Population structure and genetic diversity of mango (*Mangifera indica* L.) germplasm resources as revealed by single-nucleotide polymorphism markers

Qingzhi Liang[1,2], Hongbing Pan[3], Xiaolong He[2], Songbiao Wang[2], Yuanhua Hou[1,4], Hua Xiao[5], Guangzhao Xu[1], Runhua Yi[1], Dongbo Lin[1] and Zhuanying Yang[1]*

[1]College of Coastal Agricultural Sciences, Guangdong Ocean University, Zhanjiang, China, [2]South Subtropical Crops Research Institute, Chinese Academy of Tropical Agricultural Sciences, Zhanjiang, China, [3]Fruits Research Institute, Panzhihua Academy of Agricultural and Forestry Sciences, Panzhihua, China, [4]College of Tropical Crops, Yunnan Agricultural University, Puer, China, [5]College of Electronic and Information Engineering, Guangdong Ocean University, Zhanjiang, China

Introduction: Mango is a vital horticultural fruit crop, and breeding is an essential strategy to enhance ongoing sustainability. Knowledge regarding population structure and genetic diversity in mango germplasm is essential for crop improvement.

Methods: A set of 284 mango accessions from different regions of the world were subjected to high-throughput sequencing and specific-locus amplified fragment (SLAF) library construction to generate genomic single-nucleotide polymorphism (SNP).

Results: After filtering, raw data containing 539.61 M reads were obtained. A total of 505,300 SLAFs were detected, of which, 205,299 were polymorphic. Finally, 29,136 SNPs were employed to dissect the population structure, genetic relationships, and genetic diversity. The 284 mango accessions were divided into two major groups: one group consisted mainly of mango accessions from Australia, the United States, Cuba, India, Caribbean, Israel, Pakistan, Guinea, Burma, China, and Sri Lanka, which belonged to the Indian type (P1); the other group contained mango accessions from the Philippines, Thailand, Indonesia, Vietnam, Cambodia, Malaysia, and Singapore, which belonged to Southeast Asian type (P2). Genetic diversity, principal component analysis (PCA), and population structure analyses revealed distinct accession clusters. Current results indicated that the proposed hybridization occurred widely between P1 and P2.

Discussion: Most of the accessions (80.99%) were of mixed ancestry, perhaps including multiple hybridization events and regional selection, which merits further investigation.

KEYWORDS

Mangifera indica, SLAFs, population structure, genetic diversity, SNPs

Introduction

Mango (*Mangifera indica* L., 2n = 40) is an important tropical and subtropical fruit that belongs to the Anacardiaceae family (Wang et al., 2020; Zhou et al., 2020). Mango, the "king of tropical fruits," is an economically important and the fifth largest fruit crop in the world. The tropical area Southeast Asia is the main growing and production center for mangoes (Luo et al., 2016; Lora and Hormaza, 2018; Liang et al., 2022). At present, globally more than 100 countries grow mangoes, spanning between 30° north and 30° south latitude. India has the largest area of mango cultivation, which accounts for 40%, followed by China, approximately 17% (Lal et al., 2017; Niu et al., 2021; Liu et al., 2023). Approximately 1,370 years ago, mango cultivation started in China and mango was introduced to Southeast Asia (Mukherjee and Litz, 2009; Gao et al., 2011). Mainly five species of mango were cultivated in China, and approximately 200 cultivars belonging to *Mangifera indica* L. were released for cultivation (Zhou et al., 2020).

Around 69 species of the mango genus, mainly found in India, Sri Lanka, China, and Philippines (Bompard, 2009; Gao et al., 2011; Lora and Hormaza, 2018), one specie (*Mangifera indica*) and more than 1,000 mango cultivars were currently cultivated worldwide (Wang et al., 2020). Botanically, there are two main groups: (i) monoembryonic type, where the seed has only one embryo and only one seedling emerges after sowing, (ii) polyembryonic type: where the seeds have multiple embryos and after sowing several seedlings can grow—mango accessions from Philippine varieties and Thailand varieties belong to this type (Mukherjee and Litz, 2009; Warschefsky and Wettberg, 2019).

Mango had been cultivated in India and Indochina for more than 4,000 years before being introduced to Africa, South America, and other continents (Popenoe, 1920; Mukherjee, 1949; Singh et al., 2016; Warschefsky and Wettberg, 2019). Mangoes arrived in Miami from the West Indies in 1862 or 1863 (Litz, 2009). During the twentieth century, with the implementation of a breeding program in South Florida, more and more elite commercial cultivars were released; some of them are still the dominant varieties in mango-producing regions (Knight et al., 2009). Historically, mango was introduced to China; two possible routes were proposed. The first recorded introduction of mangoes was in the Tang Dynasty (Mukherjee and Litz, 2009; Gao et al., 2011); mango was secondly introduced in China from India and South Asia in the fifteenth century AD by sea during Zheng He's Voyages to the Western Seas (Ming Dynasty. Xu, 1992; Yang, 2019). The mango breeding program in China had gone through two stages: the first stage was seedling breeding, where a large number of original varieties were selected from seedlings of Indian or Southeast Asian type varieties (P1 or P2); the second one was cross breeding, where the artificial cross breeding between P1 and P2 produced many excellent varieties which had been widely promoted in China (Luo et al., 2010; Huang et al., 2013; 2012).

