

Integrative Genomics for Mango Genetics and Breeding

Bilal Ahmad^{1, 2}, Ying Su^{1, 2}, Rida Arshad^{1, 2}, Tayyaba Razzaq^{1, 2}, Yi Zhang^{1, 2}, Ting Hou²,
Chaochao Li^{1, 2}, Zhongxin Jin¹, Chengjie Chen¹, Peng Wang¹, Melanie J. Wilkinson⁴, Yibo Bai¹,
Yeyuan Chen⁵, Yu Zhang⁶, Zhiguo Dang¹, Yongfeng Zhou^{1, 2, *}, Xinmin Tian^{3, *}, Jianfeng
Huang^{1, *}

1 State Key Laboratory of Tropical Crop Breeding, Tropical Crops Genetic Resources
Institute, Chinese Academy of Tropical Agricultural Sciences, Haikou, Hainan 571101,
China

2 State Key Laboratory of Tropical Crop Breeding, Shenzhen Branch, Guangdong
Laboratory of Lingnan Modern Agriculture, Key Laboratory of Synthetic Biology,
Ministry of Agriculture and Rural Affairs, Agricultural Genomics Institute at Shenzhen,
Chinese Academy of Agricultural Sciences, Shenzhen, China

3 Key Laboratory of Ecology of Rare and Endangered Species and Environmental Protection
(Ministry of Education) & Guangxi Key Laboratory of Landscape Resources Conservation
and Sustainable Utilization in Lijiang River Basin, Guangxi University Engineering
Research Center of Bioinformation and Genetic Improvement of Specialty Crops, Guangxi
541006, China

4 School of the Environment, The University of Queensland, Brisbane, Qld, 4072, Australia
Australian Research Council Centre of Excellence for Plant Success in Nature and
Agriculture, The University of Queensland, Brisbane, Qld, 4072, Australia

5 Sanya Research Institute of Chinese Academy of Tropical Agricultural Sciences, Sanya
572025, China

6 Guangxi Subtropical Research Institute, Guangxi Academy of Agricultural Sciences,
Nanning 530001, China

© The Author(s) 2025. Published by Oxford University Press. This is an Open Access article
distributed under the terms of the Creative Commons Attribution License
<https://creativecommons.org/licenses/by/4.0/>, which permits unrestricted reuse, distribution,
and reproduction in any medium, provided the original work is properly cited.

* Author for correspondence: zhouyongfeng@caas.cn (Y.Z.);
tianxm333333@foxmail.com (X.T); huangjian1984xy@163.com (J. H)

Abstract

Mango is the second most important tropical fruit crop. Due to ever-changing environmental conditions, world mango production is facing challenges such as diseases (anthracnose and mango malformation), physiological disorders (alternate bearing), low fruit setting, poor fruit quality, short shelf life, and climate change adaptation. Breeding efforts are hindered by the long juvenile period, outdated breeding system, and high heterozygosity, resulting in a slow pace of mango improvement programs. However, over the last decade, significant advances in high-quality genome assemblies, pangenomics, genetic mapping, multi-omics data, and phenomics of large populations have accelerated crop genetics and breeding. Here, we summarize recent progress on the origin and domestication of mango, advancements in genome assemblies, development of genetic maps, functional and comparative genomics, evolutionary insights, and assessments of global phenotypic and genotypic diversity, including species at risk. We also discuss the integration of multi-omics approaches with quantitative genetics for crop improvement. Further, we highlight the key research gaps that limit breeding efficiency and propose integrative strategies combining pangenomics, multi-omics, and machine learning with improved transformation protocols and multi-environment testing to accelerate the development of climate-resilient, high-quality mango cultivars.

Keywords: OMICS; Domestication; Evolutionary history; Breeding; Population Genetics

Introduction

Mango (*Mangifera indica* L.) is a member of the family Anacardiaceae, and the genus *Mangifera*, which includes 69 species, 27 of which produce edible fruits^{1,2}. It is well-known as the “king of fruits” for its attractive aroma, appearance, high nutrition, and delicious taste³. It is widely cultivated in tropical and warmer subtropical regions^{3,4}. It is grown across all continents (excluding Antarctica) and in more than 103 countries. In 2021, global mango production exceeded 51 million metric tons (Mt). India is the leading producer, followed by China. Interestingly, Mexico is 4th largest producer but a leading exporter of mango⁵. Mango is a highly heterozygous and cross-pollinated fruit plant².

Based on the origin, mangoes are classified into two primary cultivar types: Indian and Indochinese. Indian cultivars have a distinct color change when ripe, transitioning to shades of orange or red. They also have rounded shapes, strong-flavored fibrous flesh, and monoembryonic seeds. Unlike Indian cultivars, Indochinese mangoes have polyembryonic seeds, which contain a single zygotic embryo and multiple embryos derived from maternal nucellar tissue¹. This nucellar embryony trait is uncommon among angiosperms; however, it has been reported in three other *Mangifera* species: *M. laurina*, *M. odorata*, and *M. casturi*^{6,7}. The majority of polyembryonic fruits have a green to yellow pericarp, with the Thai mango (*M. siamensis* warbg. ex Craib) being the most common example^{8–10}. The monoembryonic cultivars hold significant commercial importance in Africa, India, Florida, and South America due to their superior fruit quality characteristics. On the other hand, polyembryonic mango cultivars are favored in Australia, Central America, Haiti, Southeast Asia, Hawaii, and South Africa^{11,12}.

The domestication and selection of mango varieties has a rich history spanning thousands of years. Traditional varieties have generally been developed through vegetative propagation by grafting mutant branches. Conventional breeding is not new for mango but became popular after its introduction in the USA, Australia, and China¹³. New fruit cultivars with improved and desirable characteristics have been developed through traditional breeding. However, traditional breeding is unsuitable for developing new desirable mango cultivars due to polyembryony, heterozygosity, long juvenility, more fruit drop (low fruit setting), and long breeding cycles¹⁴. Integrative genomics is required to accelerate mango breeding progress.

Initially, mango was considered an allotetraploid plant with a genome size of 450 Mb^{15,16}. However, several authors^{9,17,18} have reported that genetic markers for mango are inherited in a disomic manner, indicating that mango may be better treated as a diploid (Two complete

sets of chromosomes one from each parent). Mango has $2n = 2x = 40$ chromosomes, suggesting that there are 20 haploid chromosomes and 20 linkage groups¹

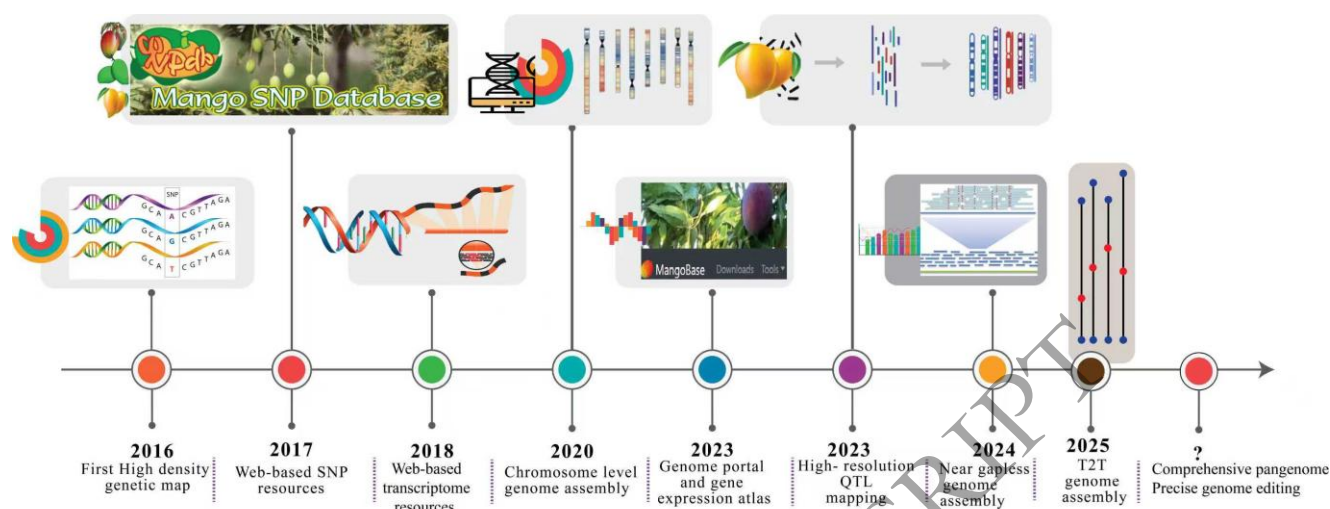


Figure 1. Important developments in mango genomics and future priorities. (Mango SNP Database and MangoBase images are downloaded from relevant websites).

Mango has been genomically understudied and lacks a detailed understanding of the genetic basis of complex agronomic traits. However, significant progress has been made in the last decade in reference genome assembly^{2,19–24}, genetic maps with all 20 linkage groups²⁵, transcriptomes^{26–32}, molecular analyses of germplasm collections^{30,33,34}, gene function³⁵ structural variation identification³⁶ and QTL mapping to identify loci underlying key traits³⁷ (Fig. 1). Recent advancements in genomics, machine learning, and genome editing have led to new concepts and discoveries in crop domestication and breeding. The integration of these innovations will enhance the efficiency of breeding³⁸. Here, we summarize the progress made in mango breeding by genetic mapping, genomics, genetics, and functional genomics and a way forward for future breeding programs.

The origin of the genus *Mangifera*

The genus *Mangifera* belongs to the Anacardiaceae family, which has 73 genera and 850 species. Mostly of tropical origin; however, some have temperate origins. The Malay Peninsula (South of the Kangar Pattani line to the Bismarck archipelago east of New Guinea) contains the highest number of species of the Anacardiaceae family³⁹. The Anacardiaceae family diverged from its sister family, Burseraceae, during the late Cretaceous (~100–66 million years ago) due to continental vicariance, with their most recent common ancestor likely distributed across the Northern Hemisphere. Both families share similar stem age, but the speed of

diversification differs significantly—Anacardiaceae shows a broader range of fruit morphologies and dispersal strategies, and has species that can resist freezing temperatures. This endorses that Anacardiaceae adapted more frequently to shifting climates and niches, likely due to greater intrinsic evolutionary lability or exposure to heterogeneous selective pressures ⁴⁰.

Within Anacardiaceae, the genus *Mangifera* stands out for its ecological and economic significance. Based on the existing diversity in Borneo, the Malay Peninsula, and Sumatra, this region is considered the origin of the genus *Mangifera* ⁴¹, where it is believed to have emerged during the Eocene or possibly the Cretaceous period. The most recent classification of *Mangifera* is based on floral morphology. Accordingly, genus *Mangifera* has two sub-genera—*Mangifera* and *Limus* (Marchand) Kosterm. The evolutionary origins of the genus *Mangifera* and its species remain uncertain, and even cladistic analyses cannot resolve their precise ancestry. It is surmised that two subgenera may have different origins. Out of a total of 69 species, 47 are included in subgenus *Mangifera*, 11 in *Limus* (Marchand), and the position of 11 species is still uncertain. Subgenus *Mangifera* is subdivided into four sections—*Marchandora* Pierre, *Euantherae* Pierre, *Rawa* Kosterm, and *Mangifera* Ding Hou. These sections contain 1, 3, 9, and 34 species, respectively ^{6,41}. Further details on classification, conservation status, and species with important breeding traits are provided in Table 1.

Mango was grown in Indo-Burmese and Southeast Asian regions for an estimated 4000 years and later spread to other parts of the world (Fig. 2) ³⁴ such as America and Africa. The cultivated mango (*M. indica* L.) is thought to have come from Eastern India, particularly in Assam, to Burma or the Malayan region. Mango may have naturally originated in India as a wild fruit ⁴². Indo-Burma was also considered the center of origin ⁴³. The historical records, 60-million-year-old fossil remains of carbonized mango leaves ⁴⁴, distribution of related species and genera, and the existence of multiple varieties and wild species strongly endorse that Indian subcontinent is the geographical origin of *Mangifera*. The fossil remains of leaf imprints established that Assam was the origin of *M. pentandra* ⁴⁵. Some studies report the origin of wild common mango trees in Andaman and Nicobar Islands, Bangladesh, Myanmar, and Northeastern India. Many semi-wild mango germplasm resources are found in the forests of the Indian subcontinent. Northeast India, Borneo, Peninsular Malaya, and Sumatra have the maximum species diversity of the *Mangra* genus. The forests of Borneo, Thailand, the Malay Peninsula, Indo-China, the Indonesian archipelago, and the Philippines house seedling

147 variations with primitive traits i.e. dwarf growth habit and polyembryony, along with a rich
 148 array of species.

149 **Table 1. Taxonomy of Genus *Mangifera* (Kostermans & Bompard, 1993): Edibility, Conservation**
 150 **Status, and Breeding Traits**

Subgenus	Section	Species
<i>Mangifera</i>	<i>Marchandora</i> Pierre	<i>M. gedede</i> Miq (1)
	<i>Euantherae</i> Pierre	<i>M. calneuron</i> Kurz, <i>M. cochinchinensis</i> Engler, <i>M. pentandra</i> Hooker f. (2)
	<i>Rawa</i> Kosterm	<i>M. andamanica</i> King, <i>M. gracilipes</i> , <i>M. griffithii</i> Hooker f., <i>M. merrillii</i> Mukherji*, <i>M. microphylla</i> Griff. ex Hooker f., <i>M. minutifolia</i> Evard*, <i>M. nicobarica</i> Kosterm*, <i>M. paludosa</i> Kosterm*, <i>M. parvifolia</i> Boerl. & Koorders
	<i>Mangifera</i> Ding Hou	<i>M. altissima</i> Blanco (3), <i>M. applanta</i> Kosterm., <i>M. austro-indica</i> Kosterm*, <i>M. austro-yunnanensis</i> Hu, <i>M. casturi</i> Kosterm* (4), <i>M. collina</i> Kosterm*, <i>M. dewildei</i> Kosterm*, <i>M. dongnaiensis</i> Pierre*, <i>M. flava</i> Evard., <i>M. indica</i> L., <i>M. lalijwa</i> Kosterm*, <i>M. laurina</i> Bl. (5), <i>M. linearifolia</i> (Mukherji) Kosterm., <i>M. longipetiolata</i> King, <i>M. magnifica</i> Kochummen, <i>M. minor</i> Bl., <i>M. monandra</i> Merr. (6), <i>M. mucronulata</i> Bl., <i>M. oblongifolia</i> Hooker f., <i>M. orophila</i> Kosterm., <i>M. pedicellata</i> Kosterm., <i>M. pseudo-indica</i> Kosterm., <i>M. quadrifida</i> Jack, <i>M. rigida</i> Bl., <i>M. rubropetala</i> Kosterm., <i>M. rufocostata</i> Kosterm. (7), <i>M. similis</i> Bl. (8), <i>M. sulawesiana</i> Kosterm. **, <i>M. sumbaewanensis</i> Kosterm*, <i>M. sylvatica</i> Roxb., <i>M. swintonioides</i> Kosterm. (7), <i>M. timorensis</i> Bl., <i>M. torquenda</i> Kosterm*, <i>M. zeylanica</i> (Bl.) Hooker f.
<i>Limus</i> (Marchand) Kosterm.		<i>M. blomesteinii</i> Kosterm., <i>M. caesia</i> Jack (9), <i>M. decandra</i> Ding Hou (1), <i>M. foetida</i> Lour. (10), <i>M. kemanga</i> Bl., <i>M. lagenifera</i> Griff., <i>M. leschenaultii</i> Marchand, <i>M. macrocarpa</i> Bl., <i>M. odorata</i> Griff. (11), <i>M. pajang</i> Kosterm. (12), <i>M. superba</i> Hooker f.
Uncertain		<i>M. acutigemma</i> Kosterm., <i>M. bompardii</i> Kosterm., <i>M. bullata</i> Kosterm., <i>M. campospermoides</i> Kosterm*, <i>M. hiemalis</i> Liang Jian Ying, <i>M. maingayi</i> Hooker f., <i>M. persiciformis</i> Wu & Ming, <i>M. subsessifolia</i> Kosterm., <i>M. taipu</i> Buch.-Hamilton, <i>M. transversalis</i> Kosterm., <i>M. utana</i> Utana
Description of Special Traits and Symbol Legend (1) Tolerant to waterlogged conditions and can be used as rootstock. <i>M. reba</i> Pierre, <i>M. camptosperma</i> Pierre, and <i>M. inoarpoides</i> Merr. and Perry are its synonyms. (2) High and consistent fruit yield due to the prevalence of hermaphrodite flowers. (3) Insects tolerant (Seed borers and hoppers). (4) Heavy and regular bearer with small, black, and sweet fruit. (5) Also known as <i>M. longipes</i>, it is immune to anthracnose. (6) Unripe fruit can be eaten (7) Off-season bearing (8) Stone-free fruits (9) The pulp is soft, milky-white, sweet, and fragrant. (10) Less fiber and sweet taste (11) Distinct aroma, firm texture, sweet-sour taste, and can grow in high-rainfall and humid conditions (12) Fruit peel can be easily removed, like a Banana. Species names in bold denote edible species * Endangered species ** Extinct in the wild		

151 Sources. 6,41,46,47

The genus *Mangifera* evolved over a wide geographic range, with the Malay Peninsula considered as the origin's hub and Siam, Burma, Indo-China, and Southeast Asia as the region with the greatest diversity^{8,48}. The Malay Peninsula, Sunda Islands, and Eastern Peninsula have reported 19, 16, and 14 *Mangifera* species respectively, indicating that these regions have the maximum diversity. The noted species include *M. indica* and its close relatives like *M. sylvatica*, *M. odorata*, *M. pentandra*, *M. foetida*, *M. caesia*, and *M. caloneura*. The region between Indo-China and Myanmar is one of the major centers of diversification, due to the presence of many Euantherae (*M. cochinchinensis*, *M. caloneura*, and *M. pentandra*) section species⁴⁹.

