

# Assessing the Genetic Diversity of Indonesian Mango (*Mangifera indica* L.) Germplasm Based on Morphological Character

Sukartini Sukartini, Hadiati Sri \*, Budiyanti Tri, Mansyah Ellina, Riska Riska, Jumjunidang Jumjunidang, and Ihsan Farihul

Research Center for Horticulture, Research Organization for Agriculture and Food, Indonesia National Research and Innovation Agency (BRIN). Jl. Raya Jakarta – Bogor Km. 46, Cibinong, 16811, Bogor, West Java, Indonesia.

**Abstract.** The dissemination of mango (*Mangifera indica* L.) into Indonesia has led to remarkable genetic and phenotypic diversification. This diversity primarily results from the species' open pollination system, which enables natural hybridization and promotes the emergence of new cultivars with distinct characteristics. Over extended periods of adaptation to diverse agroecological conditions, these genotypes have evolved into stable local cultivars that reflect strong environmental interactions and selection pressures. Currently, Indonesia conserves approximately 187 local and introduced mango cultivars in a centralized germplasm collection. This extensive diversity serves as a valuable genetic resource for pre-breeding, breeding, and conservation programs aimed at enhancing mango productivity, adaptability, and fruit quality. Morphological and phenotypic characterization provides a cost-effective and reliable method for assessing genetic variation, serving as a preliminary step before employing molecular or genomic approaches. This study aimed to evaluate the genetic diversity of 187 Indonesian mango accessions based on phenotypic traits. Principal Component Analysis (PCA) revealed that nine quantitative characters contributed predominantly to total variation, explaining 48.023% of the diversity (PC1). In contrast, qualitative traits showed no single character with a dominant influence on overall variability. Cluster analysis based on qualitative characteristics further grouped the 187 accessions into three major clusters, indicating the existence of phenotypic similarities that may reflect underlying genetic relationships. These findings highlight the substantial genetic richness of Indonesian mango germplasm and its potential utilization in future genetic improvement and conservation strategies.

## 1 Introduction

Mango (*Mangifera indica* L.) is one of the most economically important tropical fruit crops, cultivated widely across tropical and subtropical regions. Its significance is reflected not only

---

\*Corresponding author: [shadiati@yahoo.com](mailto:shadiati@yahoo.com)

in the scale of production but also in its contribution to global trade and nutrition. Over the past five years, mango production in Indonesia has shown a steady increase, rising from 2,203,789 tons in 2017 to 2,835,442 tons in 2021. Correspondingly, the export value also increased from US\$ 2.02 millions to US\$ 4.56 millions during the same period [1]. Such growth demonstrates the expanding market demand and the important role of mango as a high-value horticultural commodity. Nutritionally, mango fruit provides a rich source of carbohydrates, proteins, lipids, dietary fiber, minerals, vitamin C, carotenoids, and phenolic compounds, which collectively contribute to growth, development, and health maintenance [2]. Additionally, mango peel contains abundant fiber, minerals, and antioxidants, while the seeds are rich in starch, lipids, sterols, and tocopherols [3], making almost all parts of the fruit valuable for both food and industrial applications.

Taxonomically, mango belongs to the family *Anacardiaceae*, genus *Mangifera*, which comprises approximately 45 to 69 species worldwide [4]. The genus is predominantly distributed in tropical Asia, with its highest diversity in western Malesia. The domestication of *M. indica* is estimated to have occurred at least 4000 years ago, originating from the Assam Valley near the western border of the Indochina region of Myanmar during the Quaternary period. From this center of origin, the species spread across the Indian subcontinent and subsequently throughout the tropics. Besides *M. indica*, at least 26 other *Mangifera* species produce edible fruits traditionally consumed in Southeast Asia, including *M. altissima*, *M. caesia*, *M. foetida*, *M. kemang*, *M. laurina*, *M. odorata*, *M. pajang*, and *M. pentandra* [5].

Globally, more than 1,000 mango cultivars have been identified and classified into two major groups: the Indochinese and West Indian types. The Indochinese group is generally characterized by flat, kidney-shaped fruits with light green or yellow skin and little red coloration, whereas the West Indian group typically bears rounder, plumper fruits with bright red or orange hues [6]. The high morphological variability among cultivars reflects complex interactions between genetic factors, hybridization events, and diverse environmental adaptations [2], [6], [7], [8], [9]. Assessing genetic and phenotypic diversity is therefore essential for the effective management, conservation, and utilization of germplasm resources. Such studies form the basis for selecting genotypes with superior agronomic traits and for developing improved cultivars through targeted breeding programs.