Mango cultivation has greatly promoted the income of farmers in tropical regions; however, its molecular biology research was relatively lagging behind (Warschefsky and Wettberg, 2019). In recent years, with the development of molecular biology, the development of molecular markers and genetic diversity evaluation of mango germplasm resources have made some progress (Nadeem et al., 2017; Zhou et al., 2020). Previous studies used nuclear DNA sequence fragments such as RAPD (Schnell et al., 1995), AFLP (Yamanaka et al., 2006); ISSR (Singh et al., 2007), SSR (Shamili et al., 2012), SCoT (Zhou et al., 2020), and SNP (Sherman et al., 2015; Kuhn et al., 2019) markers to evaluate the genetic diversity of mango germplasm resources. However, further research on mango biology was limited due to the limited number of polymorphic molecular markers. The development of mango chromosome-level whole genome sequencing (Wang et al., 2020; Ma et al., 2021; Mango Genome Consortium et al., 2021) and high-throughput sequencing technology has made it possible for mango to develop bulk molecular markers quickly and cheaply.

Specific-locus amplified fragment sequencing (SLAF-seq) technology was an efficient method, which was developed by Beijing Biomarker Technologies Corporation (Sun et al., 2013). To date, SLAF-seq has been used successfully to dissect accurately the genetic diversity and population structure for many crops (Shen et al., 2017; Zeng et al., 2017; Chang et al., 2019). The results of such techniques were further used in plant breeding (Jones et al., 2013; Boeven et al., 2016; Wu et al., 2016; Zhang et al., 2018; Bruce et al., 2019; Cui et al., 2020).

In the current study, 284 mango germplasm resources were collected from different regions worldwide. However, the genetic diversity of these mango germplasm resources in the collection remained molecularly uncharacterized. Therefore, this research analyzed the genetic diversity and population structure of 284 mango germplasm resources using SNP markers in order to generate highly important information, which together with previous research results led to deeper insight on the mango gene pool for mango breeders worldwide.

Materials and methods

Sampling

In order to dissect the population structure and genetic diversity of mango germplasm resources in China, 284 mango accessions were selected from National Field Genebank for Tropical Fruit (Zhanjiang, China) that originated in different geographical regions: a total of 18 countries or regions of major mango growing areas of the world, mainly including India, Pakistan, Thailand, Burma, Sri Lanka, Indonesia, the Philippines, Australia, the United States, China, etc. These accessions were collected from different geographical areas worldwide since the 1920s (from the earliest introduction of new mango varieties by overseas Chinese to the later introduction and exchange of germplasm resources worldwide); we tried to collect more diverse and distinctive mango germplasm resources in morphology and geography from every tropical region, including historical varieties particularly. Additionally, we collected also landrace, new varieties, and elite lines identified through the Chinese breeding program (Supplementary Table S1).

DNA extractions and SLAF-seq

Genomic DNA from young leaves of 284 mango accessions were extracted using the modified CTAB method described by He et al. (2005). The quality and purity were determined with 1.2% agarose gel electrophoresis and spectrophotometry at a wavelength of 260/280 nm using a BioPhotometer (D30, Eppendorf, Germany), respectively. In the present study, 284 mango germplasm resources were subjected to molecular marker development using the SLAF-seq approach, which was developed by Beijing Biomarker Technologies Corporation to obtain genome-wide molecular markers (Sun et al., 2013). Except that the genomic DNA was digested with double restriction enzymes Hpy166II and EcoRV. Sequences with an enzyme digestion length of 264 bp–414 bp were defined as SLAF tags. These SALF tags were aligned with the latest mango genome sequences (https://www.ncbi.nlm.nih.gov/assembly/GCA_016746415.1) and localized to the corresponding chromosomes. SNPs were obtained using GATK and SAMtools (Li and Durbin, 2009; McKenna et al., 2010). SNPs with minor allele frequency (MAF) < 5% and integrity < 80% were excluded from the genotype data sets of all the accessions.

Data analysis

The mean effective numbers of observed heterozygosity (Ho), expected heterozygosity (He), and fixation index (Fst) for each SNP marker and accession were estimated using software PLINK v1.90b6.21. Population genetic structure analysis is an important tool for genetic relationship analysis, which can provide information about the origin and composition of individual lineages. Based on 29,136 polymorphic SNP markers, the population structure of 284 germplasm resources was analyzed using STRUCTURE V2.3.4 software (Alexander et al., 2009; Excoffier and Lischer, 2010). To determine the number of hypothetical clusters (K), define populations, and assign individual accessions to certain subpopulations based on genetic data, an admixture and shared allele frequencies model was employed (Pritchard et al., 2000). Numbers in the range from 1 to 10 were assumed for K. The initial burn-in period, for each run, was set to 10,000 with 100,000 MCMC (Markov chain Monte Carlo) iterations, with no previous information on the source of accessions. The most suitable value of K was calculated using the ΔK method as used in the Structure Harvester web page (Evanno et al., 2005).

A phylogenetic tree was used to represent the evolutionary relationship among mango germplasm resources. According to the distance of the relationship between various organisms, all kinds of organisms were placed on the branched tree chart, which showed the evolutionary process and relationship of organisms. Based on SNP markers, the phylogenetic tree of 284 mango accessions was analyzed by RAxML version 8.2.12 software and the neighbor-joining algorithm (Saitou and Nei, 1987; Tamura et al., 2011). A bootstrap consensus tree was obtained from 1,000 replicates.