Multiple genomic tests confirm that *M. indica* exhibits exceptional genetic diversity compared to many horticultural crops, particularly among Southeast Asian cultivars genetically divergent from other *M. indica* germplasms. Phylogenetic studies have yet to decipher the relationship between *M. indica* and its closely related species, but studies like³⁴ have made significant progress in understanding global mango diversity and domestication history. According to the maximum-likelihood topology, it is suggested that mango cultivars from Southeast Asia diverged earlier than those from other regions worldwide. Previous hypotheses include independent domestication in India and Indochina and initial cultivation in Southeast Asia, with subsequent improvements in India. In a recent study, multiple population genetic analyses were conducted on 208 accessions of *M. indica*⁵⁰. The results revealed the presence of several genetic clusters, with one particular cluster showing significant genetic differentiation. Notably, this genetically differentiated cluster comprised accessions that had been imported from the Southeast Asia region. There are controversies about the center of origin of cultivated mango, some argue for India as a single origin^{51,52} while others proposed Southeast Asia^{1,41}. However, based on recent studies^{19,34,53} it is commonly accepted that mango has two centers of origin: India and Southeast Asia (Fig. 2).

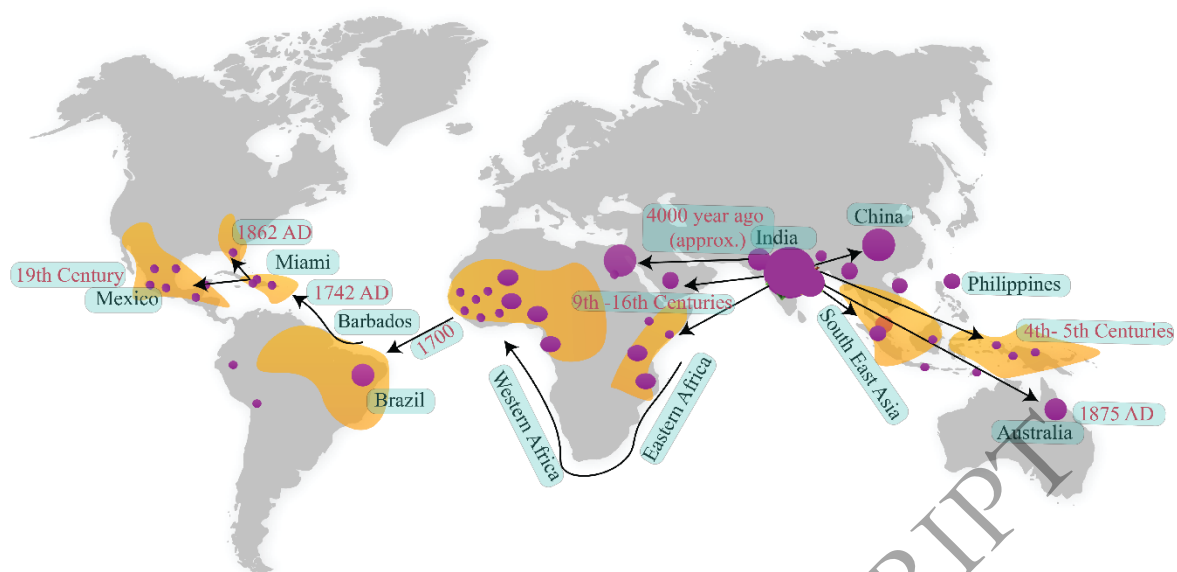


Figure 2. A holistic exploration of the origins and global dispersal patterns of mangoes. Yellow shapes represent regions including multiple countries, while the size of the purple circles indicates germplasm diversity. Adapted and expanded from concepts previously discussed in ⁵⁴.

Apart from the two main centers of origin, mango diversity can also be observed in other parts of the world, such as Africa, Australia, and United States of America (USA). After the first mango arrived on Florida's shores in 1833, it emerged as a major secondary center for mango production, in the late 19th and early 20th century. Monoembryonic Indian cultivar (Malgoa) showed a polyembryonic nature in the USA ¹¹. Early seedling selections from 'Malgoa' were designated 'Haden', and later on world-leading monoembryonic cultivars 'Keitt', 'Kent', and 'Tommy Atkins' were developed from this ⁵⁵. Numerous new Floridian cultivars were introduced and promoted to meet the taste preferences of Western consumers. These cultivars were selectively bred for their appealing red blush coloration, mild flavor, and high fruit weight. Despite the popularity of these cultivars, there remains a demand for further improvements in taste and quality. As such, breeding programs in Australia, South Africa, Israel, Brazil, and China are actively engaged in developing new and improved mango cultivars. These programs are dedicated to creating mango varieties that are even more delicious, visually appealing, and adaptable to a wider range of growing conditions ⁵⁶. Table 2 summarizes the phenotypic diversity of leading mango cultivars, showing variations in fruit morphology (size, shape, skin, pulp color), fruit quality (pulp texture, sweetness, aroma, fiber content), shelf life,

and yield potential across major mango-growing regions. This global diversity displays the crop's adaptability and value for targeted breeding.

Table 2. Global Overview of Phenotypic Diversity in Mango Cultivars

Trait	Phenotypic Diversity	Cultivar(s)	Country/Region
Fruit Size	Small (100-300 g)	Ataulfo Zafiro, Ataulfo Diamante, Alphonso, Dashehari, GuiQI Mango, Yuexi NO.1	Mexico, India, China
	Medium (300-450 g)	Angie, Mallika, Nam Doc Mai, Calypso, Orli, Langra, Guifei, Nan Klang wun	USA, India, Thailand, Israel, Pakistan, China,
	Large (>450 g)	Noa, Keitt, Kent, Omer, Tommy Atkins, Sindhri, Chenab Gold, JinHuang mango, Longjing mango, Hongxiangya	USA, Israel, Pakistan, China
Fruit Shape	Oblong	Ataulfo, Duncan, Harders, Tota Pari, JinHuang mango, Hongxiangya, Chenab Gold	Mexico, USA, India, China, Pakistan
	Elliptic	Kensington Pride, Amrapalli, Zihua	Australia, India, China
	Roundish	Angie, Orli, Agam, R2E2, Renong NO.1, Keitt	USA, Australia, Israel, China
	Ovoid	Kent, Cogshall, Chaunsa, Langra, GuiQI Mango, Tainong NO.1, Siji mango, Sensation, Carabao	USA, India, China, Pakistan, Thailand
	Obovoid	Nam Doc Mai, Edward, Alphonso, Sannian Mango	Thailand, USA, India, Philippines
Pulp Texture	Soft	Ataulfo, Angie, Langra, Dashehari, GuiQI mango, Hongyu, Nan Klang wun, Chenab Gold	Mexico, India, China, Thailand, Pakistan
	Intermediate	Calypso, Kent, Nam Doc Mai, Alphonso, JinHuang mango, Guifei, Dashehari, Sufaid Chaunsa	Australia, USA, Thailand, China, India, Pakistan
	Firm	Keitt, Duncan, Gouveia, Tommy Atkins, Haden, Tainong NO.1, Renong NO.1	USA, Brazil, India, China
Skin Color (ripe)	Green	Kensington Pride, Langra, GuiQi Mango, Hongyu, Rad	Australia, Pakistan, China, Thailand
	Yellow	Ataulfo Zafiro, Nam Doc Mai, Duncan, Mallika, Sindhri, Chenab Gold	Mexico, Thailand, India, Pakistan
	Greenish-Yellow	Totapari, JinHuang mango, Siroli	India, China, Pakistan
	Green with red blush	Tommy Atkins	USA

	Green with purple patches	Cavallini	Israel, Australia
	Dark red	Ah Ping, Agam, Sensation, Irwin, Tommy Atkins	USA, Israel,
	Light Yellow	Favo de Mel, Nang Klang Wan, Baiyu mango	Brazil, Thailand, China
	Golden Yellow	Kent, Keitt, Ataulfo Diamante, Ataulfo Zafiro, Edward, Kensington, Carabao	USA, Mexico, Australia, Philippines
	Yellow orange	Kent, Mallika, Palmer, Edward, Guifei, Hongyu	USA, India, China
Pulp Color	Orange	Madame Francis, Ah Ping, Duncan, Mallika	Canary Islands, USA, India
	Yellow	Tommy Atkins, Ataulfo, JinHuang mango, Sannian mango	USA, Mexico, China,
	Light orange	Kensington Pride, Maha Chinook	Australia, Thailand
	Dark orange	Tali, Mcheso, Dashehari	Israel, Myanmar, China
	Low (10–14.0)	Ataulfo Zafiro, Edward, Renong NO.1, Hongyu	Mexico, USA, China
TSS (°Brix)	Medium (14.1–18.0)	Nam Doc Mai, Kent, Totapari, Guifei	Thailand, USA, India, China
	High (18.1–22.0)	Alphonso, Chaunsa, Angie, Mallika, Favo de Mel, JinHuang mango, Tainong NO.1	India, Pakistan, USA, Brazil, China,
	Very High (>22.0)	Papo de Peru 2, GuiQi Mango, Azeem Chaunsa	Brazil, China, Pakistan
	Mild	Kent, Keitt, Omer, Maha Chinook, Renong NO.1, Nan Klang wun	USA, Israel, China
Aroma	Intermediate	Calypso, Cogshall, Nam Doc Mai, JinHuang Mango, Guifei Hongyu	Australia, USA, China, Thailand
	Strong	Alphonso, Kensington Pride, Mallika, GuiQi Mango, Repin NO.16, Tainong NO.1, Sannian mango	India, Australia, China
	Short (<7 days)	Alphonso, Carabao, GuiQi Mango, Guifei	India, Philippines, China
Shelf Life	Medium (7–14 days)	Nam Doc Mai, Angie, Ataulfo, Sannian mango, Keitt, Alphonso	Thailand, USA, Mexico, India
	Long (>15 days)	Keitt, Tommy Atkins, Kent, JinHuang mango, Tainong NO.1, Renong NO.1, Chenab Gold	USA, China, Pakistan

Yield Potential	High	Tommy Atkins, Keitt, Calypso, Sindhri, GuiQi mango, JinHuang mango, Nan Klang wun	USA, Australia, Pakistan, China, Thailand
	Moderate	Mallika, Nam Doc Mai, Kensington Pride, Hongyu, Sannian mango	India, Thailand, Australia, China
	Absent	Ataulfo, Alphonso, Nan Klang wun, Nam Doc Mai	Mexico, India, Thailand
Fiber Quantity	Low	Kensington Pride, Maha Chinook, Keitt, Mallika, Chaunsa, Dasher, JinHuang mango	USA, Australia, India, Pakistan, China
	Intermediate	Cogshall, Kent, 'Papo de Peru 2, Madame Francis, GuiQi mango, Tainong NO.1	USA, Brazil, Canary Islands, China
	High	Totapuri, Sannian mango, Dongpo mango	India, China

Sources. ^{57–63}

Domestication

The domestication history of fruit crops has played a vital role in shaping their characteristics and availability ⁶⁴. Central to this process is the concept of domestication syndrome—a suite of heritable traits that distinguish cultivated plants from their wild progenitors, arising from human selection ⁶⁵. Historical records suggest that mango trees initially have small, thin-fleshed fruits, while centuries of human selection led to larger fruits with more edible pulp ⁴⁹. These morphological changes mark some of the earliest domestication traits. Modern breeding programs have expanded the domestication syndrome to include traits such as reduced tree vigor, regular and high yields, disease resistance, improved transportability, and extended shelf life, aligning selection goals with both production efficiency and consumer preferences ^{14,66}. However, studies into the primary bottleneck related to mango's initial domestication are hindered by the inaccessibility of wild *M. indica* populations, and the history of human-mediated migration into new areas allows for investigating whether secondary genetic bottlenecks occurred during successive

founder events. Although many studies have shed light on the molecular diversity and genetic structure of mango cultivars within specific regions, including Brazil ⁶⁷, China ⁶⁸, Kenya ⁶⁹, Myanmar ⁷⁰, Colombia ⁷¹, Iran ⁷², and particularly India ^{73–75}, only a few studies have investigated mango cultivars across a wide geographical range.

Studies performed in the 2000's based on the microsatellite genetic markers found Southeast Asian mango cultivars were more differentiated from other populations ^{18,30}. Some studies revealed that the domestication history of mango is more complex than previously

thought. It proposes two trends observed in perennial fruit crop domestication: multiple domestications and hybridization with related species. These trends are common in perennial crops and occur over a broader geographical scale and a longer period compared to annual species. This indicates the presence of two cultivated gene pools in mango, contributing to regions of increased diversity outside the presumed centers of origin. Furthermore, it suggests that not all the genetic variation present in current mangoes can be traced back to India ³⁴.

Over the past decade, numerous national and international initiatives have been dedicated to large-scale genotyping of plant genetic resources. Table 3 presents a compilation of studies conducted in different countries, aiming to investigate the genetic resources available in those regions or comparison with genetic resources available in other countries. Overall, SSRs are marker types widely used for examining levels of diversity and polymorphism ⁷⁶. The SNP genotyping arrays are very helpful for pedigree identification and have been developed in many horticultural crops, such as cherry, peach, grape, avocado, apple, potato, rose, walnut, pear, and strawberry ^{77–80}. SNP arrays have not been specifically designed for parentage analysis and have been limited to a few SSR markers, and due to this many popular open-pollinated mango cultivars either have a single known parent or lack parentage information entirely. In recent years, the availability of diverse genomic resources, SNP genotyping arrays, and cost reductions in sequencing have made the genotyping of open-pollinated seedlings more attractive for mango breeding programs.

Table 3. Details of studies performed using various techniques for genomic diversity analysis in different parts of the world.