Morphological characterization remains a fundamental and widely applied approach in genetic diversity studies. It offers several advantages, including cost-effectiveness, ease of implementation, and applicability in both field and controlled conditions. However, the expression of morphological traits can be strongly influenced by environmental factors and their inheritance is often complex, limiting their use as sole indicators of genetic variability [10], [11], [12]. Hence, morphological analysis is considered a preliminary but indispensable step before applying more advanced molecular and biochemical techniques.

Indonesia possesses a rich and diverse mango germplasm collection, comprising 208 varieties (298 accessions) maintained at the Cukurgondang Experimental Field, Pasuruan, East Java. This collection, established during the Dutch East Indies period in 1941, includes both local and introduced cultivars representing a broad spectrum of phenotypic variation. Understanding the relationships among fruit-related traits and clarifying the extent of phenotypic diversity within this collection are essential for germplasm evaluation and breeding efforts. Therefore, the present study aims to assess the phenotypic diversity and genetic relationships among 187 mango accessions using morphological descriptors. The findings are expected to provide valuable insights for germplasm management and the development of improved mango cultivars in Indonesia and other tropical regions.

## 2 Materials and Methods

A total of 187 mango (*Mangifera indica* L.) accessions, including both Indonesian local cultivars and introduced varieties aged between 30 and 80 years, were evaluated in this study. All accessions were maintained in the Cukurgondang Experimental Field, Pasuruan, East Java, Indonesia (Appendix 1). The trees were planted in uniform field conditions and managed under standard agronomic practices. For morphological characterization, ten samples of fully expanded mature leaves and ten samples of physiologically ripe fruits were collected from each accession to represent phenotypic variation.

Morphological characterization was conducted following the *Mango Descriptor List* developed by IPGRI [13], encompassing 15 quantitative and 12 qualitative traits. The quantitative traits included two leaf characters—leaf length (LL) and leaf width (LW); eight fruit traits—fruit weight (FW), fruit length (FL), fruit diameter (FD), pulp thickness (FT), edible portion (EP), and total soluble solids (TSS); four stone traits—stone weight (STW), stone length (STL), stone width (STD), and stone thickness (STT); and three seed traits—seed weight (SW), seed length (SL), and seed width (SD).

The qualitative traits observed were fruit shape (FS), skin colour of ripe fruit (SC), pulp colour (PC), pulp texture (PT), pulp aroma (PA), seed shape (SS), leaf attitude relative to the branch (LA), leaf texture (LT), leaf blade shape (LS), leaf margin (LM), leaf apex shape (LAS), and leaf base shape (LBS). Qualitative descriptors were scored according to the established IPGRI scale.

Descriptive statistical analyses, including the calculation of minimum, maximum, mean, standard deviation (SD), and coefficient of variation (CV), were conducted for quantitative traits to determine the extent of morphological variation. The quantitative data were further analyzed using Principal Component Analysis (PCA) to identify the most influential traits contributing to overall variation among accessions. For qualitative data, frequency distribution and diversity indices were calculated to assess character variability.

The dominance index (C) was calculated to determine the relative contribution of each character, categorized as follows:  $0 < C < 0.5$  = low dominance;  $0.5 < C \leq 0.75$  = moderate dominance; and  $0.75 < C \leq 1.0$  = high dominance. The Shannon–Wiener diversity index ( $H'$ ) was employed to evaluate population-level phenotypic diversity, using the following thresholds:  $H' \leq 1.0$  = low diversity;  $1.0 < H' \leq 3.0$  = moderate diversity; and  $H' \geq 3.0$  = high diversity [14]. Cluster analysis based on qualitative data was performed using the unweighted pair group method with arithmetic mean (UPGMA) and Euclidean distance coefficients to examine relationships among accessions. The robustness of clustering was tested using bootstrap resampling (Boot N = 9999) to ensure the reliability of the grouping patterns [15]. All statistical analyses were conducted using PAST software version 4.04 kasih [16]

## 3 Results

### 3.1 Quantitative characters

The high coefficient of variation values for fruit, stone and seed weight traits, namely 53.52%, 40.58% and 36.16% respectively, indicate high data diversity for each of these characters. The edible portion of flesh value showed the lowest variation (9.53%). Leaf traits (length and width) had moderate variation. Most fruit traits had higher diversity than leaf traits and edible portions (Table 1).