Principal component analysis (PCA) of the selected SNPs was performed with Plink v1.90b6.21 software. Based on SNP markers

and cluster software, principal component analysis (PCA) was carried out to obtain the clustering of 284 mango germplasm resources. PCA can be used to determine whose samples were relatively close or whose samples were relatively distant, which can assist evolutionary analysis (Dunteman, 1989; de Hoon et al., 2004). Meanwhile, to explore the similarities and dissimilarities among samples, principal coordinate analysis (PCoA) was also conducted. A distance matrix was calculated using Euclidean distance from the PCA matrix. PCoA was then performed using the cmdscale function in R, retaining two dimensions (k = 2). The resulting coordinates were transformed into a data frame, and sample groups were assigned for visualization purposes. Regions of interest were identified based on the data distribution, with specific coordinates highlighting dissimilar and similar areas. Visualization of the PCoA results was achieved using the ggplot2 package, incorporating enhanced point size, transparency (alpha = 0.6), and a minimal theme. Text size and font were adjusted, and colors were manually set for different groups. Rectangular annotations were used to highlight the identified regions of interest.
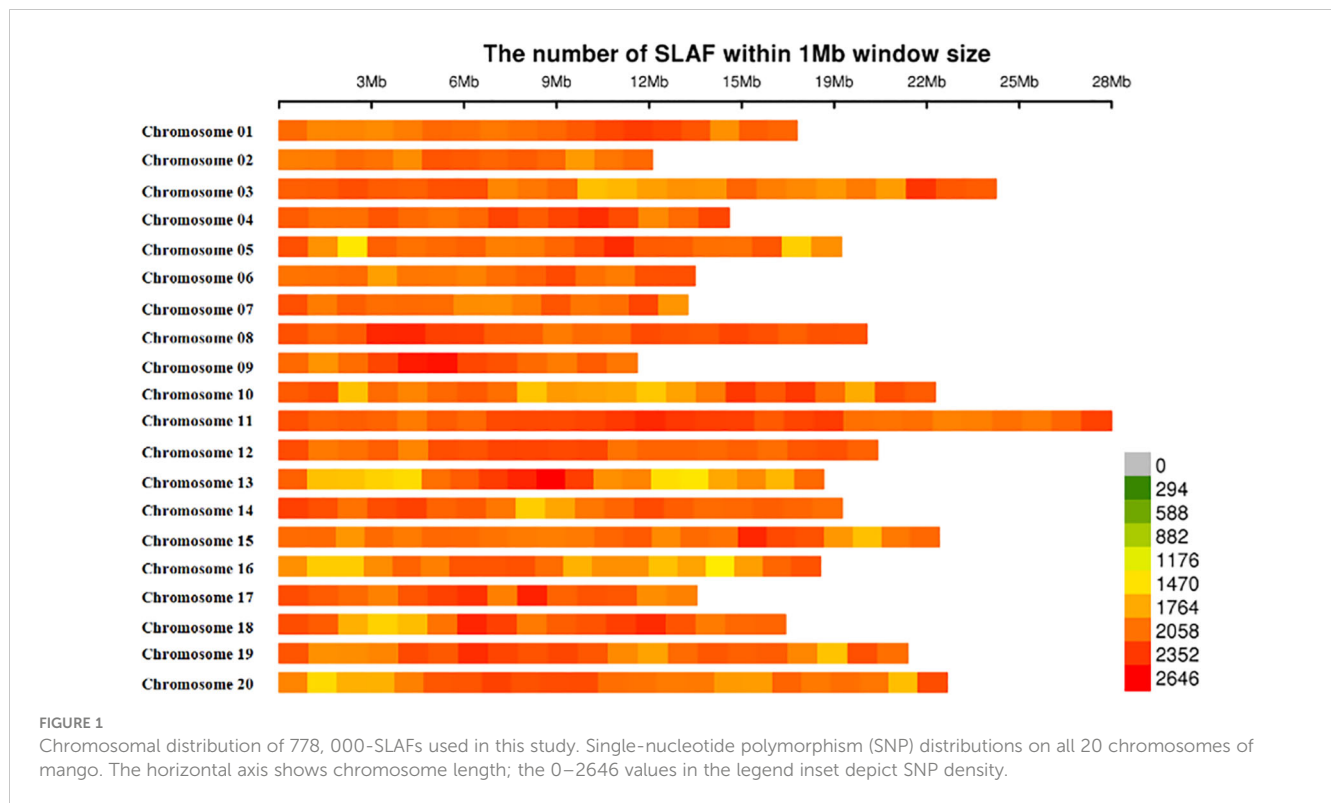
Results

SLAF-seq genotyping

Using the SLAF-seq approach developed recently, a total of 539.61-M paired-end reads for the 284 mango accessions were obtained. For the reads, on average, Q30 was 88.94% and the GC content was 36.12% (Supplementary Table S2). High-quality reads were aligned to the reference genome of mango (https://www.ncbi.nlm.nih.gov/assembly/GCA_016746415.1).

The number of reads for each mango accession was not equal, which ranged from 1,058,866 to 11,753,604, with an average of 1,900,029. After read clustering, 505,300 SLAF markers were detected and the average sequencing depth was 9.98X. Of the 505,300 SLAF markers, 205,299 SLAF markers were polymorphic, with the polymorphism rate reaching up to 40.63%, and the remaining 300,001 SLAF markers were non-polymorphic or repetitive. A total of 778,000 SLAFs were successfully invoked to evaluate the MAF and integrity based on accessions, which were distributed evenly throughout the mango genome (Supplementary Table S2; Figure 1).

Those 778,000 SLAF markers evenly distributed across all 20 chromosomes, and the number of SLAFs on each chromosome corresponds to its physical length (Figure 1; Supplementary Table S4). The largest number of SLAFs was detected on chromosome 11 (61,172 SLAFs), followed by chromosome 3 (48,775 SLAFs), whereas the smallest number of SLAFs was found on chromosome 9 (26,173 SLAFs).

Observed heterozygosity (Ho) varied from 0.14 to 0.33 with a mean of 0.22 (Supplementary Table S3), whereas the mean expected heterozygosity (He) was 0.11 with individual values per locus ranging from 0.07 to 0.16 (Supplementary Table S3). In the case of fixation indices, the minimum, maximum, and mean values for F_{ST} were 0.03, 0.72, and 0.47, respectively (Supplementary Table S3; Supplementary Figure S1).