Country	Technique	Total Cvs and their origin	Mango and their origin	Conclusion	References
Australia	SSR Markers	254 accessions of <i>Mangifera indica</i> L. and related <i>Mangifera</i> species from 12 diverse geographic areas		133 alleles were detected using 11 SSR markers. The accessions were divided into 4 major nodes, indicating geographical origins. 24 unique genotypes were recognized for 50 trees previously assigned different accession names. No relationship was detected between SSR markers and embryony.	³⁰

Bangladesh	Multivariate analysis	23 local genotypes	The analysis divided the genotypes into five clusters, with 15 confined to just two.	81
Brazil	RAPDS	35 accessions (Brazil), 6 (USA), 1 (India)	The unweighted pair group method utilizing arithmetic average cluster analysis identified 55 polymorphic loci and suggested 5 genotypic categories.	82
China	SCoT markers	168 mango accessions (52 Cvs introduced from different countries i.e., USA, Australia, and 116 indigenous accessions present at Guangxi University)	45 primers generated 337 fragments, and 244 (72.4%) exhibited polymorphism. The dendrogram divided accessions into 2 clusters with 89% of genetic diversity occurring within populations versus 11% among populations.	83
China	WGS	224 mango accessions (wild and cultivars) from 22 countries	<i>M. himalis</i> J.Y. Liang was reclassified into <i>M. indica</i> . Population structure and differentiation analysis divided Chinese accessions into 2 distinct pools.	53
Egypt	SSR and EST-SSR	28 Mango Cvs (4 native Egyptian CV & 24 being grown in Egypt)	Genetic polymorphism and morphological traits both showed significant variation, ranging from 0.71% to 100%.	84
India	Multiple markers (65 RAPD, 25 ISSR, 23 DAMD, 22 SCoT, 33 CBDP, and 40 SSR)	70 selected Indian mango genotypes based on geographic and fruit status.	The highest genetic diversity was seen in the East Indian populations. The indigenous genotypes showed more potential for exploiting the unique and favorable alleles. The maximum genetic variability was detected with SSRs.	76
Indonesia	Morphology and RAPD markers	82 native Cvs	Morphological studies divided all Cvs into 3 main groups i.e., Kidang Kencono, Kopek, and Karang. RAPD suggested a 2 to 31% range of genetic variability, and 9 clusters having 75% similarity were formed.	85
Japan	SSR Markers	120 mango genetic resources available in the country (imported from	46 polymorphic SSR markers were used, and 10 sets of 3 SSR markers distinguished 83 genotypes, excluding synonymous and identical accessions. The analysis revealed three clusters, showing genetic differentiation	10

		17 different countries)	between Indian and Southeast Asian varieties, with Floridian varieties being closer to the Indian variety.	
Mexico	AFLP	41 local mango Cvs	High levels (> 84 %) of polymorphisms were detected among accessions. Based on AMOVA analysis, significant genetic differences were detected among and within mango accessions. Genetic differences were found to be related to their geographical origins.	86
Nigeria	SSR markers	17 selected mango Cvs present in Nigeria	The seven polymorphic primers identified a total of 21 alleles, with each locus containing 2 to 5 alleles and an average of 3.0 alleles per locus.	63
Pakistan	SSR Markers	15 Indigenous Cvs	Utilizing 12 SSR primers, the accessions showed coefficients of similarity ranging from 75 to 100%, indicating inbreeding and a few parents.	87
Thailand	SSR anchored primers	24 mango Cvs (13, 4, 3, 1, and Thai, Florida, Indian, Indonesian, and Philippine, respectively)	40 primers were used, and 7 primers produced polymorphic DNA amplification patterns. Two Thai Cvs (Nong Saeng and Nang Klangwan) showed more genetic diversity than other Cvs, and based on the dendrogram, 11 closely related Thai Cvs were divided into 3 groups.	88
USA, Senegal, Thailand, Australia	Simplified visual analysis plus SNP markers	Different <i>Mangifera</i> species, accessions from half-sibling populations, and 2232 accessions	272 SNP markers identified over 520,000 Genotypes. The affinity propagation method divided genotypes into 258 groups. A simple visual method suggested that no more than 30 SNPs are required for the accurate identification of a CV.	33
USA	Microsatellite Markers	208 accessions of mango (203, <i>M. indica</i> : 2, <i>M. griffithii</i> Hook f.: and 3, <i>M. odorata</i> Griff.)	For the <i>M. indica</i> population, 25 microsatellite loci displayed an average of 6.96 alleles per locus and an average polymorphism information content score of 0.552. Estimates put the collection's overall propagation error at 6.13%. The Florida Cvs were more closely related to Southeast Asian Cvs than to those from India.	18
Vietnam	RAPD and ISSR markers	10 mango samples from different regions	Using 10 ISSR and 10 RAPD markers high polymorphic levels were detected; ISSR (98.1%) and RAPD (91.9%), and 145 and 162 loci were generated, respectively.	89

249 **Abbreviations:** Cvs (Cultivars), SSRs (Simple Sequence Repeats), RAPDS (Random
250 Amplified Polymorphic DNA), SCoT (Start Codon Targeted), EST-SSR (Expressed Sequence
251 Tag-Simple Sequence Repeat), DAMD (Directed Amplification of Minisatellite-region DNA),

ISSR (Inter Simple Sequence Repeat), CDBP (CAAT box- derived polymorphism) AFLP (Amplified Fragment Length Polymorphism), SNP (Single Nucleotide Polymorphism), WGS (Whole-Genome re-Sequencing).

Progress and advances in genomics-based methods

Mango Genomic Databases

Over the last decade, many mango genomes have been sequenced and several SSR and SNP markers have been produced. Mango genomic databases have been created to effectively store this growing genomic data. In 2017, the first SNP database for mango, known as the *Mangifera indica* SNP database (MiSNPDb; <http://webtom.cabgrid.res.in/mangosnps/>), was developed⁹⁰. This database provides information about 1.25 million SNPs discovered from 84 distinct mango cultivars using ddRAD-Seq approach. In 2018, Qamar-ul-Islam and his team developed a cutting-edge Mango Genomic database called MGdb after a comparative transcriptome analysis of four mango cultivars, namely ‘Langra’ (Pakistan), ‘Shelly’ (Israel), ‘Zill’ (China), and ‘Kent’ (Mexico). The database includes transcriptomic resources (e.g., 30,000–85,000 unigenes) that,⁹¹ while limited compared to modern telomere-to-telomere (T2T) assemblies, offer foundational insights into mango’s genetic diversity. In 2023, a comprehensive resource for mango genomics called MangoBase (<https://mangobase.org/>) was established⁹². This platform is solely dedicated to mango genetics and provides an array of interactive bioinformatics tools, sequences, and annotations to analyze, visualize, and retrieve omics data. One of its notable features is a gene expression atlas that comprises 12 datasets and 80 experiments analyzed based on the ‘Tommy Atkins genome’ reference, which is linked to their gene references and annotations. Overall, these powerful tools will help researchers gain a deeper understanding of the genetic makeup of mangoes and explore new possibilities for developing improved cultivars with desirable traits.

Genome assemblies and comparative genomics

Plant genomes provide evidence of evolutionary history, whole-genome duplication, population processes, and other factors⁹³. Early attempts to assemble the mango genome were hindered by its high heterozygosity (~2.5%)⁹⁴. However, current sequencing and assembling technologies are capable of assembling highly heterozygous genomes into two separate haplotypes. The first attempt to assemble the mango cultivar ‘Amrapali’ genome was made using Illumina MiSeq overlapping paired-end reads (100-150 bp). However, the assembly size of 492 Mb in 211,141 contigs exceeded the estimated size of the mango genome (439 Mb),

suggesting redundancies in the assembled contigs⁹⁵. Later, the raised issues were resolved by utilizing PacBio single-molecule real-time (SMRT) long sequence reads for genome assembly and allowed for long overlaps of 500 bp and a high mismatch rate of 15% to accommodate the heterozygosity⁵⁴. This approach resulted in the first draft genome assembly of 323 Mb for the mango cultivar ‘Amrapali’ (NCBI: LMWC000000000 v.1). Later in the same year, the draft genome assembly (407 Mb) of the prominent Australian mango cultivar ‘Kensington Pride’ was developed based on Illumina short-read sequencing⁹⁶, the issue of heterozygosity was not addressed, and contig redundancy remained a concern.

Due to the recent advancements in sequencing techniques and bioinformatics tools, significant advancements have been made in the complete genome sequencing of mango cultivars. The complete genome sequencing of the ‘Hong Xiang Ya’ cultivar resulted in a high-quality genome assembly, comprising 371.6 Mb of genomic sequence (Table 4). Utilizing the available genetic map, the genome was assembled onto its respective chromosomes, achieving a remarkable anchoring of approximately 98.77% of the genome assembly onto 20 pseudo-chromosomes²². Further, a better chromosome-scale genome assembly of the mango cultivar ‘Alphonso’, was successfully generated and is being used as a reference genome¹⁹. Moreover, the 10× genomics long-read sequencing was used for generating a high-quality genome assembly of ‘Tommy Atkins’. The assembly included ~86% of the ~ 439 Mb haploid mango genome. ~ 3.3 M SNPs were identified using the mapping population of ‘Tommy Atkins’ and ‘Kensington Pride’, and 28 potential genes associated with fruit size were identified². Wijesundara et al., (2024)²¹ generated a good quality, near gapless chromosome-level (365 Mb) assembly of ‘Irwin’ with only high coverage (204×) HiFi. Recently, a gapless, chromosome-level genome assembly (377.6 Mb) of the mango landrace ‘San Nian Mang’, consisting of a single contig per chromosome, was reported²⁴. In addition, a haplotype-resolved, telomere-to-telomere (T2T) genome assembly of the cultivar ‘Alphonso’ has been generated using a combination of PacBio HiFi, Oxford Nanopore (ONT), and Hi-C sequencing data²³, providing the high-quality mango reference genome to date.

Table 4. Comparative statistics of available mango genomes.

Mango Cv.	Amra pali	Hong Xiang Ya	Alpho nso	Irwin	Tomm y Atkins	Irwin Hap1	Irwin Hap2	Irwin Collapsed	Alphon so	Alphon so
--------------	--------------	---------------------	--------------	-------	---------------------	---------------	---------------	--------------------	--------------	--------------

										ALT2T _HAP1	ALT2T _HAP2
Sequencing Technology	PacBio o, Illumi na, BioNa no	PacBio o CLR+ Illumi na	PacBi o		Illumi na HiSeq	PacBio HiFi	PacBio, HiFi	PacBio HiFi	PacBio HiFi HiC, ONT	PacBio HiFi	
			Sequel II + HiSeq 2000 and MiSeq + HiC	PacBi o							
Assembly method	FALC ON (v. 0.2.2)	Falco n versio n (v. 0.3.0)	Canu (v. 1.8)	MEC AT (v. 1.2); Polis h (v. 1.22)	DeNo vo Magic (v. 3.0)	HiFiasm (v.1)	HiFiasm (v.1)	HiFiasm (v.1)	HiFias m (v. 0.16), Juicer (v. 1.5), 3D- DNA (v.180 922)	HiFias m (v. 0.16), Juicer (v. 1.5), 3D- DNA (v.180 922)	
Genome coverage (x)	289.6	388	240	61.96	180	204	204	204	Hifi:88 .4 ONT:1 49.4	Hifi:88 .12 ONT:1 48.88	
Genome size (including unplaced scaffolds) (Mb)	411.3	371.6 2	392	396	374.8	-	-	-	-	-	
Assembly Size (Mb)	290	367	357	375	328	354	354	365	346.22	347.59	
Assembly level	Chro moso me	Chro moso me	Chro moso me	Chro moso me	Chrom osome	Chromos ome	Chromos ome	Chromos ome	Chrom osome	Chrom osome	
Number of Contigs	3,714	120	420	3022	17,187	4711	1515	4642	397	70	
Contig N50 (Mb)	0.955	4.82	3.50	1.03	0.040	13.21	15.45	14.98	17.03	17.96	

Number of scaffolds	2314	-	252	1,305	2,565	23	23	23	390	65
N50 scaffold (Mb)	12.1	18.78	17.6	1.03	16.2	—	-	-	18.06	19.25
Telomers	-	-	-	-	-	37	37	40	37	36
BUSCO% (Embryophyta_odb10)	97.4	93.3	95.9	—	94.6	98.1	99	99.2	99.1	99.1
No. of predicted genes	46,395	34,529	41,251	36,756	26,616	34,659	33,230	35,220	41,423	40,794
Accession number	PRJNA300605	PRJCA002248	PRJNA487154	CRA004336	PRJCA005296	GWHESEJ000000	GWHESEJ000000	GWHESEJ000000	PRJNA1218505	PRJNA1218506
Public source	NCBI	NCBI	NCBI	NCBI	NCBI	NCBI	NCBI	NCBI	NCBI	NCBI
Reference	94	22	19	20	2	21	21	21	23	23

Abbreviations: BUSCO (Benchmarking Universal Single-Copy Orthologs), HiFiasm (Highly Efficient Assembler for Long-Reads Sequencing Data), ONT (Oxford Nanopore Technologies), NCBI (National Center for Biotechnology Information), NGDC (National Genomics Data Center)

Due to the wide genetic diversity in *M. indica*, single or multiple reference genomes may not be enough for the in-depth genetic diversity analysis of the species, for this a pangenome is required. A pangenome refers to a detailed collection of genomes from a species, achieved by comparing numerous genomes. The construction of a pangenome facilitates the determination of various essential aspects, such as the size of a core genome (the conserved

portion among related genomes), the overall size of the pangenome, and the extent and characteristics of variations within a species or genus ⁹⁸. This in-depth understanding contributes to advancements in our knowledge of species/genus evolution and aids in studying agronomic traits. Recently, a draft pangenome of mango was constructed using three cultivars ⁹⁹ but not of high quality. Identification of disease-resistance alleles in mango cultivars and their wild relatives is important for improving mango crop quality. This can be achieved by constructing a pangenome using wild and cultivated accessions. In the future, more T2T genomes and a comprehensive pangenome of the genus *Mangifera* are needed for more insights into mango genomics and breeding.

With the accessibility of complete genome sequencing data, opportunities for comparative genome analysis are opening up to uncover gene networks, transposable elements, intron-exon boundaries, novel biological processes, molecular markers associated with economically significant traits for breeding purposes, and structural variants ^{36,100–103}. Cortaga et al. (2022) ³⁶ sequenced the whole genomes of three mango species native to the Philippines (Carabao, Huani, and Paho). Genome-wide variants were identified by mapping 93–95% of the quality-filtered reads with sequence coverage of 2.91× to 4.30× to the Alphonso and Tommy Atkins mango genomes. They identified over 2.3 million SNPs, over 199 thousand insertions, and over 199 thousand deletions in each species. To identify genomic and structural variants, few studies have been performed, and more studies are thus needed in the future.

Besides, the mitochondrial and chloroplast genomes also provide valuable information for understanding the evolutionary history of fruit plants ^{104,105}. Recently, phylogenetic relationships within the genus *Mangifera* were determined using whole chloroplast and nuclear genome sequences ¹⁰⁶. Thirteen samples representing eleven *Mangifera* species were sequenced, revealing that the chloroplast genomes of *M. altissima*, *M. caloneura*, *M. applanata*, and *M. laljiwa* were nearly identical to that of *M. indica* (99.9% sequence similarity). The close genetic relationship indicates a common ancestry and potential cross-hybridization between *Mangifera* wild relatives and *M. indica*.

Genetic basis of complex agronomic traits

Plant breeding traditionally relies on phenotypic selection but can be time-consuming, especially when dealing with fruit trees. The most time-consuming phase in this process is the transition from the juvenile stage to a fruit-bearing tree, which typically takes several years in mango trees ^{8,107,108}. By using molecular genomic tools, there is an opportunity to assess the

genetic diversity of prospective parent plants, identify markers linked to significant horticultural traits, and enhance the speed of mango breeding programs ⁶⁶.

Early selection and use of molecular markers through marker-assisted selection (MAS) can significantly streamline the breeding process. A high-density genetic map is required for quantitative trait loci (QTL) mapping. Kashkush et al. (2001) ¹⁰⁹ made an initial attempt to generate a genetic map of mangoes. Subsequently, attempts were made by other scientists ^{110,111} for improvements, and a high-resolution genetic map with the potential for future use was generated in 2016 ¹¹². A comparison of the average distance between this map and others demonstrated the possibility of constructing a high-density genetic map capable of fine QTL mapping and MAS in mango. Later on, Kuhn et al. (2017) ⁶⁶ developed a high-density genetic map for mango, allowing for the exploration of significant associations between traits and SNP markers (Table 5).

Table 5. Details of available genetic Maps in Mango

Genotyping Technology	Population size	Number of markers	Total Genetic distance (cM)	Distance between markers (cM)	Number of LGs	Linked QTLs	References
RFLP and AFLP	31	847	1437.7	10.4	-	-	113
AFLP	29	34	NA	4.75	-	-	109
AFLP	60	81	354.1	4.37	-	-	111
Microsatellite and RFLP	31	76	529.9	6.97	-	-	114
SLAF-seq	173	6594	3148.28	0.48	20	-	112
SNP markers	775	726	2890.6	4.13	20	Polyembryony, branch habit, bloom, ground skin color, blush intensity, beak shape, pulp color	66
SNP Markers	94	4,361	2,982.75	0.68	20	Fruit color and firmness	115

Abbreviations: cM (centiMorgan— unit of genetic linkage), RFLP (Restriction Fragment Length Polymorphism), AFLP (Amplified Fragment Length Polymorphism), SLAF-seq (Specific-Locus Amplified Fragment Sequencing), SNP (Single Nucleotide Polymorphism) LGs (Linkage Groups), QTLs (Quantitative Trait Loci).