**Table 1.** Descriptive statistics of 15 quantitative traits of 187 mango accessions

No	Quantitative Traits	Minimum	Maximum	Mean	SD	CV (%)
1.	Fruit weight (g)	75.00	1066.80	337.35	180.53	53.52
2.	Fruit length (cm)	4.95	21.00	11.09	3.02	27.23
3.	Fruit diameter (cm)	4.54	10.67	7.47	1.29	17.26
4.	Fruit pulp thickness (cm)	0.99	3.06	1.90	0.45	23.80
5.	Stone weight (g)	12.00	128.00	39.44	16.00	40.58
6.	Stone length (cm)	4.13	17.75	8.65	2.54	29.30
7.	Stone width (cm)	2.40	6.28	3.73	0.62	16.55
8.	Stone thickness (cm)	1.27	2.95	2.00	0.28	13.95
9.	Seed weight (g)	5.53	65.00	19.93	7.21	36.16
10.	Seed length (cm)	3.30	10.20	6.20	1.02	16.52
11.	Seed width (cm)	1.93	5.00	2.98	0.52	17.41
12.	Edible portion (%)	47.30	84.45	71.15	6.78	9.53
13.	TSS (Brix)	8.00	25.00	15.49	2.94	19.00
14.	Leaf length (cm)	8.50	29.30	19.53	3.28	16.81
15.	Leaf width (cm)	3.70	8.30	5.38	0.80	14.88

PCA analysis uses the correlation menu, namely standardized variance-covariance (because the units of measurement for each data are different) with a bootstrap N of 9999. The cumulative four main components are 77.572. Since the variance is mostly due to one component (PC1), this PCA is successful (Table 2).

PCA analysis produces eigenvalue and % variance. Eigenvalues below the red line (broken stick) represent non-significant components (Figure 1). Based on this, all data is plotted in one coordinate system (Figure 2). Projection of variables or characters into a scattergram visualizes PCA loadings or coefficients (Figure 3). Table 3 explains the relationship between the original variables and the new variables (principal components) formed using principal component analysis.

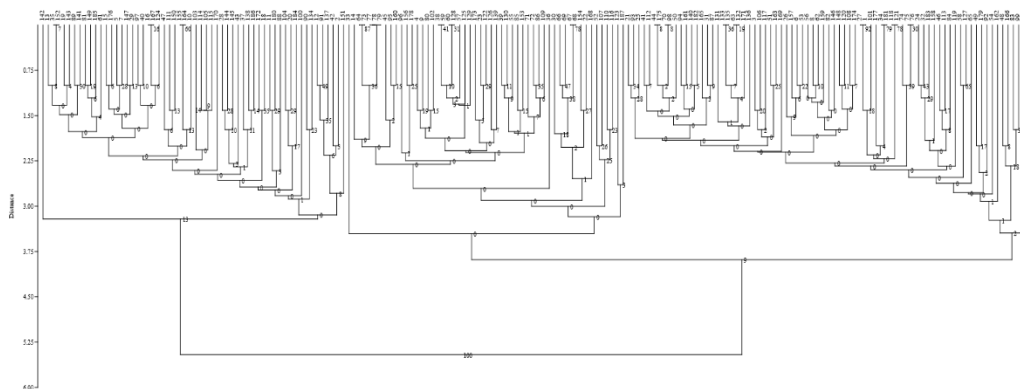
### 3.2 Qualitative Characters

**Table 2.** Eigenvalues and % variance of PC1-15.

PC	Eigenvalue	% variance	Eig. 2.5%	Eig. 97.5%
1	7.203	<b>48.023</b>	43.711	52.536
2	1.773	<b>11.823</b>	10.068	14.371
3	1.416	<b>9.440</b>	7.759	11.022
4	1.243	<b>8.286</b>	6.709	10.449

From 187 accessions or taxa\_S and 12 qualitative characters, the dominance\_D value and Shannon\_H' diversity value were known (Table 4). Grouping was also carried out for all accessions based on the observed qualitative characters (Figure 2).





**Figure 4.** UPGMA with Euclidian Similarity Index of 12 Qualitative Characters.

**Table 3.** Loading/Coefficients Principal Component

Characters	PC 1	PC 2	PC 3	PC 4
FW	<b>0.33965</b>	0.15201	-0.10954	0.10685
FL	<b>0.29865</b>	0.32282	-0.1623	-0.15237
FD	<b>0.32746</b>	0.0038577	-0.026552	0.27733
FT	0.26953	0.12768	0.027917	<b>0.43646</b>
EP	0.16292	0.39086	-0.037651	<b>0.49307</b>
STW	<b>0.31718</b>	-0.23681	0.0061208	-0.177
STL	0.27914	<b>0.29395</b>	-0.24271	-0.2841
STD	<b>0.27984</b>	-0.36029	0.00093119	0.1896
STT	<b>0.23323</b>	-0.31197	0.18481	-0.09107
SW	<b>0.30683</b>	-0.23146	0.062124	-0.22025
SL	<b>0.28517</b>	0.02413	-0.086587	-0.24801
SD	<b>0.2626</b>	-0.34101	0.04836	0.10171
LL	0.14803	0.3595	<b>0.42004</b>	-0.36567
LW	0.10177	0.18111	<b>0.62703</b>	-0.032452
TSS	-0.033183	-0.038622	<b>0.53138</b>	0.20839