FIGURE 1
Chromosomal distribution of 778, 000-SLAFs used in this study. Single-nucleotide polymorphism (SNP) distributions on all 20 chromosomes of mango. The horizontal axis shows chromosome length; the 0−2646 values in the legend inset depict SNP density.
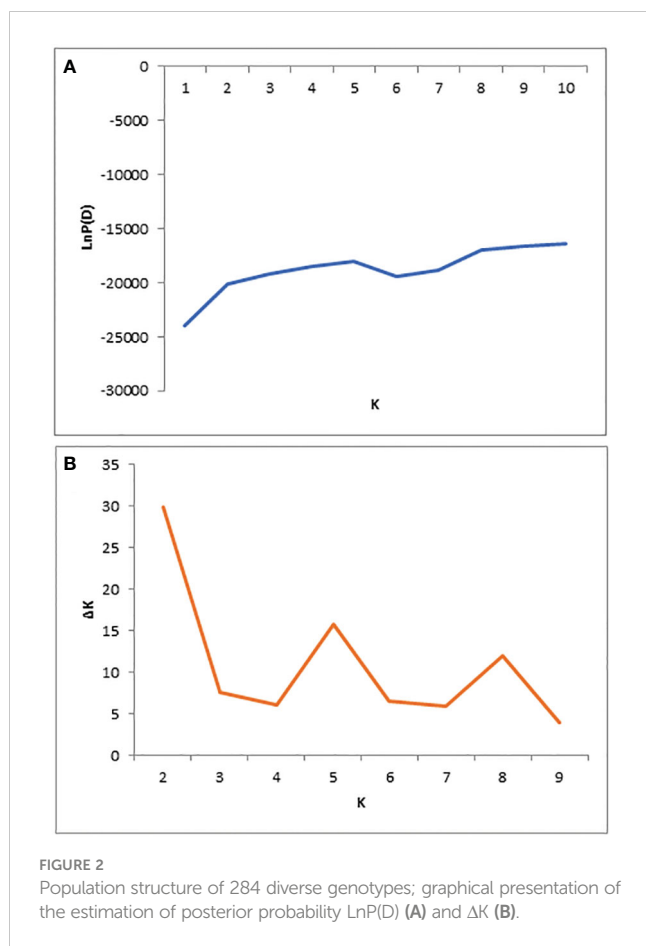
Population structure and admixture analysis

The population structure of 284 mango germplasms was evaluated with 29,136 SNP markers using STRUCTURE V2.3.4 software (Figures 2, 3). The LnP(D) score for the number of populations (K) is shown in Figure 2A; no significant inflation point was found. The peak of ΔK occurred when K was 2 (Figure 2B). According to the principle of maximum likelihood value, combined with the method of K value determination by Evanno et al. (2005), the best K value was judged to be equal to two; therefore, 284 mango germplasm resources were classified into two subpopulations (Figures 2, 3).

To understand the origin and breeding history of mango, ancestry proportions of 284 mango accessions were estimated by the Bayesian clustering algorithm. In the present study, 284 mango accessions were classified into two subpopulations based on Q (the probability that the genomic variation of a material A originates from population K) less than or equal to 0.5 (Figure 3; Supplementary Table S5). Of 284 accessions, 155 (54.58%) accessions were assigned to P1 and the remaining 129 (45.42%) accessions were assigned to P2. These results showed that the mango accessions analyzed, regardless of their current distribution, had ancestors in both populations. Admixture analysis identified 284 accessions in the attribution and showed that the gene flow of cross-pollination derived from a mixed ancestral origin (Supplementary Table S3; Figure 3). Of those, 25 accessions of P1 kept their homogeneous genetic background; these accessions were either introduced directly from India historically or hybrids of Indian varieties (Figure 3; Supplementary Table S1). Similarly, 29 accessions with a homogeneous genetic background

come from P2; these accessions were either introduced directly from India-China historically or hybrids of Indian-Chinese varieties (Figure 3; Supplementary Table S5). In addition, 130 and 100 cultivars had a mixed pedigree belonging to P1 and P2, respectively (Figure 3; Supplementary Table S5). The recent introgression from P1 into P2 indicated that some of different populations did not reflect their geographical distributions.

For K equaling two, population 1 mainly from Florida, India, Australia, Sri Lanka, Cuba, Israel, and the Caribbean had a similar composition, with a high-level ancestry of P1 and a moderate admixture level ancestry of P2. The majority of accessions in Indonesia and Southeast Asia (Philippines, Thailand, Cambodia, Malaysia, Singapore, and Vietnam) had the highest level ancestry of P2 and a moderate admixture level ancestry of P1. In addition, there were 108 mango accessions from China, 59 of whose accessions had high-level ancestry from P1, and the remaining 49 mango accessions had high-level ancestry from P2, with great variability in inferred ancestry across mango accessions; these results showed most of mango accessions from China with a high level of admixture from the two populations, indicative of the ongoing exchange of mango germplasm resources around the world.