Some studies have highlighted the important role of genetic structure in trait mapping accuracy and breeding efficiency. Wilkinson et al. (2022)⁵⁰ identified four distinct genetic clusters in the Australian mango gene pool, with the most differentiated group comprising Southeast Asian accessions. This population structure was strongly linked with key traits like trunk circumference and fruit blush color, demonstrating how shared ancestry can create spurious genotype-phenotype linkages. For accurate QTL mapping and genomic selection, it is therefore essential to account for population structure to avoid false marker-trait associations. These findings underscore that breeding programs must characterize genetic backgrounds (e.g., the unique Southeast Asian gene pool) to improve predictive accuracy for complex traits like fruit quality. Fruit size, firmness, aroma, and color are important fruit quality traits that need to be examined at the genomic level for cultivar improvement². Fruit quality-related traits are quantitative traits and are influenced by multiple genes. Transcriptomic studies^{116,117} have identified some genes potentially involved in fruit color, flavor, ripening duration, and storage period. However, the lack of functional validations suggests that the genetic basis of these traits is largely unknown. Recently, Srivastav et al. (2023)¹¹⁸ constructed two high-density linkage maps based on segregating female and male parents and performed high-resolution QTL mapping. They reported that QTLs for fruit peel color are located at Chr 3 (99.63 cM) and 18, and pulp firmness on Chr 11 (94.82 to 95.19 cM) and 20.

Flowering at a favorable time gives a desirable crop yield, while the flowering mechanism in mango is very complex and challenging for growers, physiologists, and breeders¹¹⁹. Like many other perennial fruit crops, mango tends to be alternate (Heavy fruit in one year (ON) and less fruit (OFF)) or irregular bearing. This phenomenon poses severe losses to farmers and is also undesirable from a breeding point of view. Alternate bearing is controlled by environmental, genetic, physiological, and biochemical factors¹²⁰. In mango exploring the genetic basis of alternate bearing is difficult due to some inherent factors influencing the ability to collect the fruit data required to assess alternate bearing, long juvenility period, and heavy fruit drop, etc.¹²¹. Nakagawa et al. (2012)¹²² identified the FLOWERING LOCUS T-like gene and genes related to gibberellin metabolism in biennial-bearing mango trees. *FT*, *API*, and *LFY* genes showed up-regulation in mango leaves during flower induction. Twenty-six genes were differentially expressed (DEGs) in Neelum (regular) and Dashehari (Irregular) cultivars, suggesting their roles in the alternate bearing. The DEGs belonged to hormone synthesis, alternate bearing, and carbohydrate metabolism pathways. Further, three genes (SPL-like gene, Rumani GA-20-oxidase-like gene, and LOC103420644) showed significant changes in

expression patterns in both cultivars, suggesting their potential roles in alternate bearing. These findings provide initial information about the genetic basis of alternate bearing, and we surmise that multiple pathways are involved in alternate bearing¹²⁰. Some other flowering-related genes (Table 6) have been overexpressed in Arabidopsis for functional studies. A few sporadic type studies have tried to explore the genetic basis of alternate flower bearing in mango, but detailed studies are lacking.

Table 6. Identification or functional analysis of flower, fruit, and stress-related candidate genes.

Trait/Role	Cultivar	Type of study	Genes identified/characterized	Reference
Flowering and malformation	Amrapali	Transcriptomic and qRT-PCR Analysis	<i>GA20OX3</i> , <i>AGL24</i> and <i>LDL2</i>	123
Alternate bearing/Flowering	Neelum and Dashehari	Comparative transcriptomic and qRT-PCR Analysis	<i>SPL-like gene</i> , <i>Rumani</i> <i>GA-20-oxidase-like gene</i> , and <i>LOC103420644</i>	120
Alternate bearing	Irwin	Comparative expression analysis	<i>MiFT</i>	122
Flowering/Alternate bearing	Dashehari	Comparative transcriptomic analysis	Identified 85 DEGS in bearing and non-bearing tissues	124
Flowering	SiJiMi	Overexpression in Arabidopsis	<i>MiTFL-1</i> , <i>MiTFL1-2</i> , <i>MiTFL1-3</i> and <i>MiTFL1-4</i>	125
Flowering	SiJiMi	Overexpression in Arabidopsis	<i>MiLFY</i>	126
Flowering	SiJiMi	Overexpression in Arabidopsis	<i>MiCOL1A</i> and <i>MiCOL1B</i>	127
Flowering	SiJiMi	Overexpression in Arabidopsis	<i>MiFT1</i> , <i>MiFT2</i> , <i>MiFT3</i>	128
Flowering	SiJiMi	Overexpression in Arabidopsis	<i>MiSVP1</i> and <i>MiSVP2</i>	129
Flowering	SiJiMi	Overexpression in Arabidopsis	<i>MiAPI-1</i> and <i>MiAPI-2</i>	35
Flowering	SiJiMi	Overexpression in Arabidopsis	<i>MiCOL16A</i> and <i>MiCOL16B</i>	130

Fruit color	Amrapali	Transcriptomic and phylogenetic analysis	<i>MiCHS</i> , <i>MiCHI</i> , and <i>MiF3H</i>	131
Fruit development and ripening	Kaituk', 'Nam Dok Mai No.4', and 'Nam Dok Mai Sithong'	Content and qRT-PCR analysis	<i>MiPSY</i> , <i>MiPDS</i> , <i>MiZDS</i> , <i>MiCRTISO</i> , <i>MiLCYb</i> , <i>MiLCYe</i> , <i>MiHYb</i> , <i>MiZEP</i> , <i>MiCCD1</i> , <i>MiNCED2</i> , and <i>MiNCED3</i>	132
Fruit ripening and softening	Alphonso	Transcriptomic and qRT-PCR Analysis	38 genes of metabolic pathways	116
Fruit ripening and softening	Dashehari	Comparative Transcriptomic and qRT-PCR Analysis	SAM synthase, ACC synthase, ACC oxidase, <i>ETR</i> , <i>EIN3</i> , endo beta 1–4 glucanase, expansin, pectate lyase, and polygalacturonase	133
Fruit	Dashehari	Expression analysis	<i>MiExpA1</i>	134
Fruit ripening	Kent	Transcriptomic and qRT-PCR Analysis	<i>ACO</i> , <i>ETR1</i> , <i>ERS1</i> , <i>EIN4</i> , <i>EXP</i> , <i>AGAL</i> , <i>LCYB</i> , and <i>CHYB2</i>	27
Disease Resistance (<i>Colletotrichum Gloeosporioides</i>)	Zill	Transcriptomic and qRT-PCR Analysis	<i>ERFs</i> , <i>NBS-LRR</i> , <i>NPR1</i> , and <i>PR</i>	135
Salt and drought tolerance	SiJiMi	Overexpression in Arabidopsis	<i>MiPRIA</i>	136
Salinity and drought	Jin Huang	Overexpression in Arabidopsis	<i>MiCOL9A</i> and <i>MiCOL9B</i>	137
Drought	SiJiMi	Overexpression in Arabidopsis	<i>MiCOL1A</i> and <i>MiCOL1B</i>	138
Drought	SiJiMi	Overexpression in Arabidopsis	<i>MiCOL16A</i> and <i>MiCOL16B</i>	130

415 **Abbreviations:** qRT-PCR (Quantitative Reverse Time-Polymerase Chain Reaction), DEGS
416 (Differentially Expressed Genes).

417

418 Mango malformation and anthracnose are the most devastating diseases of mango
419 worldwide. The malformation is caused by different species of the fungus *Fusarium*, including
420 *F. mangiferae*, *F. mexicanum*, and *F. sterilihyphosum*. The susceptibility of different mango

cultivars to mango malformation disease varies significantly, with conflicting reports and no experimental evidence to support them. For example, one source reported the 'Ewais' cultivar as moderately susceptible, but showed low and high susceptibility in other references ¹³⁹. Anthracnose appears in field and storage (pre and postharvest stages) and causes a significant reduction in yield and fruit quality. The relationship between mango and *Colletotrichum* species (causative agent of anthracnose) is still unknown. Few mango cultivars and *Mangifera* species show resistance against anthracnose ^{140,141}. No studies have been performed to identify the novel resistance genes in these genotypes for pyramiding genetic resistance in mangoes. Hong et al. (2016) ¹³⁵ investigated the effect of borate on postharvest *C. gloeosporioides* infection and reported some defense-related genes having a role in the anthracnose resistance mechanism. However, functional validation of the genes has not been performed. β -1,3-GLU2 has critical roles in the primary defense response of mango against anthracnose. AS-PCR assays and phenotyping suggested that two alleles of SNP 21881933 are linked with anthracnose resistance ¹⁴². Studies clarifying the genetic basis of mango malformation and anthracnose are not available.

Marker Selection using GWAS

Genome-wide association studies (GWAS) have been recognized as a promising approach for trait association studies in various fruit crops, including mango. GWAS offers a valuable avenue to regenerate and revise the genetic understanding of mango, enabling researchers to identify key genomic regions associated with desirable traits. By harnessing the power of GWAS, breeders can expedite the development of improved mango varieties by targeting specific genomic markers linked to traits of interest. GWAS studies in fruit crops can establish associations between genotypes and phenotypes by examining and testing variations in allelic frequencies of genetic variants. This approach enables researchers to assess the differences in allele distribution and their potential impact on observable traits ^{143,144}. GWAS has greatly contributed to the identification of robust associations for various traits and diseases of interest ¹⁴⁴. This powerful technique has been instrumental in uncovering the underlying biology of phenotypes, estimating their heritability, and calculating genetic correlations, among other valuable insights. For example, GWAS analysis has been utilized to map the short chilling requirement, flowering time, and diverse fruit quality traits in peaches, leading to the identification of candidate genes associated with these traits ¹⁴⁵. Moreover, candidate genes regulating fruit ripening have been identified through GWAS in apples. Ma et al. (2024) ⁵³ used genome re-sequencing data from 224 mango accessions and performed GWAS analysis for

flowering ability, volatile compounds related to aroma, and fruit weight. Twenty-six loci linked with these parameters were identified, and 422 genes linked with identified loci were also predicted. These methods have been applied to investigate numerous traits in various fruit crops. However, there are fewer studies on mango, as it requires many samples to make associations, and difficult to obtain values in mango. Recently, GWAS analysis has been extended to include metabolomics data, known as mGWAS, which provides a more comprehensive understanding of metabolite biosynthesis. Several recent reports have demonstrated the successful application of these techniques in perennial crops. Consequently, the potential of mGWAS in studying mango crop populations appears promising and should be further explored.

Transcriptomics-Based Insights into Mango Breeding

In the last decade, significant advancements have been made in mango breeding, such as the swift selection of parent plants and their offspring, streamlining breeding programs, and accurately assessing the genetic relationships between different mango cultivars. The implementation of modern plant breeding techniques has revolutionized mango breeding and paved the way for remarkable progress in the field¹⁴⁶. Significant progress has recently been made in understanding the mechanisms behind the softening and ripening process in mango fruit cultivars. However, there is still a lack of knowledge, and advancements in omics-based biotechnological approaches, including genomics, transcriptomics, metabolomics, and proteomics, can play a crucial role in deciphering the ripening phenomena of mango fruit and understanding its postharvest physiology.

The availability of extensive transcriptomic data for different mango cultivars, developmental stages, tissues, and treatments has greatly facilitated the investigation of functional genes. Transcriptomic studies have reported genes and proteins involved in processes and pathways such as spongy tissue development¹⁴⁷, fruit ripening¹⁴⁸, aroma production¹⁴⁹, and disease resistance¹⁵⁰. For example, the RNA-Seq analysis of red and green fruit following *C. gloeosporioides* infection on their respective red and green sides revealed that resistance in the red fruit is linked to the upregulation of the brassinosteroid, ethylene, and phenylpropanoid pathways¹⁵⁰. To identify the genetic network involved in the defense response of mango fruit, the host's reactions against *C. gloeosporioides* were examined using Illumina paired-end sequencing technology, and expression analysis of 35 defense-related genes was investigated. The study revealed that a majority of defense-related genes, for example, nucleotide binding site-leucine-rich repeats (NBS-LRRs), ethylene response factors (ERFs), 6 pathogenesis-related proteins (PRs), and nonexpressor of pathogenesis-related genes

(NPRs) were upregulated following pathogen infection¹³⁵. These studies highlight the value of transcriptomic data in understanding the genetic networks underlying fruit quality, disease resistance, and overall crop productivity, and pave the way for future studies.

Metabolomics and Proteomics Applications in Mango Research

While transcriptomics provides insight into gene-level regulation, metabolomics provides a snapshot of the biochemical outcomes, enabling a more complete understanding of fruit traits and stress responses. Regarding specific agronomical traits in mangoes, metabolomics analysis studies have explored various aspects beyond growth and development. One important quality attribute of fruits is their flavor, which encompasses taste and aroma. The volatile compounds can provide information about the aroma of a fruit. Researchers examined mango cultivars with diverse genetic backgrounds and geographical origins by identifying natural compounds and analyzing their quantitative and qualitative variations. The cultivars exhibited variations in their volatile composition, suggesting that those from different regions have been subjected to diverse selection criteria regarding aroma^{151,152}. Such studies are useful for the mango industry and breeding programs aiming to develop cultivars with specific flavors. Many ions and compounds (methanol, ethanol, acetaldehyde, and esters) linked with fruit maturity have been identified¹⁵³. Metabolite profiling of cultivar ‘Alphonso’ and 2 different natural mango populations revealed a total of 66 metabolites that were unique to either mature unripe or ripened pulp tissue¹⁵⁴. These studies emphasize the significance of investigating natural mango populations with diverse genetic backgrounds.

Furthermore, target-based metabolomics is an essential tool for uncovering key pathways. Suh et al. (2022)¹⁵⁵ investigated the biosynthetic mechanisms of important flavor compounds in three mango cultivars (Saigon, Glenn, and Mamme) and focused on five core metabolic pathways: phenylalanine biosynthesis and metabolism, terpenoid backbone biosynthesis, linoleic and linolenic acid route, butanoate metabolism, and carbon fixation and sucrose metabolism. A comparative transcriptome analysis and targeted metabolomics of high total soluble solids in mango cultivar ‘Tainong-1’ and the low total soluble solids in cultivar ‘Renong-1’ showed that the high sugar content in ‘Tainong-1’ is likely influenced by metabolic and biosynthetic processes, potentially regulated by transcription factors such as MYB and NAC¹⁵⁶.

Antioxidant compounds in the peels of high-quality elite and local mango cultivars were determined at different stages of postharvest using metabolomics-based approaches¹⁵⁷.

Metabolomics studies are crucial for understanding abiotic defense mechanisms. In the case of ‘Chaunsa White’ cultivar, GC-MS analysis was performed on various fruit samples to investigate metabolic pathways at different stages of development. This analysis provides valuable insights into how plants respond to abiotic stress¹⁵⁸. Moreover, metabolomic fingerprinting is an essential aspect of studying the domestication history and a valuable tool for assessing metabolite differences through comparative profiling and fingerprinting analyses. For example, Farag et al. (2022)¹⁵⁹ performed a metabolomic fingerprinting study in Egypt for a better understanding of the distinct characteristics and chemical composition of Egyptian mango cultivars grown in different regions.