**Table 4.** Dominance, and Shannon Index.

Characters	Taxa S	Dominance D	Shannon H'
FS	187	0.003272	4.973
SC	187	0.003267	4.859
PC	187	0.005809	4.966
PT	187	0.002859	4.953
PA	187	0.002627	4.800
SS	187	0.003582	<b>5.033</b>

LA	187	0.002398	4.833
LT	187	0.000995	4.728
LS	187	0.001819	4.776
LM	187	0.002932	4.904
LAS	187	0.003182	4.991
LBS	187	0.002678	4.842

## 4 Discussions

### 4.1 Quantitative Characters

Based on the observation results, the 15 quantitative traits exhibited low to high levels of variability. Similar studies on leaf morphology among several mango accessions have also reported limited diversity. Ethnobotanical and morphological investigations indicate that certain morphotypes are difficult to distinguish solely based on leaf characteristics. Nevertheless, the wide range of variation observed among accessions provides valuable opportunities for selection. Such information is essential for breeders in identifying optimal parental combinations for hybridization and in selecting accessions as potential superior cultivars. The expression of phenotypic traits is determined by genetic factors, environmental influences, and their interactions. A higher coefficient of variation indicates greater diversity and a richer genetic background.

Generative characters, particularly those related to fruit morphology, are more reliable as discriminating variables within the Indonesian mango collection. The effectiveness of fruit-related traits as distinguishing markers among accessions has been previously emphasized by several researchers [6], [9]. The observed high diversity in fruit traits is attributed to genetic composition and genotype–environment interactions. The Indonesian mango germplasm therefore represents an important source of parental material for breeding programs aimed at increasing the edible portion of the fruit. Improvement in flesh proportion can be achieved by reducing the size of the stone and seed through artificial hybridization. However, fruit weight alone does not determine fruit quality, as consumers generally prefer medium-sized fruits.

The cumulative variance explained by the first four principal components (PCs) was close to 80%, which is generally acceptable. These four PCs were selected based on eigenvalues  $\geq 1$ . A principal component with the largest variance, particularly one with significant contribution, should be retained for interpretation. A large variance indicates that the variable has a strong association with the principal component. The selected loading values represent the highest coefficients among the four components, reflecting the variables that most strongly contribute to morphological diversity (indicated in bold).

PC1 included nine traits that exhibited the highest variation among the 15 characters analyzed. As shown in Figure 3, the clustering pattern of data reflects the influence of a single dominant component (PC1). The large angles formed between character vectors in the scatterplot indicate a close relationship among traits in shaping the observed diversity in mango accessions.

### 4.2 Qualitative Characters

The level of diversity of Indonesian mangoes collected in Cukurgondang Pasuruan can be based on calculations of within-community species diversity ( $\alpha$ -diversity). In this case,  $\alpha$

diversity indicates the richness of species at the level of individual community base on qualitative characters. Among all the indices at the level of  $\alpha$ -diversity, Shannon's index presents remarkable characteristics and is used extensively.  $\alpha$ -diversity can be studied essentially for its two most important characteristics: dominance and diversity.

The results of calculating the dominance and diversity values for 12 qualitative characters show that there is not a single qualitative character that dominates the other qualitative characters and each qualitative character has high diversity among all the accessions observed. The  $H'$  value is almost the same for all qualitative characters and slightly greater for the SS character. The greater the  $H'$  value indicates the more homogeneous the data on that character. Thus, the grouping of accessions is influenced by almost all the characters, or almost all the characters are very diverse. Grouping based on qualitative characters has divided Indonesian mango collections into three large groups (distance=3.75). Based on the same data, there are several accessions that are very similar to each other (Fig.4).

The diversity of Indonesian mango germplasm based on quantitative data is more influenced by 9 characters from PC1, namely 48.023%. As for qualitative characters, no characters were found that most influenced the diversity of mango germplasm. However, the 187 accessions can be grouped into 3 main groups based on similar qualitative characters.