Principal component analysis (PCA) was used to analyze the attribution of 284 mango accessions to better understand the relationship among 284 mango germplasm resources, and PCA scores were used to assess genetic variation (Figure 4). The first principal component (PC1) explained 31.77% of the variance, whereas the second (PC2) explained 24.30% (Figure 4). According to the first two components, the accessions were divided into two groups: Indian type (P1) and Southeast Asian type (P2). Once again, the results of the PCA were consistent with the results of the
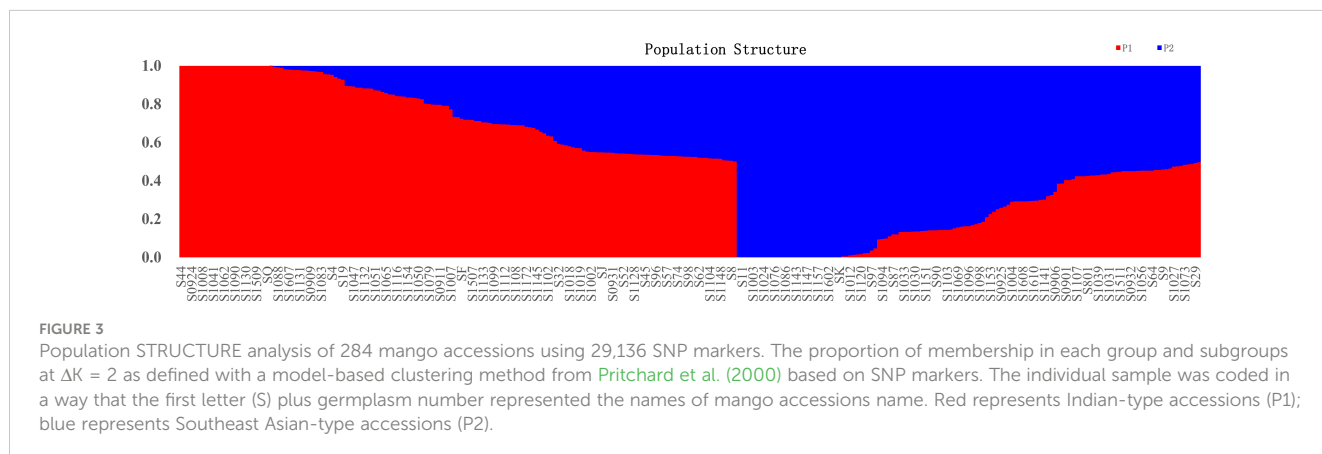
type and the Indochinese mango type, respectively. These results indicated that mango (*Mangifera indica* L.) had two places of origins: one was India, and the other was Indochina Peninsula. From the embryonic point of view, mango cultivars of the United States and Australia belonged mostly to the Indian type (P1). Both populations showed a mixture of polyembryonic and monoembryonic seed-type varieties in their cluster.

Genetic diversity and genealogy analysis

DARwin 6 software was used to calculate the genetic distance coefficients between collected mango accessions based on the proportions of shared alleles obtained from SNP markers for each accession. Genetic distances between pairs of the 284 accessions varied from 0.002 to 0.450 (on a scale of 0–1, 0 means no parentage relationship at all and 1 means the same accession), with an overall average of 0.207 (Supplementary Table S6). The 39,798 pairwise comparisons are summarized in Figure 5, where most pairs had distances ranging from 0.100 to 0.400; however, 372 pairs of genotypes (0.93%) had genetic distance ≤0.05 and 273 pairs differed by ≤2% of the total number of alleles in present analyses. Majority of pairs of genotypes (95.24%) had a genetic distance between 0.101 and 0.400. No pairs of accessions had a genetic distance over 0.500; these results suggested 284 accessions with high levels of diversity (Figure 5; Supplementary Table S6).

The genetic diversity and genetic relationship of 284 mango accessions were further elucidated using the Nei's genetic distance-based unweighted group averaging cluster analysis method (UPGMA). The present results indicated that 284 mango accessions were divided into two distinct clusters, mainly based on their geographical origin (Figure 6).

Cluster I included 141 mango accessions (C1), majority of which originated from regions and countries such as Florida, Australia, India, Cuba, Sri Lanka, Caribbean, and Israel (P1) (Supplementary Table S1; Figure 6). Notably, of 42 American mango accessions, 39 mango accessions were grouped into cluster I; similarly, 16 of 19 Indian mango accessions and 7 of 9 Australian mango accessions were grouped into cluster I, respectively. Meanwhile, cluster I contained 49 mango accessions from China. Cluster II contained 143 mango accessions (C2) originating mainly from Thailand,

population structure. The two groups can still be clearly distinguished, although a few accessions overlap between P1 and P2, following a trajectory suggestive of admixture between P1 and P2.

Comparison by horticultural type (monoembryonic vs. polyembryonic) was characterized for the 284 germplasm resources (Supplementary Table S1). Both clusters showed a mixture of polyembryonic and monoembryonic seed-type varieties in their cluster. However, most germplasm resources (99 accessions) in P1 were single-embryonic seeds, whereas majority of germplasm resources (99 accessions) in P2 were polyembryonic seeds (Supplementary Table S1), which is consistent with the characteristics of the Indian mango
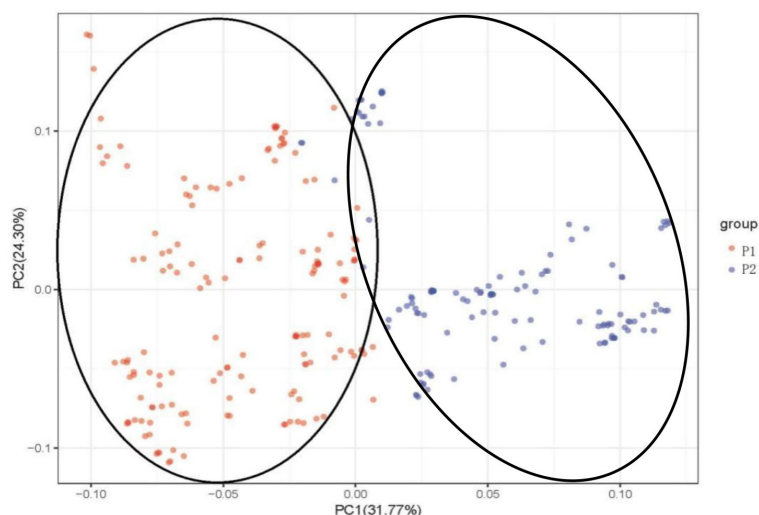
FIGURE 4
Principal component analysis (PCA) of 284 mango germplasm resources collected from different tropical regions worldwide. The percentage of variation explained by the corresponding PC is marked on the axis. Red represents Indian-type accessions (P1); blue represents Southeast Asian-type accessions (P2).