Proteomics can identify the proteins essential for fruit development and ripening and determine their impact on nutritional and sensory characteristics. Proteomic techniques have successfully identified numerous proteins and their associated metabolic pathways that have a role in mango fruit development, fruit coloration, maturity, ripening stages, and shelf life. For example, a mass spectrometry-based approach aligned with database searches of mango-derived Expressed Sequence Tags and proteins, led to the identification of 1001 peptides in mango leaves, which matched 538 known proteins¹⁶⁰. Fruit ripening is a complex process, and comparative proteomic studies can help identify the proteins involved in maturity and ripening. In a comparative proteomic analysis, a decrease in proteins linked with hormone biosynthesis and carbon fixation during ripening was identified, and proteins involved in catabolism and stress response (oxidative stress and defense against abiotic factors and pathogens) were observed to accumulate¹⁶¹. Further, 2-dimensional gel electrophoresis (2D-GE) in combination with MALDI-TOF/TOF analysis was used to identify proteins with differential abundance during the ripening process of two mango varieties (Chokanan and Golden Phoenix). Proteins differentially expressed in the mango fruit mesocarp during the ripe and unripe stages were categorized into three groups: proteins involved in cell wall disintegration, ethylene synthesis, and aromatic volatiles. In addition to these categories, a fourth distinct protein category related to energy and carbohydrate metabolism was identified in the ‘Chokanan’¹⁶². LC-MS/MS-based proteomics analysis of the pericarp and pulp of ‘Zill’ mango fruit successfully identified a total of 7,536 peptides that corresponded to 2,754 proteins¹⁶³.

Transcriptome studies have contributed to the identification of genes involved in different processes. By integrating transcriptome and metabolome data, mango germplasm resources can be evaluated to uncover the genetic and biochemical variability linked with quality traits. The integration of metabolomics and transcriptomics will enhance mango

breeding efforts. Through multi-omics exploration, researchers have identified that the development of abiotic stress, particularly oxidative stress, induces disturbances in various metabolic pathways and their interconnectedness, ultimately forming spongy tissue in mangoes¹⁴⁷. This has shed light on the alterations in ethylene and flavonoid biosynthesis, cell wall degradation, fruit ripening, and flavor formation, which collectively compromise the distinctive characteristics of mangoes affected by spongy tissue disorder.

Research limitations and future perspectives

Mango holds a special position in fruit crops. Conventional breeding methods for mango improvement have been hindered by the mango tree's lengthy juvenile phase. As a result, efforts for genetic improvement of mangoes using genomics and molecular biology-based advanced techniques have opened new possibilities, such as genome editing. These tools can significantly influence the breeding process for perennial crops. Shortening the juvenile period, the use of advanced NGS methods, gene editing, and transfer technologies will aid in exploring the genetic potential of mango germplasm resources and provide a foundation for improvement (Figs. 3 and 4).

The full potential of mango germplasm resources has not been harnessed. Nearly 90% of the collected germplasm resources have undergone basic phenotypic trait evaluation. Similarly, 30-50% and less than 10% of germplasm resources have been evaluated for physicochemical traits and genotype description, respectively. Therefore, many germplasm resources lack accurate identification, resulting in insufficient excavation of excellent traits and key genes. These factors collectively hinder the mango breeding progress. A strategy for harnessing the potential of diverse germplasm resources has been described in Figure 3.

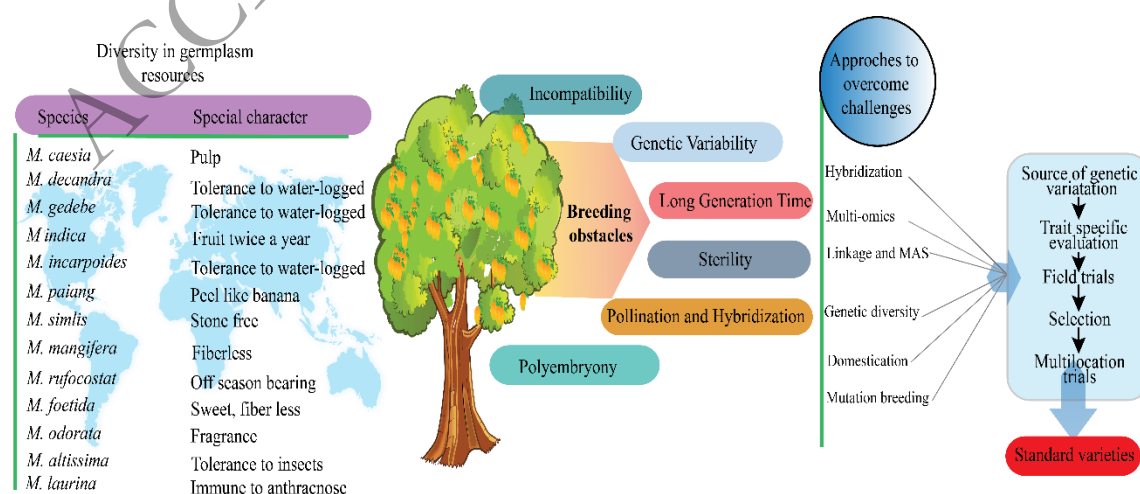


Figure 3. Genetic resources utilization and possible role of genetic resources in crop improvement.

The figure describes the genetic diversity of *Mangifera* species, highlighting special characteristics such as soft pulp, tolerance to waterlogging, twice fruiting per year, easily removable peel, stone-free fruits, sweetness with less fiber, increased fragrance, and resistance to diseases and pests. These traits can be harnessed through strategies like grafting, hybridization, multi-omics approaches, and mutation breeding to develop superior mango cultivars with better fruit quality, high yield, adaptability to diverse climatic conditions, and resistance to diseases and pests.

Advancements in single-cell multi-omics technologies can speed up the evaluation of germplasm potential by investigating the complex molecular makeup of mangoes. These methods have revolutionized research in molecular cell biology. However, progress in mango research has not been on par with these advancements. Integrated omics approaches are necessary for understanding mango plants' temporal and spatial changes under various stresses. However, there is a lack of epigenomic and functional genomic studies in mangoes. The recently developed haplotype-resolved T2T genome now enables more precise genetic studies, but challenges remain in tissue culture and genetic transformation that hinder functional validation. The development of a comprehensive pangenome will be essential to facilitate genetic diversity studies and marker-assisted breeding in future mango improvement programs. Similarly, understanding epigenetic changes and functional validation are necessary for understanding the response of genes to different stresses. The use of functional genomics with other omics disciplines will help to uncover the connections between various genomes and phenotypes under specific stress conditions. However, single-cell integrated multi-omics methods face challenges, such as high sequencing costs and limited cell coverage ¹⁶⁴. Overcoming these challenges will facilitate the development of sophisticated strategies and the creation of comprehensive atlases that integrate multiple omics datasets and different timescales, thereby increasing our understanding of mango genetics and mango breeding efforts.

In fruit plants, a long juvenile period also hinders progress efforts and can be reduced by maintaining environmental conditions that negatively affect plant vigor ^{165,166}. In Apple's manipulation of the environmental (optimal or controlled) conditions shortened the juvenility

period from 5 years to 10 months¹⁶⁷. Moreover, direct transgenic expression of flowering time genes, like *FT*, has induced early flowering in various woody perennial species, such as citrus and plum, where constitutive *FT* expression drastically shortened the juvenile phase and generation time^{168,169}. Although information regarding alteration in the juvenile period in mango is not available in the literature, the juvenile period can be reduced by optimizing growth conditions and modulating signal transduction pathways for flowering. For example, the use of transgenic rootstock expressing FT can promote early flowering. Integrating these approaches with genomic prediction can significantly shorten the juvenility period and accelerate cultivar development.

Genomic prediction allows to skip some early evaluation steps, thereby reducing the large-scale phenotyping, resources, cost, and shortening the breeding cycle^{170,171}. For example, Sun et al. (2024)¹⁷² predicted high breeding values of pear at an early stage using eight fruit traits. In apples, accessions with a favorable fruit texture were predicted¹⁷³. Brault et al. (2024)¹⁷⁴ compared two ongoing grape breeding programs to identify factors influencing the prediction accuracy. For the first time, they used a multi-trait selection index to select superior genotypes. This study provided an avenue for integrating genomic prediction into grapevine breeding programs. Compared to other fruit crops (grapes, apple, and citrus), the mango genome has been less extensively studied, and advanced methods have not been widely used in mango breeding. However, the availability of high-quality mango genomes paves the way for the integration of genomic prediction in mango breeding programs to overcome the breeding obstacles, notably the long juvenile period.

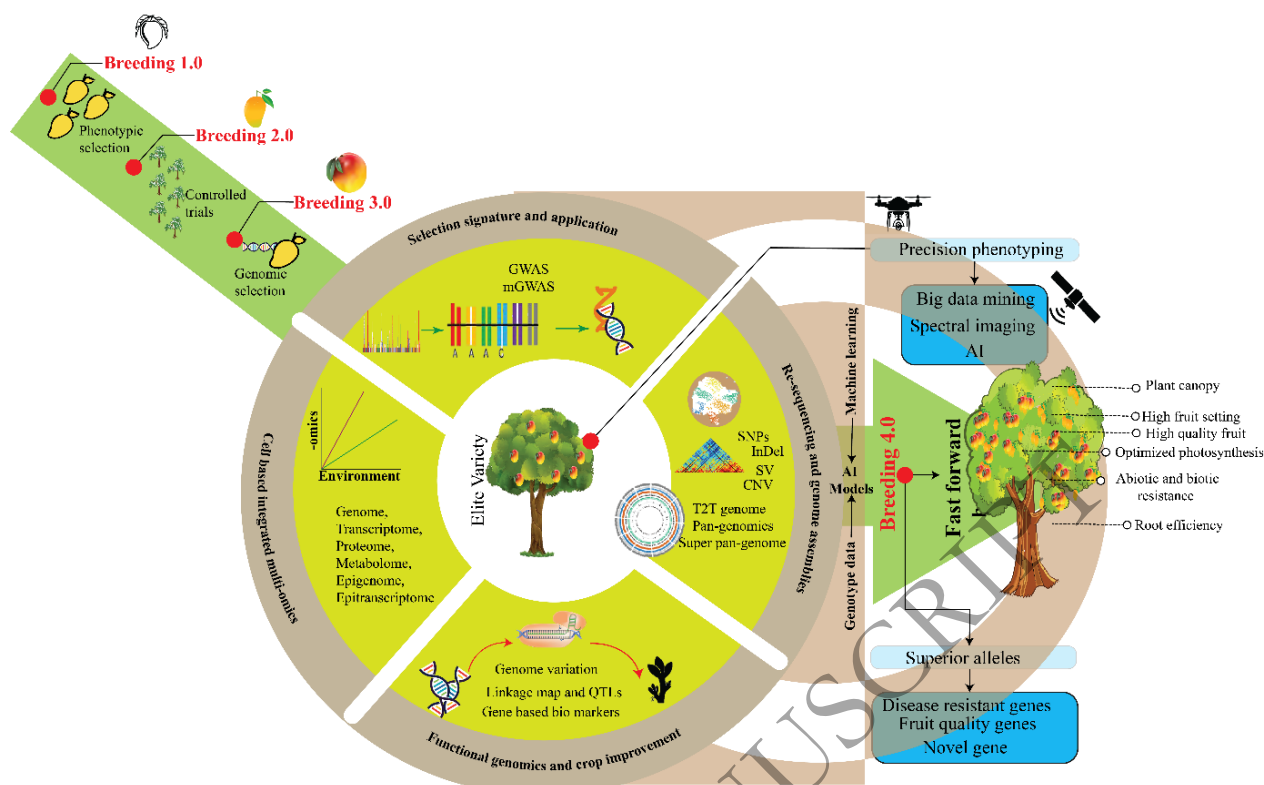


Figure 4. Multiple dimensional overviews of genomic and functional crop improvement strategies in mango. This figure explains the evolution of mango breeding from phenotypic selection (Breeding 1.0) and controlled trials (Breeding 2.0) to genomic selection (Breeding 3.0). It illustrates the integration of genome-wide association studies, multi-omics, population genetics, and functional genomics with big data mining, spectral imaging, and AI for developing mango cultivars having high fruit quality, disease resistance, and optimized photosynthesis. **Abbreviations:** GWAS (Genome-Wide Association Studies), mGWAS (Metabolite Genome-Wide Association Study), SV (Structural variations), SNP (Single Nucleotide Polymorphism), InDels (Insertions and Deletions), CNV (Copy Number Variation), T2T (Telomere to Telomere), AI (Artificial Intelligence).

Breeding efforts can progress rapidly by accurately adopting genome editing techniques, as efficient genome editing has been carried out in different horticultural crops for crop improvement^{175,176}. However, no progress has been made in mangoes due to the lack of efficient regeneration and genetic transformation systems. The gene transformation system has not been developed due to the late availability of the reference genome. The availability of mango genomes has paved the way for developing genetic transformation and genome editing protocols. Successful development of these protocols will facilitate mango breeding programs.

Integration of genomic prediction with genome editing can revolutionize mango breeding efforts.

Conventional plant breeding has reached its limitations in sustaining the continuously expanding global population and alleviating the impact of global climate change¹⁷⁷. Based on the methods and techniques used, scientists have divided breeding into four generations. Initially, people relied on phenotypic selection but lacked professional breeding technologies (Breeding 1.0). Later, with some advancements, breeders introduced purposeful strategies with Mendelian genetics (Breeding 2.0). Advancements in techniques and sequencing technologies enabled integration of molecular markers and genomic data for trait selection (Breeding 3.0)¹⁷⁸. Recently, some crops, like potato, have achieved Breeding 4.0 generation (identified and deleted deleterious alleles—hybrid potato genome design), while others, like tomato and rice, are expected to achieve it soon^{98,179}. Unfortunately, mango breeding is still lagging due to the long juvenile period, high heterozygosity, and complex genome. To overcome these challenges, we need to learn and adopt techniques used in potato genome design.

The precise estimation of genomic breeding values is crucial in genomic prediction and requires the analysis of thousands of markers across the genome¹⁸⁰. With the use of machine learning tools, genomic prediction will be easier, and the breeding cycle can be shortened—the main obstacle of mango breeding. Moreover, integrating Interactome Big Data (an effective approach for representing interactions within an organism) with machine learning can identify new functional genes based on already known genes. This integrative strategy has advantages over forward and reverse genetics approaches, as it eliminates the need to create multiple large populations and undergo stable transformation. The required data can be retrieved from public databases, so this is a cost- and time-effective approach¹⁸¹. Due to the advancements in sequencing technologies and assembly methods, the publicly available omics data for mango is increasing. Integration of Interactome Big Data with machine learning can fuel the breeding progress of mangoes (Fig. 4). In this way, mango's main breeding obstacles can be addressed (long breeding cycle and generation of large populations).

Concluding Remarks

During the last decade, significant progress has been made in mango genomic studies, which has provided a platform for improvement in breeding practices. These advancements offer important resources for mango crop improvement and domestication studies with a prime focus on QTLs linked with disease resistance, fruit quality, and shelf life. Importantly, getting insights

into the domestication and genetic improvement of a crop at the whole-genome level can greatly contribute to future endeavors aimed at enhancing yield and other desirable traits. A comprehensive pangenome needs to be built using both cultivars and wild genomes for a better understanding and improvement of mango. This will be particularly valuable for addressing genotype-environment interactions and identifying climate-resilient traits. Additionally, functional validation of mango QTLs identified in GWAS will help to understand the genetic basis of complex agronomic traits. However, progress in this area remains limited by challenges in tissue culture and genetic transformation systems. Traditional mango breeding is time-consuming and may require decades to develop a desirable variety. Modern phenomics, genomics, synthetic biology tools, and machine learning can accelerate the mango breeding process through gene validation and precision genome editing. Future efforts should focus on integrating these advanced technologies while developing solutions for the current limitations in transformation protocols and multi-environment testing.

Author Contributions

Y. Z. and X.T. conceived the project. B.A. wrote the initial draft of the manuscript, and B.A., Y.S., R.A., C. C., P. W., M.J.W., Y. C., Y.B., Z.D., J.H., X.T., and Y. Z. thoroughly edited the paper. B.A., T.R., Y.Z., T.H., Y.Z., C.L., and Z. J. performed the figures and tables drawing. All authors contributed to the proofreading of the manuscript.

Conflict of interest statement

The authors declare that they have no conflicts of interest.

Funding

This work was supported by the National Key Research and Development Program of China (No. 2024YFE0214700), the Central Government Guides Local Science and Technology Development Projects of China (No. 2023ZYZX1224), the National Natural Science Foundation of China (No. 32372662), and the Hainan Mango Agriculture Research System (Grant No. HNARS-07). Additional support was provided by the Project of National Key Laboratory for Tropical Crop Breeding (No. NKLTCB-RC202501) and the State Key Laboratory of Tropical Crop Breeding (No. SKLTCBYWF202507).

