. Assessment of the recurrent mutagenesis efficacy of sesame crosses followed by isolation and evaluation of promising genetic resources for use in future breeding programs. *Open Agric.*, **9**(1): 20220311.

6. Kumar, A., Kumar, A., Prakash, S., Singh, S., Pandey, V., Raghav, D. K. and Pal, D. 2023. Assessment of different elite mango varieties suitable for North western plain zones of Uttar Pradesh. *Environ. Conserv. J.*, **24**(4): 214-220.

7. Kumar, A., Kumar, A., Prakash, S., Kumar, V., Singh, A., Singh, M. K. and Singh, S. 2024. Evaluation on morphological and quality attributes of mango cultivars in Western Uttar Pradesh. *J. Adv. Biol. Biotechnol.*, **27**(9): 485-497.

8. Kumar, A., Gupta, A. K., Kanth, N., Marboh, E. S. and Yadav, S. K. 2024. Assessment of variability and genetic divergence of longan genotypes (*Dimocarpus longan* Lour.) based on qualitative and quantitative traits. *Euphytica*, **220**(5): 72.

9. Lal, N., Pandey, S. K., Pandey, S. D. and Nath, V. 2023. Genetic diversity and grouping of litchi genotypes based on 83 qualitative and quantitative traits. *Erwerbs-Obstbau*, **65**(4): 1003-1012.

10. Ma, X., Wu, H., Liu, B., Wang, S., Zhang, Y., Su, M. and Luo, X. 2024. Genomic diversity, population structure, and genome-wide association reveal genetic differentiation and trait improvements in mango. *Hortic. Res.*, **11**(7): 153.

11. Mahalanobis PC 1936. On the generalized distance in statistics. *Proc. Nat. Inst. Sci.* **2**: 49-55.

12. Mukherjee, S. K. 1948. The varieties of mango (*Mangifera indica* L.) and their classification. *Bull. hot. Soc.* **2** (2): 33-19.

13. Plathia, M., Bakshi, P., Sharma, N., Jamwal, M. and Jasrotia 2022. Character association and divergence studies in seedling population of Jamun (*Syzygium cumini* L.). *Int. J. Adv. Biochem. Res.*, **8**(2): 595-600.

14. Rajpoot, M., Sharma, T. R. and Sharma, R. M. 2024. Genetic variability in mango clones of langra cv. for physical-bio chemical parameters at Kymore Plateau and Satpura Hills of Madhya

Pradesh, India. *Int. J. Plant Soil Sci.*, **36**(8): 364-373.

15. Rao, C.R. 1952. Advanced statistical methods in biometrical research, Wiley, New York.

16. Sankaran, M., Dinesh, M. R., Gowda, D. C. S. and Venugopalan, R. 2020. Genetic Analysis in mango (*Mangifera indica* L.) based on fruit characteristics of 400 genotypes. *J. Hortic. Sci.*, **15**(2): 161-172.

17. Sheidai, M., Noormohammadi, Z. and Vazifeshenas, M. R. 2024. A new look at genetic variability in pomegranate: Landscape genetics and structural equation modeling (SEM). *Sci. Hortic.*, **334**: 113323.

18. Singh, R.K. and Chaudhary, B.D. 1977. Biometrical methods in quantitative genetic analysis. New Delhi. Kalyani Publication: 318 p.

19. Umesh Kanna, S., Parthiban, K. T., Senthilraja, K., Venkatesan, S., Udhaya Nandhini, D., Mohan Kumar, S. and Geethalakshmi, V. 2024. Genetic diversity and structure of *Terminalia bellerica* (Gaertn. Roxb.) population in India as revealed by genetic analysis. *Plants*, **13**(4): 470.

20. Vincent, L. and Anushma, P. L. 2024. Mango 14. *Fruit and Nut Crops*, 393.

(Received : December, 2025; Revised : March, 2026; Accepted : March, 2026)