Burma, Cambodia, the Philippines, and Indonesia (P2), which also included 59 mango accessions from China. These results indicated that compared with the mango germplasm resources of Southeast Asia, American mango varieties had a closer genetic relationship with Indian mango varieties, and mango germplasm resources in China originated from the Indian type and Southeast Asian type; simultaneously, these results were consistent with the road map of mango's introduction into China in history.

As expected, it can be seen from the dendrogram that the genetic relationship between derived varieties/landraces and their parent varieties was closely similar. There were 25 of 39 American accessions grouped into cluster I-1 (Supplementary Table S1; Figure 6), of which 13 accessions resulted from seedlings of Haden or derived from Haden hybrids, the remaining accessions came from hybrids of early mango introduction in America such as 'Mulgoba,' 'Turpentine,' 'Sandesina,' and 'Bombay'. Population

structure analysis showed that these American mango varieties had a mixed and higher proportion ancestry of Indian mango type (P1) (Supplementary Table S5). Given that 13 accessions had a much broader genetic base, and the ancestors of the world's most popular commercial varieties (Olano et al., 2005; Singh et al., 2016), these 13 accessions possessed very unique genes involved with a wide range of adaptability and production stability in growing regions around the world (Mukherjee and Litz, 2009). For this reason, we might refer that the American varieties in cluster I-1 constituted a heterotic group (H1). Exceptionally, only a mango accession 'Edward' was grouped into cluster II (Southeast Asian type, P2); pedigree analysis indicated that 'Edward' derived from 'Haden'× 'Carabao', a hybrid between P1 and P2 selected under south Florida conditions. Population structure analysis also showed that 'Edward' had a mixed ancestry, 54.6% Southeast Asian-type ancestry (P2) and 45.4% Indian-type ancestry (P1). The results indicated that 'Edward' was on the
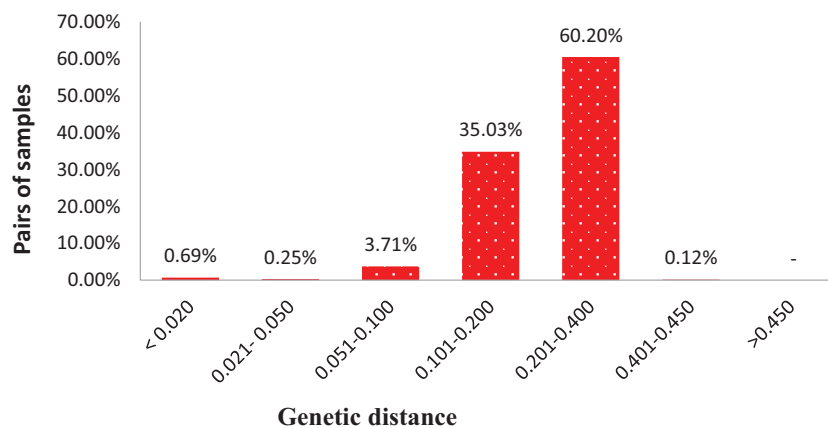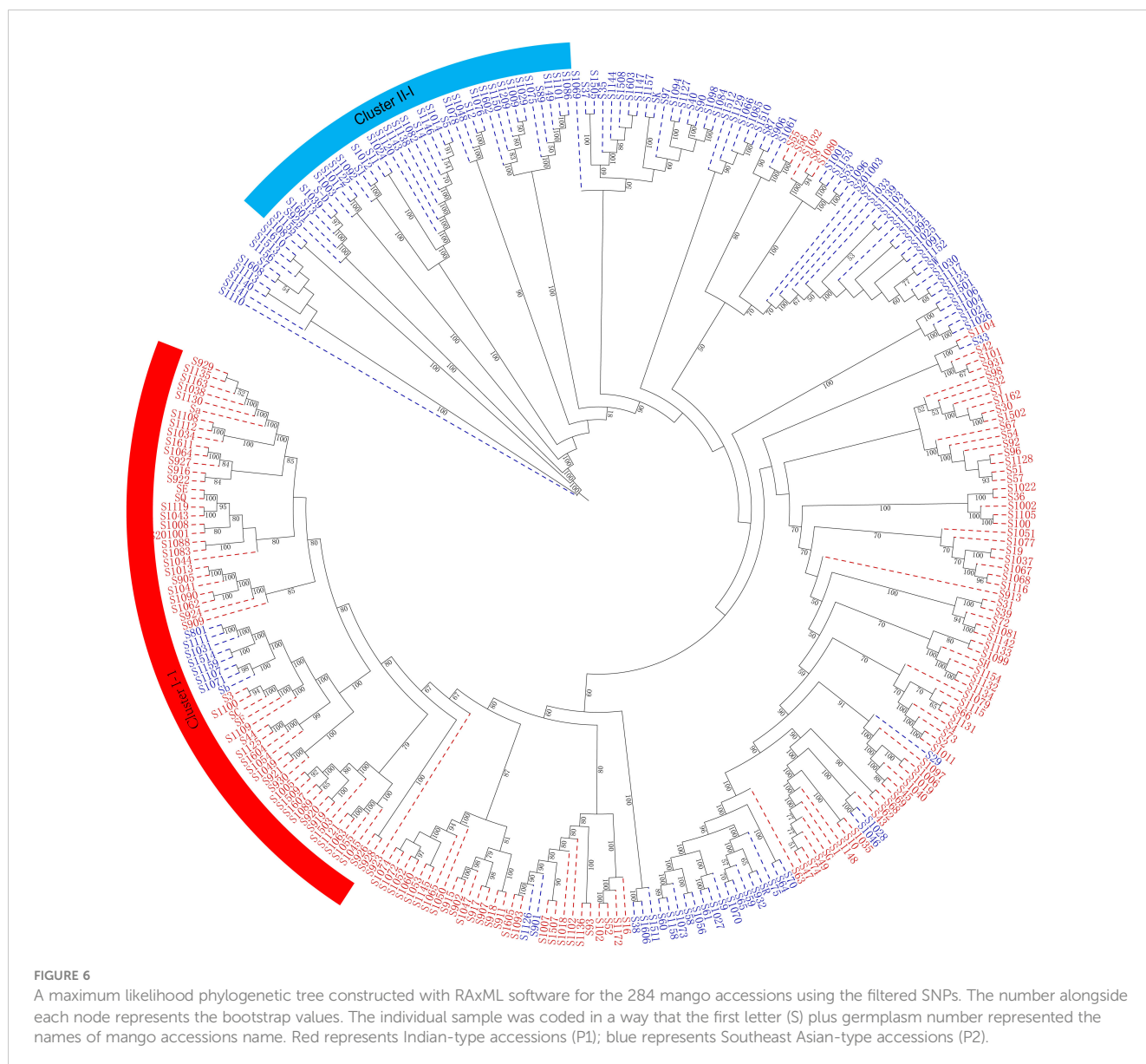