References

- 1 Mukherjee SK, Litz RE. *Introduction: botany and importance*. CABI International, 2009 doi:10.1079/9781845934897.0001.
- 2 Mango Genome C *et al*. The ‘Tommy Atkins’ mango genome reveals candidate genes for fruit quality. *BMC Plant Biol* 2021; **21**: 108.
- 3 Singh Z, Singh RK, Sane VA, Nath P. Mango - Postharvest Biology and Biotechnology. *CRC Crit Rev Plant Sci* 2013; **32**: 217–236.
- 4 Naidua GM, Naidu GR. Marketing strategies for exporting mangoes and mango products from India. In: *Acta Horticulturae*. 2009 doi:10.17660/actahortic.2009.820.6.
- 5 FAO. Major Tropical Fruits: Market Review 2020. 2021.
- 6 Kostermans AJGH, Koestermans AGH, Bompard JM. *The mangoes: Their botany, nomenclature, horticulture and utilization*. Academic Press: San Diego, CA, 1993 doi:10.1016/b978-0-12-421920-5.x5001-x.
- 7 Lim TK. *Mangifera odorata*. In: T. K. L (ed). *Edible Medicinal and Non-Medicinal Plants: Volume 1, Fruits*. Springer Netherlands: Dordrecht, 2012, pp 127–130.
- 8 Iyer CPA, Dinesh MR. Advances in classical breeding and genetics in mango. In: *Acta Horticulturae*. 1997 doi:10.17660/ActaHortic.1997.455.34.
- 9 Viruel MA, Escribano P, Barbieri M, Ferri M, Hormaza JI. Fingerprinting, embryo type and geographic differentiation in mango (*Mangifera indica* L., Anacardiaceae) with microsatellites. *Molecular Breeding* 2005; **15**. doi:10.1007/s11032-004-7982-x.
- 10 Yamanaka S *et al*. Genetic diversity and relatedness of mango cultivars assessed by SSR markers. *Breed Sci* 2019; **69**: 332–344.
- 11 Shukla M, Babu R, Mathur VK, Srivastava DK. Diverse genetic bases of Indian polyembryonic and monoembryonic mango (*Mangifera indica* L) cultivars. *Curr Sci* 2004; **85**: 870–871.
- 12 Singh LB. The Mango. Botany, cultivation, and utilization. 1960 <https://api.semanticscholar.org/CorpusID:113531234>.
- 13 Knight RJ, Schnell RJ. Mango introduction in florida and the haden cultivars significance to the modern industry. *Econ Bot* 1994; **48**: 139–145.
- 14 Bally ISE, Dillon NL. Mango (*Mangifera indica* l.) breeding. In: *Advances in Plant Breeding Strategies: Fruits*. 2018 doi:10.1007/978-3-319-91944-7_20.
- 15 Mukherjee SK. Mango: its allopolyploid nature. *Nature* 1950; **166**: 196–197.
- 16 Arumuganathan K, Earle ED. Nuclear DNA Content of Some Important Plant Species Nuclear DNA Content Material and Methods. *Plant Mol Biol Report* 1991; **9**.

- 17 Duval MF, Bunel J, Sitbon C, Risterucci AM. Development of microsatellite markers for mango (*Mangifera indica* L.). *Mol Ecol Notes* 2005; **5**. doi:10.1111/j.1471-8286.2005.01076.x.
- 18 Schnell RJ, Brown JS, Olano CT, Meerow AW, Campbell RJ, Kuhn DN. Mango genetic diversity analysis and pedigree inferences for Florida cultivars using microsatellite markers. *Journal of the American Society for Horticultural Science* 2006; **131**: 214–224.
- 19 Wang P *et al.* The genome evolution and domestication of tropical fruit mango. *Genome Biol* 2020; **21**. doi:10.1186/s13059-020-01959-8.
- 20 Ma X *et al.* Chromosome-Scale Genome and Comparative Transcriptomic Analysis Reveal Transcriptional Regulators of beta-Carotene Biosynthesis in Mango. *Front Plant Sci* 2021; **12**. doi:10.3389/fpls.2021.749108.
- 21 Wijesundara UK, Masouleh AK, Furtado A, Dillon NL, Henry RJ. A chromosome-level genome of mango exclusively from long-read sequence data. *Plant Genome* 2024. doi:10.1002/tpg2.20441.
- 22 Li W *et al.* SMRT sequencing generates the chromosome-scale reference genome of tropical fruit mango, *Mangifera indica*. *bioRxiv.org* doi:10.1101/2020.02.22.960880.
- 23 Ahmad B *et al.* Mango pangenome reveals dramatic impacts of reference bias on population genomic analyses. *Hortic Res* 2025; **12**. doi:10.1093/hr/uhaf166.
- 24 Zhang C *et al.* Gapless genome assembly and population genomics highlights diversity of mango germplasms. *Hortic Res* 2025; **12**. doi:10.1093/hr/uhaf007.
- 25 Sherman A *et al.* Mango (*Mangifera indica* L.) germplasm diversity based on single nucleotide polymorphisms derived from the transcriptome. *BMC Plant Biol* 2015; **15**. doi:10.1186/s12870-015-0663-6.
- 26 Azim MK, Khan IA, Zhang Y. Characterization of mango (*Mangifera indica* L.) transcriptome and chloroplast genome. *Plant Mol Biol* 2014; **85**. doi:10.1007/s11103-014-0179-8.
- 27 Dautt-Castro M *et al.* Mango (*Mangifera indica* L.) cv. Kent fruit mesocarp de novo transcriptome assembly identifies gene families important for ripening. *Front Plant Sci* 2015; **6**: 62.
- 28 Dautt-Castro M *et al.* Mesocarp RNA-Seq analysis of mango (*Mangifera indica* L.) identify quarantine postharvest treatment effects on gene expression. *Sci Hortic* 2018; **227**: 146–153.
- 29 Luria N *et al.* De-novo assembly of mango fruit peel transcriptome reveals mechanisms of mango response to hot water treatment. *BMC Genomics* 2014; **15**. doi:10.1186/1471-2164-15-957.

- 30 Dillon NL, Bally ISE, Wright CL, Hucks L, Innes DJ, Dietzgen RG. Genetic diversity of the Australian National Mango Genebank. *Sci Hort* 2013; **150**: 213–226.
- 31 Dillon NL, Innes DJ, Bally ISE, Wright CL, Devitt LC, Dietzgen RG. Expressed sequence tag-simple sequence repeat (EST-SSR) marker resources for diversity analysis of mango (*Mangifera indica* L.). *Diversity (Basel)* 2014; **6**. doi:10.3390/d6010072.
- 32 Kuhn DN *et al.* Estimation of genetic diversity and relatedness in a mango germplasm collection using SNP markers and a simplified visual analysis method. *Sci Hort* 2019; **252**. doi:10.1016/j.scienta.2019.03.037.
- 33 Kuhn DN, Dillon NL, Innes DJ, Wu LS, Mockaitis K. Development of single nucleotide polymorphism (SNP) markers from the mango (*Mangifera indica*) transcriptome for mapping and estimation of genetic diversity. In: *Acta Horticulturae*. 2016 doi:10.17660/ActaHortic.2016.1111.45.
- 34 Warschefsky EJ, von Wettberg EJB. Population genomic analysis of mango (*Mangifera indica*) suggests a complex history of domestication. *New Phytologist* 2019; **222**: 2023–2037.
- 35 Yu H *et al.* Isolation and characterization of two APETALA1-Like genes from mango (*Mangifera indica* L.). *Sci Hort* 2020; **259**: 108814.
- 36 Cortaga CQ, Lachica JAP, Lantican D V, Ocampo ETM. Genome-wide SNP and InDel analysis of three Philippine mango species inferred from whole-genome sequencing. *J Genet Eng Biotechnol* 2022; **20**: 46.
- 37 Wilkinson MJ *et al.* Centromeres are hotspots for chromosomal inversions and breeding traits in mango. *New Phytologist* 2024. doi:10.1111/nph.20252.
- 38 Huang X, Huang S, Han B, Li J. The integrated genomics of crop domestication and breeding. *Cell* 2022; **185**: 2828–2839.
- 39 Fosberg FR, Whitmore TC. Tropical Rain Forest of the Far East. *Taxon* 1976; **25**. doi:10.2307/1220538.
- 40 Weeks A, Zapata F, Pell SK, Daly DC, Mitchell JD, Fine PVA. To move or to evolve: Contrasting patterns of intercontinental connectivity and climatic niche evolution in ‘Terebinthaceae’ (Anacardiaceae and Burseraceae). *Front Genet* 2014; **5**. doi:10.3389/fgene.2014.00409.
- 41 Bompard JM. Taxonomy and systematics. *CABI* 2009; : 19–41.
- 42 Hooker JD. *The flora of British India*. 2011 doi:10.5962/bhl.title.678.
- 43 Matsuoka Y. Centers of origin of cultivated plants. *N I Vavilov origin and geography of cultivated plants* 1926; **55**.
- 44 Srivastava G, Mehrotra RC. Endemism due to climate change: Evidence from *Poeciloneuron* Bedd. (Clusiaceae) leaf fossil from Assam, India. *Journal of Earth System Science* 2013; **122**: 283–288.

- IV.—Records of the Geological Survey of India. *Geol Mag* 1912; **9**.
doi:10.1017/s0016756800116085.
- Iyer CPA. Recent advances in varietal improvement in mango. *Acta Hort* 1991.
doi:10.17660/actahortic.1991.291.14.
- IUCN. The IUCN Red List of Threatened Species. Version 2025-1. International
Union for Conservation of Nature. 2025.
- Popenoe W. *Manual of tropical and subtropical fruits, excluding the banana,
coconut, pineapple, citrus fruits, olive, and fig*. The Macmillan Company: New
York, 1920 doi:10.5962/bhl.title.43926.
- Mukherjee SK. The mango and its wild relatives. *Sci Cult* 1949; **26**: 5–9.
- Wilkinson MJ *et al*. The influence of genetic structure on phenotypic diversity in
the Australian mango (*Mangifera indica*) gene pool. *Sci Rep* 2022; **12**: 20614.
- Candolle A de. *Origin of cultivated plants*. K. Paul, Trench: London,
1884 <https://www.biodiversitylibrary.org/item/201281>.
- Bose TK, Mitra SK, Sanyal D. *Fruits: Tropical and subtropical*. Naya Udyog:
Calcutta, India, 2001.
- Ma X *et al*. Genomic diversity, population structure, and genome-wide association
reveal genetic differentiation and trait improvements in mango. *Hortic Res* 2024;
11: uhae153.
- Singh NK. Origin, diversity and genome sequence of mango (*Mangifera indica* L.).
Indian J Hist Sci 2016; **51**: 353–368.
- Johnson GI. Introduction of the Mango to Australia. *Proceedings of the Royal
Society of Queensland, The* 2000; **109**: 83–90.
- Jain SM, Priyadarshan PM. *Breeding plantation tree crops: tropical species*.
Springer, 2009.
- IPGRI. *Descriptors for mango (Mangifera indica L.)*. 2006.
- Singh AK *et al*. Morphological and biochemical diversity of promising mango
(*Mangifera indica* L.) indigenous cultivars in India. *Genet Resour Crop Evol* 2025.
doi:10.1007/s10722-025-02399-8.
- Song-biao Wang *et al*. Evaluation of sixteen introduced mango cultivars in
zhanjiang, CHINA. *Acta Hort* 2013; : 211–219.
- Abdul Ghaffar Grewal *et al*. Fruiting behavior and fruit quality of leading mango
cultivars grown in south Punjab-Pakistan. *Agricultural Sciences Journal* 2024; **6**:
59–72.
- Jonathan H. Crane. Expanding Mango Consumption Phase 1: Mango (*Mangifera
indica*) Cultivar Evaluation for the National Mango Board. *Proc Fla State Hort Soc*
2021; **134**: 56–59.

- 861 62 Zhang Y *et al.* Transcriptome and metabolome analyses reveal that GA3ox
862 regulates the dwarf trait in mango (*Mangifera indica* L.). *BMC Plant Biol* 2024; **24**:
863 1025.
- 864 63 Arogundade O, Matthew JO, Akinyoola OI, Akin-Idowu PE, Akinyemi SOS.
865 Phenotypic and Molecular Characterization of Mango Cultivars in Southwest
866 Nigeria. *International Journal of Fruit Science* 2022; **22**: 151–159.
- 867 64 Fuller DQ, Denham T, Allaby R. Plant domestication and agricultural ecologies.
868 *Curr Biol* 2023; **33**: R636–R649.
- 869 65 Washburn JD, Bird KA, Conant GC, Pires JC. Convergent Evolution and the Origin
870 of Complex Phenotypes in the Age of Systems Biology. *Int J Plant Sci* 2016; **177**:
871 305–318.
- 872 66 Kuhn DN *et al.* Genetic Map of Mango: A Tool for Mango Breeding. *Front Plant*
873 *Sci* 2017; **8**. doi:10.3389/fpls.2017.00577.
- 874 67 dos Santos Ribeiro ICN, Lima Neto FP, Santos CAF. Allelic database and accession
875 divergence of a Brazilian mango collection based on microsatellite markers.
876 *Genetics and Molecular Research* 2012; **11**: 4564–4574.
- 877 68 Luo C *et al.* Genetic diversity of mango cultivars estimated using SCoT and ISSR
878 markers. *Biochem Syst Ecol* 2011; **39**: 676–684.
- 879 69 Sennhenn A, Prinz K, Gebauer J, Whitbread A, Jamnadass R, Kehlenbeck K.
880 Identification of mango (*Mangifera indica* L.) landraces from Eastern and Central
881 Kenya using a morphological and molecular approach. *Genet Resour Crop Evol*
882 2014; **61**: 7–22.
- 883 70 Hirano R, Oo TH, Watanabe KN. Myanmar mango landraces reveal genetic
884 uniqueness over common cultivars from Florida, India, and Southeast Asia.
885 *Genome* 2010; **53**: 321–330.
- 886 71 Diaz-Matallana M, Schuler-Garcia I, Ruiz-Garcia M, Hodson de Jaramillo E.
887 Analysis of diversity among six populations of Colombian mango (*Mangifera*
888 *indica* L. cvar. Hilacha) using RAPDs markers. *Electronic Journal of*
889 *Biotechnology* 2009; **12**. doi:10.2225/vol12-issue3-fulltext-10.
- 890 72 Shamili M, Fatahi R, Hormaza JI. Characterization and evaluation of genetic
891 diversity of Iranian mango (*Mangifera indica* L., Anacardiaceae) genotypes using
892 microsatellites. *Sci Hortic* 2012; **148**: 230–234.
- 893 73 Ravishankar K V, Anand L, Dinesh MR. Assessment of genetic relatedness among
894 mango cultivars of India using RAPD markers. *Journal of Horticultural Science &*
895 *Biotechnology* 2000; **75**: 198–201.
- 896 74 Ravishankar K V *et al.* Genetic diversity and population structure analysis of
897 mango (*Mangifera indica*) cultivars assessed by microsatellite markers. *Trees-*
898 *Structure and Function* 2015; **29**: 775–783.