**Acknowledgement:** We would like to thank the Ministry of Agriculture for allowing us to observe the Indonesian mango germplasm collection in the Cukurgondan Pasuruan Mango Collection Garden, East Java.

## References

1. Kementerian Pertanian Republik Indonesia, *Statistik Pertanian. Agricultural Statistics 2021*. Pusat Data dan Sistem Informasi Pertanian Kementerian Pertanian Republik Indonesia. Center for Agricultural Data and Information System Ministry of Agriculture Republic of Indonesia., 2021.
2. P. E. Akin-Idowu, U. G. Adebo, K. O. Egbekunle, Y. O. Olagunju, O. I. Aderonmu, and A. O. Aduloju, "Diversity of Mango (*Mangifera Indica* L.) Cultivars Based on Physicochemical, Nutritional, Antioxidant, and Phytochemical Traits in South West Nigeria," *International Journal of Fruit Science*, vol. 20, no. S2, pp. S352–S376, 2020, doi: 10.1080/15538362.2020.1735601.
3. V. R. Lebaka, Y. J. Wee, W. Ye, and M. Korivi, "Nutritional composition and bioactive compounds in three different parts of mango fruit," *International Journal of Environmental Research and Public Health*, vol. 18, no. 2, pp. 1–20, 2021, doi: 10.3390/ijerph18020741.
4. S. E. Bally *et al.*, "The 'Tommy Atkins' mango genome reveals candidate genes for fruit quality," *BMC Plant Biology*, vol. 21, no. 1, pp. 1–18, 2021, doi: 10.1186/s12870-021-02858-1.
5. P. Wang *et al.*, "The genome evolution and domestication of tropical fruit mango," *Genome Biology*, vol. 21, no. 1, pp. 1–17, 2020, doi: 10.1186/s13059-020-01959-8.
6. D. Igbari, G. I. Nodza, A. D. Adeusi, and O. T. Ogundipe, "Morphological characterization of mango (*Mangifera indica* L.) cultivars from south-west Nigeria," *Ife Journal of Science*, vol. 21, no. 1, p. 155, 2019, doi: 10.4314/ijfs.v21i1.13.
7. M. M. Kulkarni, M. M. Burondkar, N. V Dalvi, B. R. Salvi, P. M. Haldankar, and T. Bhattacharyya, "Mango Fruit Size Diversity found in Konkan," vol. III, pp. 43–46, 2019.

8. B. N. Tewodros, K. R. W. Fredah, M. A. Wassu, S. M. Githiri, and O. O. Willis, "Phenotype characterization and diversity assessment of mango (*Mangifera indica* L.) cultivars in Ethiopia," *Journal of Plant Breeding and Crop Science*, vol. 11, no. 2, pp. 55–67, 2019, doi: 10.5897/jpbcs2018.0784.
9. J. S. Shivran, A. K. Singh, D. C. Dimri, and R. Kumar, "Morphological characterization of mango (*Mangifera indica* L.) germplasm under Tarai region of Uttarakhand," vol. 12, no. 2, pp. 2472–2477, 2023.
10. E. O. Ibukun and O. Yomi, "Studies on variation in morphological traits of mango trees (*Mangifera indica*) growing on Kogi State University campus, Anyigba, Kogi State, Nigeria," *GSC Biological and Pharmaceutical Sciences*, vol. 11, no. 1, pp. 113–120, 2020, doi: 10.30574/gscbps.2020.11.1.0092.
11. S. S. Sobieh and R. M. Sayed, "Molecular, morphological, palynological and biochemical characterization of six accessions of two *Mangifera indica* L cultivars (Keitt and Ewais) native to Egypt," *Genetic Resources and Crop Evolution*, no. 0123456789, 2024, doi: 10.1007/s10722-024-01939-y.
12. D. H. Mursyidin, "Genetic diversity and relationship of mango and its wild relatives (*Mangifera* spp.) based on morphological and molecular markers," *Pesquisa Agropecuaria Tropical*, vol. 53, 2023, doi: 10.1590/1983-40632023v5375339.
13. IPGRI, *Descriptors for mango (*Mangifera indica* L.)*, vol. 84. 2006.
14. M. Ulfah, S. N. Fajri, M. Nasir, K. Hamsah, and S. Purnawan, "Diversity, evenness and dominance index reef fish in Krueng Raya Water, Aceh Besar," *IOP Conference Series: Earth and Environmental Science*, vol. 348, no. 1, 2019, doi: 10.1088/1755-1315/348/1/012074.
15. Ø. Hammer, "PAST PAleontological Statistics Reference Manual version 3.20," *Technology*, vol. 1, no. November, pp. 720–766, 2001.