FIGURE 5
Distribution of pairwise identity-by-state based genetic distances among pairs of 284 mango accessions.

FIGURE 6
A maximum likelihood phylogenetic tree constructed with RAxML software for the 284 mango accessions using the filtered SNPs. The number alongside each node represents the bootstrap values. The individual sample was coded in a way that the first letter (S) plus germplasm number represented the names of mango accessions name. Red represents Indian-type accessions (P1); blue represents Southeast Asian-type accessions (P2).

borderline of the genetic population between cluster I and cluster II. Meanwhile, another heterotic group was found in cluster II and 25 of 39 Thailand mango accessions were grouped into cluster II-1; a large number (N = 21) of new mango varieties had been selected using seedling breeding from those of Thailand mango germplasm resources and promoted in China. In particular, a series of varieties 'Guire No. 3,' 'Guire No. 7,' 'Guire No. 10,' 'Yuanjiang_ivory,' 'Haibao 1,' 'Guire 71,' 'Hongxiangya,' 'Jinsui,' 'Wugongci Mango,' 'Guire 11,' and 'Hwagnyu (503)' were selected from Thailand mango accessions and were released and promoted in China (Supplementary Table S1); from this perspective, Thailand mango germplasms probably constituted another heterotic group (H2). Interestingly, several important commercial mango varieties were bred by Chinese mango breeders through cross breeding between American (H1) and Thailand varieties (H2); these commercial mango varieties included 'Jinhuang,' 'Yiwen,' 'Guixiang,' and 'Guifei'; those were evenly distributed in the phylogenetic tree (Figure 6; Supplementary Table S1).

Discussion

Development of high-throughput sequencing markers

The development of abundant and reliable molecular markers was extremely important for mango germplasm resource evaluation and breeding. In the present study, 29,136 SNP markers were used to dissect the population structure, principal component analyses (PCA), and genetic diversity for 284 mango accessions. SLAF-seq technology was a newly developed technique to identify SNP markers in recent years (Kashkush et al., 2001; Sun et al., 2013). Since the development of SLAF-seq technology, it has been applied in many plants studies and has achieved remarkable achievements. Shen et al. (2017) employed SLAF-seq to identify a large number of SLAF markers. Interspecific variation was identified using these markers, and these results were useful for cotton genetics research

and molecular breeding. Yang et al. (2020) applied SLAF-seq to develop 8,738 polymorphic SLAFs and resistance genes of soybean mosaic virus strain SC9 identified from cultivar Tianlong No. 1. Approximately 148 target genes were located on chromosome 2. Similar studies were carried out in many plants, such as Japanese plum (Zhang et al., 2020), Zicaitai (Li et al., 2019), eggplant (Wei et al., 2020), Chinese Lou onion (Fang et al., 2020), and Chinese elm (Lyu et al., 2020). Compared with traditional methods, this kind of marker developed by SLAF-seq technology has the advantages of higher density, consistency and effectiveness, and lower cost. In previous studies, RAPD (Schnell et al., 1995; Rajwana et al., 2008), AFLP (Yamanaka et al., 2006), ISSR (Singh et al., 2007), SSR (Shamili et al., 2012), SCoT (Zhou et al., 2020) had been used to dissect the genetic diversity of mango. However, compared with genome-level sequencing, previous molecular markers developed using the traditional method had a lower accuracy and resolution, as well as limited number of polymorphic loci, which hindered genetic research on mango. Recently, Kuhn et al. (2019) used SNP markers to genotype 1,915 mango accessions and estimate genetic diversity and relatedness; however, only 272 SNP markers were polymorphic and the number of markers was relatively small and could not cover the whole genome, which resulted in relatively large average distances between the adjacent markers (Sherman et al., 2015). Enough unbiased SNP assessments would accurately reflect the genome-wide diversity that occurred in natural populations (Zhang et al., 2020).

In the present study, 539.61 M reads were obtained for analysis of the genetic diversity of 284 mango germplasm resources. A total of 1,272,446 SLAFs were detected, of which 156,368 were polymorphic. Finally, 29,136 SLAFs were employed to dissect population structure, principal component analyses (PCA), and genetic diversity. There were more polymorphism markers in mango developed through SLAF-seq technology than the traditional marker technology, which further suggested the potential of SLAF-seq as a low-cost technology to effectively develop a large number of reliable molecular markers in mango.