- 899 75 Surapaneni M *et al.* Population structure and genetic analysis of different utility
900 types of mango (*Mangifera indica* L.) germplasm of Andhra Pradesh state of India
901 using microsatellite markers. *Plant Systematics and Evolution* 2013; **299**: 1215–
902 1229.
- 903 76 Jena RC, Chand PK. Multiple DNA marker-assisted diversity analysis of Indian
904 mango (*Mangifera indica* L.) populations. *Sci Rep* 2021; **11**. doi:10.1038/s41598-
905 021-89470-3.
- 906 77 Bassil N V. *et al.* Development and preliminary evaluation of a 90 K Axiom® SNP
907 array for the allo-octoploid cultivated strawberry *Fragaria* × *ananassa*. *BMC*
908 *Genomics* 2015; **16**. doi:10.1186/s12864-015-1310-1.
- 909 78 Peace C *et al.* Development and Evaluation of a Genome-Wide 6K SNP Array for
910 Diploid Sweet Cherry and Tetraploid Sour Cherry. *PLoS One* 2012; **7**.
911 doi:10.1371/journal.pone.0048305.
- 912 79 Bianco L *et al.* Development and validation of the Axiom®Apple480K SNP
913 genotyping array. *Plant Journal* 2016; **86**. doi:10.1111/tpj.13145.
- 914 80 Bianco L *et al.* Development and validation of a 20K Single Nucleotide
915 Polymorphism (SNP) whole genome genotyping array for apple (*Malus* ×
916 *domestica* Borkh). *PLoS One* 2014; **9**. doi:10.1371/journal.pone.0110377.
- 917 81 Alam MS, Samsuzzaman Md, Mazed Miah MA, Akter Sumi M, Habib MA.
918 Evaluation and genetic diversity analysis of local mango germplasm. *Acta*
919 *Scientifica Malaysia* 2022; **6**. doi:10.26480/asm.01.2022.23.27.
- 920 82 Souza IGB, Valente SES, Britto FB, Desouza VAB, Lima PSC. RAPD analysis of
921 the genetic diversity of mango (*Mangifera indica*) germplasm in Brazil. *Genetics*
922 *and Molecular Research* 2011; **10**. doi:10.4238/2011.December.14.1.
- 923 83 Zhou L *et al.* Evaluation of the genetic diversity of mango (*Mangifera indica* L.)
924 seedling germplasm resources and their potential parents with start codon targeted
925 (SCoT) markers. *Genet Resour Crop Evol* 2020; **67**. doi:10.1007/s10722-019-
926 00865-8.
- 927 84 Abdelsalam NR, Elshikh MS, Ibrahim EG, Salem MZM, Ali HM. Genetic and
928 morphological characterization of *Mangifera indica* L. Growing in Egypt.
929 *HortScience* 2018; **53**. doi:10.21273/HORTSCI13084-18.
- 930 85 Fitmawati, HA. Phylogenetic Study of *Mangifera laurina* and its Related Species
931 Using cpDNA trnL-F Spacer Markers. *Hayati* 2010; **17**. doi:10.4308/hjb.17.1.9.
- 932 86 Gálvez-lópez D, Salvador-figueroa M, Adriano-anaya MDL, Mayek-Perez N.
933 Morphological Characterization of Native Mangos from Chiapas, Mexico.
934 *Subtropical Plant Science* 2010; **62**.
- 935 87 Nazish T *et al.* Molecular diversity of Pakistani mango (*Mangifera indica* L.)
936 varieties based on microsatellite markers. *Genetics and Molecular Research* 2017;
937 **16**. doi:10.4238/gmr16029560.

- 88 Eiadthong W, Yonemori K, Sugiura A, Utsunomiya N, Subhadrabandhu S. Identification of mango cultivars of Thailand and evaluation of their genetic variation using the amplified fragments by simple sequence repeat-(SSR-) anchored primers. *Sci Hortic* 1999; **82**. doi:10.1016/S0304-4238(99)00036-9.
- 89 Ho VT, Tu NT. Genetic characterization of mango accessions through rapid and SSR markers in Vietnam. *SABRAO J Breed Genet* 2019; **51**.
- 90 Iqbal MA *et al.* MiSNPDB: A web-based genomic resources of tropical ecology fruit mango (*Mangifera indica* L.) for phylogeography and varietal differentiation. *Sci Rep* 2017; **7**. doi:10.1038/s41598-017-14998-2.
- 91 Qamar-ul-Islam T, Khan MA, Faizan R, Mahmood U. MGDdb: An analyzed database and a genomic resource of mango (*Mangifera Indica* L.) cultivars for mango research. *bioRxiv* 2018.
- 92 Gomez-Olle A, Bullones A, Hormaza JJ, Mueller LA, Fernandez-Pozo N. MangoBase: A Genomics Portal and Gene Expression Atlas for *Mangifera indica*. *Plants (Basel)* 2023; **12**. doi:10.3390/plants12061273.
- 93 Soltis DE, Soltis PS. The role of phylogenetics in comparative genetics. *Plant Physiol* 2003; **132**: 1790–1800.
- 94 Singh NK, Mahato AK, Jayaswal PK. The Genome Sequence and Transcriptome Studies in Mango (*Mangifera indica* L.). In: Kole C (ed). *The Mango Genome*. Springer International Publishing: Cham, 2021, pp 165–186.
- 95 Singh A *et al.* A draft genome of the king of fruit mango (*Mangifera indica* L.). 2014.
- 96 Natalie L. Dillon and David J. Innes and Yao Ming and Xi Fang and Xusen Li and Xiaolei Gao and Rulin Zhan and Wu Hong Xia and Prasad Bajaj and Ian S. E. Bally and Alok Kumar and Rajeev Kumar Varshney. Drafting the Kensington Pride Mango Genome. 2016.
- 97 Wijesundara UK, Masouleh AK, Furtado A, Dillon NL, Henry RJ. A chromosome-level genome of mango exclusively from long-read sequence data. *Plant Genome* 2024; : e20441.
- 98 Zhang C *et al.* Genome design of hybrid potato. *Cell* 2021; **184**: 3873–3883.e12.
- 99 Tahir Ul Qamar M *et al.* Comparative genomics profiling revealed multi-stress responsive roles of the CC-NBS-LRR genes in three mango cultivars. *Front Plant Sci* 2023; **14**: 1285547.
- 100 Goodwin S, McPherson JD, McCombie WR. Coming of age: Ten years of next-generation sequencing technologies. *Nat Rev Genet.* 2016; **17**. doi:10.1038/nrg.2016.49.
- 101 Hasan N, Choudhary S, Naaz N, Sharma N, Laskar RA. Recent advancements in molecular marker-assisted selection and applications in plant breeding programmes.

- Journal of Genetic Engineering and Biotechnology. 2021; **19**. doi:10.1186/s43141-021-00231-1.
- 102 Lantican D V., Strickler SR, Canama AO, Gardoce RR, Mueller LA, Galvez HF. De novo genome sequence assembly of dwarf coconut (*Cocos nucifera* L. 'Catigan Green Dwarf') provides insights into genomic variation between coconut types and related palm species. *G3: Genes, Genomes, Genetics* 2019; **9**. doi:10.1534/g3.119.400215.
- 103 Lantican D V., Cortaga CQ, Manohar ANC, dela Cueva FM, Sison MLJ. Resistance gene analogs of mango: Insights on molecular defenses and evolutionary dynamics. *Philipp J Sci* 2020; **149**. doi:10.56899/149.3a.18.
- 104 Leister D. Retrograde signaling in plants: from simple to complex scenarios. *Front Plant Sci* 2012; **3**: 135.
- 105 Unseld M, Marienfeld JR, Brandt P, Brennicke A. The mitochondrial genome of *Arabidopsis thaliana* contains 57 genes in 366,924 nucleotides. *Nat Genet* 1997; **15**: 57–61.
- 106 Wijesundara UK, Furtado A, Dillon NL, Masouleh AK, Henry RJ. Phylogenetic relationships in the genus *Mangifera* based on whole chloroplast genome and nuclear genome sequences. *Tropical Plants* 2024; **3**: 0–0.
- 107 Bally ISE. Australian National Mango Genebank. In: The 14th Australasian Plant Breeding Conference and 11th Society for the Advancement of Breeding Research in Asia and Oceania (SABRAO). Congress 2009,. 2009.
- 108 Litz R. *Mango: Botany, Production and Uses*. 2nd ed. CABI Publishing: Wallingford, England, 2009 doi:10.1079/9781845934897.0000.
- 109 Kashkush K, Jinggui F, Tomer E, Hillel J, Lavi U. Cultivar identification and genetic map of mango (*Mangifera indica*). *Euphytica* 2001; **122**: 129–136.
- 110 Chunwongse J *et al.* Molecular mapping of mango cultivars 'Alphonso' and 'Palmer'. 2000.
- 111 Jinggui F, Dajun L, Zhengqiang M. Construction of mango (*Mangifera indica* L.) genetic map using markers for double heterozygous loci. *Mol Plant Breed* 2003; **1**: 313–319.
- 112 Luo C, Shu B, Yao Q, Wu H, Xu W, Wang S. Construction of a High-Density Genetic Map Based on Large-Scale Marker Development in Mango Using Specific-Locus Amplified Fragment Sequencing (SLAF-seq). *Front Plant Sci* 2016; **7**. doi:10.3389/fpls.2016.01310.
- 113 Chunwongse J *et al.* Molecular mapping of mango cultivars 'Alphonso' and 'Palmer'. 2000.
- 114 Chunwongse C, Phumichai C, Tongyoo P, Juejun N, Chunwongse J. Development of di-nucleotide microsatellite markers and construction of genetic linkage map in

- 1014 mango (*Mangifera indica* L.). *Songklanakarin Journal of Science and Technology*
1015 2015; **37**.
- 1016 115 Srivastav M *et al.* High resolution mapping of QTLs for fruit color and firmness in
1017 Amrapali/Sensation mango hybrids. *Front Plant Sci* 2023; **14**.
1018 doi:10.3389/fpls.2023.1135285.
- 1019 116 Deshpande AB *et al.* Transcriptional transitions in Alphonso mango (*Mangifera*
1020 *indica* L.) during fruit development and ripening explain its distinct aroma and shelf
1021 life characteristics. *Sci Rep* 2017; **7**: 8711.
- 1022 117 Srivastava P, Vyas VK, Variya B, Patel P, Qureshi G, Ghate M. Synthesis, anti-
1023 inflammatory, analgesic, 5-lipoxygenase (5-LOX) inhibition activities, and
1024 molecular docking study of 7-substituted coumarin derivatives. *Bioorg Chem* 2016;
1025 **67**: 130–138.
- 1026 118 Srivastav M *et al.* High resolution mapping of QTLs for fruit color and firmness in
1027 Amrapali/Sensation mango hybrids. *Front Plant Sci* 2023; **14**. doi:ARTN 1135285
1028 10.3389/fpls.2023.1135285.
- 1029 119 Sandip M, Makwana AN, Barad A V, Nawade BD. Physiology of flowering-the
1030 case of mango. *International Journal of Applied Research* 2015; **1**: 1008–1012.
- 1031 120 Sharma N, Singh AK, Singh SK, Mahato AK, Srivastav M, Singh NK. Comparative
1032 RNA sequencing based transcriptome profiling of regular bearing and alternate
1033 bearing mango (*Mangifera indica* L.) varieties reveals novel insights into the
1034 regulatory mechanisms underlying alternate bearing. *Biotechnol Lett* 2020; **42**:
1035 1035–1050.
- 1036 121 Lal S *et al.* Association analysis for pomological traits in mango (*Mangifera indica*
1037 L.) by genic-SSR markers. *Trees* 2017; **31**: 1391–1409.
- 1038 122 Nakagawa M, Honsho C, Kanzaki S, Shimizu K, Utsunomiya N. Isolation and
1039 expression analysis of FLOWERING LOCUS T-like and gibberellin metabolism
1040 genes in biennial-bearing mango trees. *Sci Hort* 2012; **139**: 108–117.
- 1041 123 Yadav A, Jayaswal PK, Raman K V, Singh B, Singh NK, Usha K. Transcriptome
1042 analysis of flowering genes in mango in relation to floral malformation (vol 33, pg
1043 789, 2020). *J Plant Biochem Biotechnol* 2020; **29**: 571–572.
- 1044 124 Kaur H *et al.* Comparative transcriptomics in alternate bearing cultivar Dashehari
1045 reveals the genetic model of flowering in mango. *Front Genet* 2023; **13**. doi:ARTN
1046 1061168 10.3389/fgene.2022.1061168.
- 1047 125 Wang Y-H *et al.* Overexpression of four MiTFL1 genes from mango delays the
1048 flowering time in transgenic Arabidopsis. *BMC Plant Biol* 2021; **21**: 407.
- 1049 126 Wang YH, Yu HX, He XH, Lu TT, Huang X, Luo C. Isolation and Functional
1050 Characterization of a Gene in Mango. *Int J Mol Sci* 2022; **23**. doi:ARTN 3974
1051 10.3390/ijms23073974.

- 127 Guo C, Qin L, Ma Y, Qin J. Integrated metabolomic and transcriptomic analyses of
the parasitic plant *Cuscuta japonica* Choisy on host and non-host plants. *BMC Plant*
Biol 2022; **22**. doi:10.1186/s12870-022-03773-9.
- 128 Fan Z-Y *et al.* Isolation and functional characterization of three MiFTs genes from
mango. *Plant Physiology and Biochemistry* 2020; **155**: 169–176.
- 129 Mo X *et al.* Isolation and functional characterization of two SHORT VEGETATIVE
PHASE homologous genes from mango. *Int J Mol Sci* 2021; **22**: 9802.
- 130 Liu Y *et al.* Isolation and functional characterization of two CONSTANS-like 16
(MiCOL16) genes from mango. *Int J Mol Sci* 2022; **23**: 3075.
- 131 Bajpai A, Khan K, Muthukumar M, Rajan S, Singh NK. Molecular analysis of
anthocyanin biosynthesis pathway genes and their differential expression in mango
peel. *Genome* 2018; **61**: 157–166.
- 132 Yungyuen W *et al.* Carotenoid accumulation and the expression of carotenoid
metabolic genes in mango during fruit development and ripening. *Applied Sciences*
2021; **11**: 4249.
- 133 Srivastava S *et al.* Comparative transcriptome analysis of unripe and mid-ripe fruit
of *Mangifera indica* (var. “Dashehari”) unravels ripening associated genes. *Sci Rep*
2016; **6**: 32557.
- 134 Sane VA, Chourasia A, Nath P. Softening in mango (*Mangifera indica* cv. Dashehari)
is correlated with the expression of an early ethylene responsive, ripening related
expansin gene, MiExpA1. *Postharvest Biol Technol* 2005; **38**: 223–230.
- 135 Hong K *et al.* Transcriptome characterization and expression profiles of the related
defense genes in postharvest mango fruit against *Colletotrichum gloeosporioides*.
Gene 2016; **576**: 275–283.
- 136 Li JJ *et al.* Genome-wide identification of the mango pathogenesis-related 1 (PR1)
gene family and functional analysis of genes in transgenic *Arabidopsis*. *Sci Hortic*
2023; **321**. doi:ARTN 112254 10.1016/j.scienta.2023.112254.
- 137 Liu Y *et al.* Genome-wide identification of the mango CONSTANS (CO) family
and functional analysis of two MiCOL9 genes in transgenic *Arabidopsis*. *Front*
Plant Sci 2022; **13**: 1028987.
- 138 Guo Y-H *et al.* Isolation and functional analysis of two CONSTANS-like 1 genes
from mango. *Plant Physiology and Biochemistry* 2022; **172**: 125–135.
- 139 Bastawros MB. Mango malformation in Egypt. *5th International Mango*
Symposium, Vols 1 and 2 1997; : 566–574.
- 140 Bally ISE, Akem CN, Dillon NL, Grice C, Lakhesar D, Stockdale K. Screening and
breeding for genetic resistance to anthracnose in mango. In: *IX International*
Mango Symposium 992. 2010, pp 239–244.
- 141 Sanchez-Arizpe A, Galindo-Cepeda ME, Arispe-Vazquez JL, Genis-Velazquez R,
Vazquez-Badillo ME, Antonio-Bautista A. Natural resistance of two mango

- 1091 “Mangifera indica” L. commercial cultivars to anthracnose caused by
1092 “Colletotrichum gloeosporioides” Penz. Penz. and Sacc. *Aust J Crop Sci* 2021; **15**:
1093 1198–1203.
- 1094 142 Felipe JEL *et al.* Validation and molecular analysis of β -1, 3-GLU2 SNP marker
1095 associated with resistance to Colletotrichum gloeosporioides in mango (Mangifera
1096 indica L.). *Physiol Mol Plant Pathol* 2022; **118**: 101804.
- 1097 143 Alseekh S, Kostova D, Bulut M, Fernie AR. Genome-wide association studies:
1098 assessing trait characteristics in model and crop plants. *Cellular and Molecular Life*
1099 *Sciences* 2021; **78**: 5743–5754.
- 1100 144 Wang RZ, Li XL, Sun MY, Xue C, Korban SS, Wu J. Genomic insights into
1101 domestication and genetic improvement of fruit crops. *Plant Physiol* 2023.
1102 doi:10.1093/plphys/kiad273.
- 1103 145 Li Y *et al.* Genomic analyses of an extensive collection of wild and cultivated
1104 accessions provide new insights into peach breeding history. *Genome Biol* 2019;
1105 **20**: 36.
- 1106 146 Tongyoo P, Duangjit J, Sharma N, Chunwongse J. Molecular Mapping and
1107 Breeding in Mango. In: Kole C (ed). *The Mango Genome*. Springer International
1108 Publishing: Cham, 2021, pp 153–163.
- 1109 147 Oak P *et al.* Transcriptional and translational perturbation in abiotic stress induced
1110 physiological activities and metabolic pathway networks in spongy tissue disorder
1111 of mango fruit. *Postharvest Biol Technol* 2022; **188**.
1112 doi:10.1016/j.postharvbio.2022.111880.
- 1113 148 Wu S *et al.* Metabolomic and transcriptomic analyses reveal new insights into the
1114 role of abscisic acid in modulating mango fruit ripening. *Hortic Res* 2022; **9**.
1115 doi:10.1093/hr/uhac102.
- 1116 149 Pathak G, Dudhagi SS, Raizada S, Sane VA. Transcriptomic Insight into Aroma
1117 Pathway Genes and Effect of Ripening Difference on Expression of Aroma Genes
1118 in Different Mango Cultivars. *Plant Mol Biol Report* 2023; **41**: 145–163.
- 1119 150 Sudheeran PK *et al.* Induced defense response in red mango fruit against
1120 Colletotrichum gloeosporioides. *Hortic Res* 2021; **8**. doi:10.1038/s41438-020-
1121 00452-4.
- 1122 151 Lebrun M, Plotto A, Goodner K, Ducamp M-N, Baldwin E. Discrimination of
1123 mango fruit maturity by volatiles using the electronic nose and gas chromatography.
1124 *Postharvest Biol Technol* 2008; **48**: 122–131.
- 1125 152 Pandit SS *et al.* Expression profiling of various genes during the fruit development
1126 and ripening of mango. *Plant Physiology and Biochemistry* 2010; **48**: 426–433.
- 1127 153 White IR, Blake RS, Taylor AJ, Monks PS. Metabolite profiling of the ripening of
1128 Mangoes Mangifera indica L. cv. ‘Tommy Atkins’ by real-time measurement of
1129 volatile organic compounds. *Metabolomics* 2016; **12**. doi:10.1007/s11306-016-
1130 0973-1.

- 1131 154 Dar MS *et al.* Differential Modulation in Metabolites Revealed with the
1132 Improvement in the Shelf-Life of Alphonso Fruits. *Mol Biotechnol* 2020; **62**: 508–
1133 520.
- 1134 155 Suh JH, Madden RT, Sung J, Chambers AH, Crane J, Wang Y. Pathway-Based
1135 Metabolomics Analysis Reveals Biosynthesis of Key Flavor Compounds in Mango.
1136 *J Agric Food Chem* 2022; **70**: 10389–10399.
- 1137 156 Li L *et al.* Transcriptional mechanism of differential sugar accumulation in pulp of
1138 two contrasting mango (*Mangifera indica* L.) cultivars. *Genomics* 2020; **112**: 4505–
1139 4515.
- 1140 157 Monribot-Villanueva JL *et al.* Endorsing and extending the repertoire of
1141 nutraceutical and antioxidant sources in mangoes during postharvest shelf life.
1142 *Food Chem* 2019; **285**: 119–129.
- 1143 158 Khanum Z, Tiznado-Hernandez ME, Ali A, Musharraf SG, Shakeel M, Khan IA.
1144 Adaptation mechanism of mango fruit (*Mangifera indica* L. cv. Chaunsa White) to
1145 heat suggest modulation in several metabolic pathways. *RSC Adv* 2020; **10**: 35531–
1146 35544.
- 1147 159 Farag MA *et al.* A Comparative Metabolomics Approach for Egyptian Mango
1148 Fruits Classification Based on UV and UPLC/MS and in Relation to Its Antioxidant
1149 Effect. *Foods* 2022; **11**. doi:10.3390/foods11142127.
- 1150 160 Renuse S *et al.* Proteomic analysis of an unsequenced plant - *Mangifera indica*. *J*
1151 *Proteomics* 2012; **75**: 5793–5796.
- 1152 161 Andrade J de M, Toledo TT, Nogueira SB, Cordenunsi BR, Lajolo FM, Oliveira do
1153 Nascimento JR. 2D-DIGE analysis of mango (*Mangifera indica* L.) fruit reveals
1154 major proteomic changes associated with ripening. *J Proteomics* 2012; **75**: 3331–
1155 3341.
- 1156 162 Chin CF, Teoh EY, Chee MJY, Al-Obaidi JR, Rahmad N, Lawson T. Comparative
1157 Proteomic Analysis on Fruit Ripening Processes in Two Varieties of Tropical
1158 Mango (*Mangifera indica*). *Protein Journal* 2019; **38**: 704–715.
- 1159 163 Wu H *et al.* Transcriptome and proteomic analysis of mango (*Mangifera indica*
1160 Linn) fruits. *J Proteomics* 2014; **105**: 19–30.
- 1161 164 Miao Z, Humphreys BD, McMahon AP, Kim J. Multi-omics integration in the age
1162 of million single-cell data. *Nat Rev Nephrol* 2021; **17**: 710–724.
- 1163 165 Visser T. The relation between growth, juvenile Period and fruiting of apple
1164 seedlings And its use to improve breeding Efficiency. *Euphytica* 1970; **19**: 293–
1165 302.
- 1166 166 Zimmerman RH. Juvenility and Flowering in Woody Plants: A Review1.
1167 *HortScience* 1972; **7**: 447–455.
- 1168 167 Aldwinckle HS. Flowering of Apple Seedlings 16-20 Months after Germination1.
1169 *HortScience* 1975; **10**: 124–126.

- 168 Endo T, Shimada T, Fujii H, Kobayashi Y, Araki T, Omura M. Ectopic expression
of an homolog from confers an early flowering phenotype on trifoliate orange (*L.*
Raf.). *Transgenic Res* 2005; **14**: 703–712.
- 169 Srinivasan C, Dardick C, Callahan A, Scorza R. Plum () Trees Transformed with
Poplar Result in Altered Architecture, Dormancy Requirement, and Continuous
Flowering. *PLoS One* 2012; **7**. doi:ARTN e40715 10.1371/journal.pone.0040715.
- 170 Heffner EL, Sorrells ME, Jannink JL. Genomic Selection for Crop Improvement.
Crop Sci 2009; **49**: 1–12.
- 171 Meuwissen TH, Hayes BJ, Goddard ME. Prediction of total genetic value using
genome-wide dense marker maps. *Genetics* 2001; **157**: 1819–1829.
- 172 Sun MY *et al.* Genomic selection of eight fruit traits in pear. *Hortic Plant J* 2024;
10: 318–326.
- 173 Roth M, Muranty H, Di Guardo M, Guerra W, Patocchi A, Costa F. Genomic
prediction of fruit texture and training population optimization towards the
application of genomic selection in apple. *Hortic Res* 2020; **7**. doi:ARTN 148
10.1038/s41438-020-00370-5.
- 174 Brault C *et al.* Enhancing grapevine breeding efficiency through genomic
prediction and selection index. *G3-Genes Genomes Genetics* 2024; **14**.
doi:10.1093/g3journal/jkae038.
- 175 Jia HG, Orbovic V, Wang N. CRISPR-LbCas12a-mediated modification of citrus.
Plant Biotechnol J 2019; **17**: 1928–1937.
- 176 Osakabe Y *et al.* CRISPR-Cas9-mediated genome editing in apple and grapevine.
Nat Protoc 2018; **13**: 2844–2863.
- 177 Gao C. Genome engineering for crop improvement and future agriculture. *Cell*
2021; **184**: 1621–1635.
- 178 Wallace JG, Rodgers-Melnick E, Buckler ES. On the Road to Breeding 4.0:
Unraveling the Good, the Bad, and the Boring of Crop Quantitative Genomics.
Annu Rev Genet 2018; **52**: 421–444.
- 179 Li T *et al.* Domestication of wild tomato is accelerated by genome editing. *Nat*
Biotechnol 2018; **36**: 1160–1163.
- 180 Lourenco VM, Ogutu JO, Rodrigues RAP, Posekany A, Piepho HP. Genomic
prediction using machine learning: a comparison of the performance of regularized
regression, ensemble, instance-based and deep learning methods on synthetic and
empirical data. *BMC Genomics* 2024; **25**: 152.
- 181 Wu L, Han L, Li Q, Wang G, Zhang H, Li L. Using Interactome Big Data to Crack
Genetic Mysteries and Enhance Future Crop Breeding. *Mol Plant* 2021; **14**: 77–94.

Figure Legends

Figure 1. Important developments in mango genomics and future priorities.

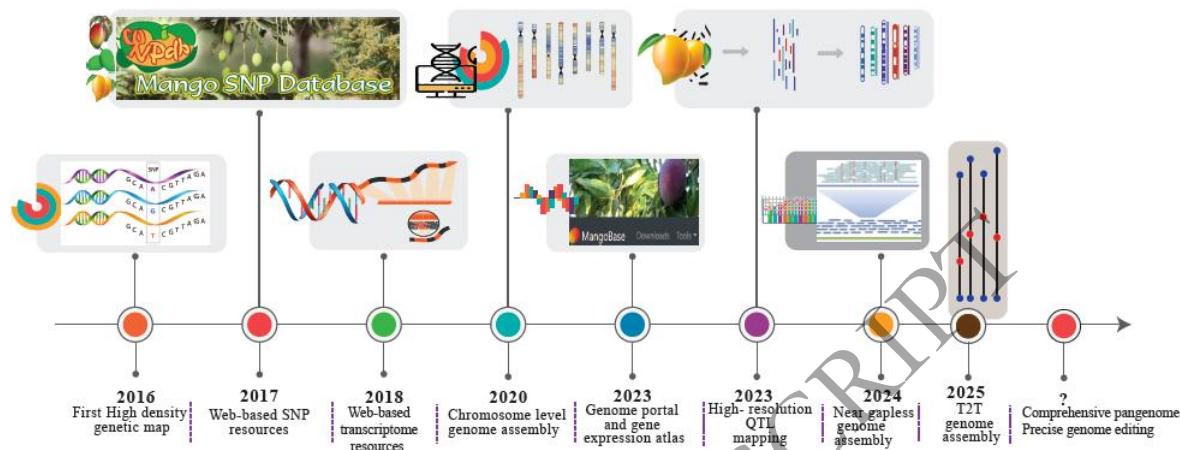


Figure 2. A holistic exploration of the origins and global dispersal patterns of mangoes. Yellow shapes represent regions including multiple countries, while the size of the purple circles indicates germplasm diversity.

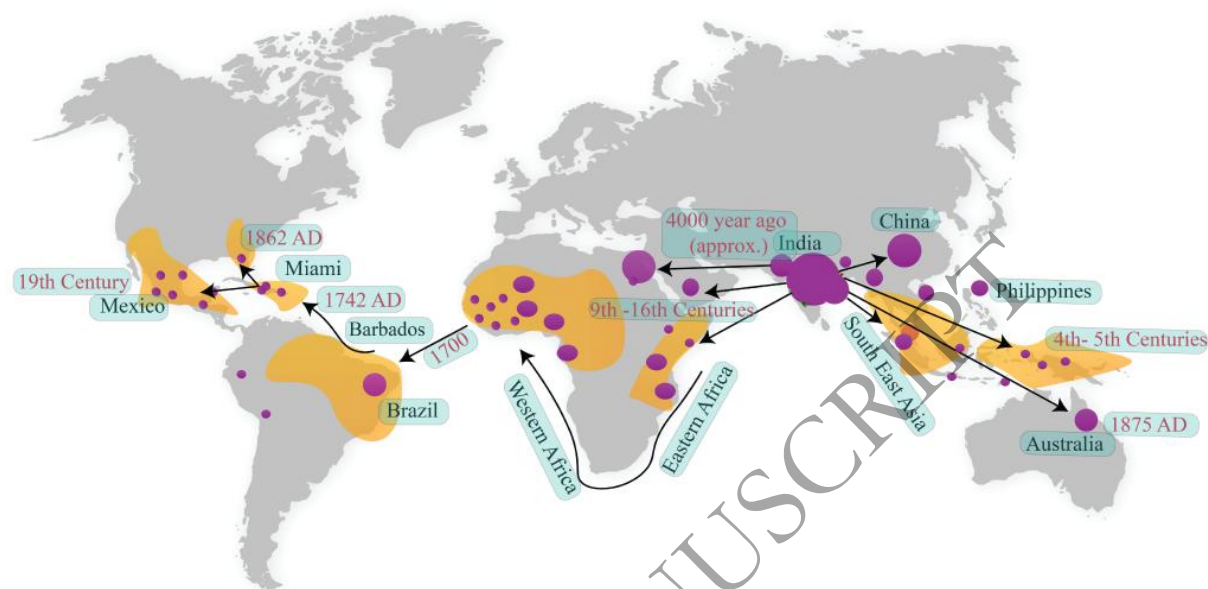


Figure 3. Genetic resources utilization and possible role of genetic resources in crop improvement. The figure describes the genetic diversity of *Mangifera* species, highlighting special characteristics such as soft pulp, tolerance to waterlogging, twice fruiting per year, easily removable peel, stone-free fruits, sweetness with less fiber, increased fragrance, and resistance to diseases and pests. These traits can be harnessed through strategies like grafting, hybridization, multi-omics approaches, and mutation breeding to develop superior mango cultivars with better fruit quality, high yield, adaptability to diverse climatic conditions, and resistance to diseases and pests.

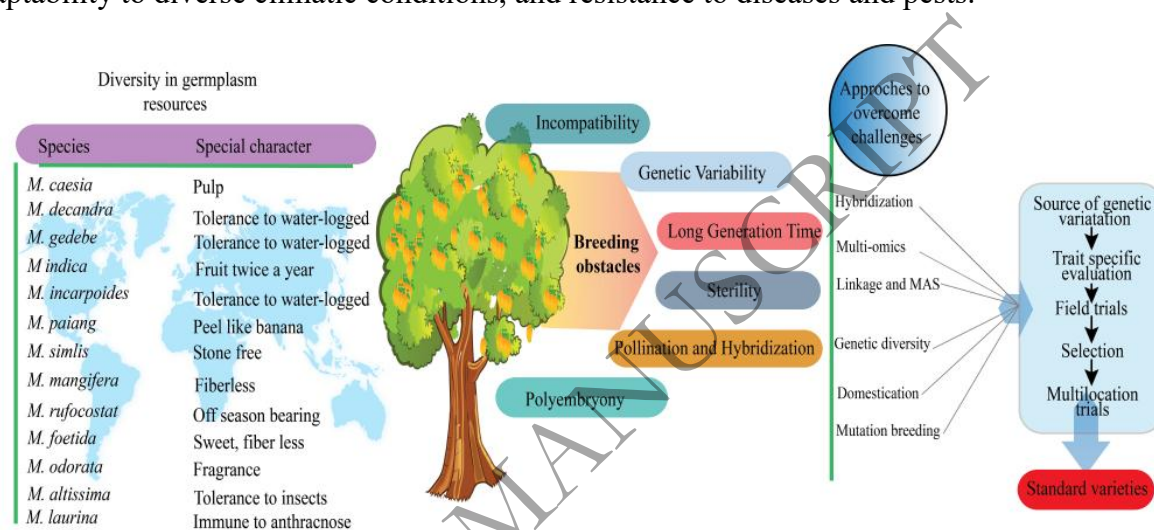


Figure 4. Multiple-dimensional overviews of genomic and functional crop improvement strategies in mango. This figure explains the evolution of mango breeding from phenotypic selection (Breeding 1.0) and controlled trials (Breeding 2.0) to genomic selection (Breeding 3.0). It illustrates the integration of genome-wide association studies, multi-omics, population genetics, and functional genomics with big data mining, spectral imaging, and AI for developing mango cultivars having high fruit quality, disease resistance, and optimized photosynthesis. **Abbreviations:** GWAS (Genome-Wide Association Studies), mGWAS (Metabolite Genome-Wide Association Study), SV (Structural variations), SNP (Single Nucleotide Polymorphism), InDels (Insertions and Deletions), CNV (Copy Number Variation) T2T (Telomere to Telomere) AI (Artificial Intelligence).