Genetic diversity and population structure of mango

The importance of understanding the genetic diversity of mango accessions in various tropical regions worldwide is critical for conservation and utilization of mango germplasm resources and assisting breeders to attain elite varieties in the mango breeding programs.

Although genetic diversity of mango germplasm resources had been previously reported using different types of markers (Schnell et al., 1995; Yamanaka et al., 2006; Singh et al., 2007; Rajwana et al., 2008; Shamili et al., 2012; Zhou et al., 2020), Schnell et al. (1995) suggested that a single RAPD marker cannot distinguish all materials and a combination of two or more markers can effectively distinguish different materials. Yamanaka et al. (2006) dissected four *Mangifera* species using AFLP markers; their results showed that AFLP markers grouped 35 materials into four subgroups, which could effectively distinguish four mango species,

and four subgroups were consistent with four species. AFLP markers also clearly revealed the genetic diversity between and within intraspecific and interspecific hybrids of mango. Samal et al. (2012) analyzed the genetic diversity of Indian mango using RAPD and ISSR markers; the results showed that 65 mango germplasm resources were clustered into eight groups, consistent with their pedigree relationships. Shamili et al. (2012) studied the genetic diversity of 41 Iranian mango germplasm resources using 16 SSR markers; cluster analysis showed that Iranian mango originated from India and Pakistan. Zhou et al. (2020) suggested that 168 mango germplasm resources were grouped into two major clusters using SCoT markers; the genetic diversity within populations was much higher than that between populations. There were 34 germplasm resources identified, most of which gathered with their parents. These results indicated that SCoT markers were useful for identification and genetic diversity analyses of mango germplasms. However, previous studies were conducted with only a few markers and limited mango accessions, too few to sufficiently elucidate the genetic diversity of mango accessions. Genetic diversity analysis was a key step in the discovery of alleles that can be used as a source of excellent traits such as high-yielding, resistant to abiotic or biotic stress (Alemu et al., 2020).

The Bayesian model-based structure analysis revealed the presence of two populations among 284 mango accessions. Of the 284 mango accessions analyzed, 155 (54.58%) accessions were assigned to P1 and the remaining 129 (45.42%) accessions were assigned to P2. These results indicated that the two populations of mango analyzed, regardless of their current geographical distribution, had ancestors in mutual populations, these results were similar to that of the phylogenetic tree and PCA analyses; meanwhile, PCoA (Supplementary Figure S2) was also performed and the result was consistent with those of PCA. Our result was further consistent with the two domestication centers that have been proposed for a long time: one was in India, and the other was in Southeast Asia (BJaS, 1997; Schnell et al., 2006; Wang et al., 2020). Admixture analysis identified 284 mango accessions showing gene flow from cross-pollination of mixed ancestral origin (Supplementary Table S5; Figure 3). Of which, 25 mango accessions of P1 kept their homogeneous genetic background; these accessions were either introduced directly from India historically, or hybrids of Indian varieties (Figure 4; Supplementary Table S4). Similarly, there were 29 mango accessions with a homogeneous genetic background coming from P2; these accessions were either introduced directly from India-China historically or hybrids of Indian-China varieties (Figure 3; Supplementary Table S5). In addition, 130 and 100 cultivars had a mixed pedigree belonging to P1 and P2, respectively (Figure 3; Supplementary Table S5). The recent introgression from P1 into P2 and vice versa had been so intensive that some of different populations did not reflect their geographical distributions. These results were consistent with previous studies (Warschefsky and Wettberg, 2019; Wang et al., 2020). Warschefsky and Wettberg (2019) identified two gene banks for cultivated mango accessions and further identified them as Indian or Southeast Asian types, but there was no significant bottleneck between the two types of gene banks. Their results indicated that mangoes had a more complicated domestication history than previously speculated. Wang et al.

(2020) reported that the genome resequencing revealed two different groups of mango germplasms; commercial varieties were clustered together with Indian germplasms, which showed an allelic admixture. Southeast Asian native varieties were in the second group. Chinese varieties had formed different branches, and some varieties showed admixture in the genome. In the present study, 25 and 29 accessions kept their homogeneous genetic background in P1 and P2, respectively, and most of accessions (80.99%) were of mixed ancestry. This mixture may be the result of human hybridization, domestication, and selection, spread through human migration or trade, which had a great impact on the diversity of population structure.

Insight into mango breeding history and breeding strategy

Understanding the source of cultivars is very important for clarifying the breeding history of cultivars, assessing genetic diversity and promoting breeding strategies. However, the spread of germplasm resources and the breeding of new varieties may be very complicated and difficult to understand, because many germplasm resources spread and the breeding process of variety previously had no detailed historical records (Nishio et al., 2014).

The mango breeding program in China has gone through two stages: the first stage was that of seedling breeding, in which a large number of original varieties were selected from seedling of Indian or Southeast Asian type varieties (P1 or P2); the second one was that of cross breeding, in which the breeding of artificial cross between P1 and P2 had produced many excellent varieties, which had been widely promoted in China (Luo et al., 2010; Huang et al., 2013; 2012).

Molecular genetics studies on the spread process of cultivars had been carried out in previous studies (BJaS, 1997; Olano et al., 2005). Parentage studies have shown that sexual cross breeding has played an important role in the emergence of new varieties in the past, and these excellent varieties had been adopted and spread by vegetative propagation (Bowers, 1999; Crespan et al., 2008). As expected, in the current research, the genetic relationship between seedlings selected cultivars or offspring of hybrid and their parent cultivars was closely related according to the dendrogram (Figure 6; Supplementary Table S1).

At the beginning of the twentieth century, the introduction of mango germplasm resources and the breeding programs of new varieties were strengthened in USA, especially in South Florida, where many of the most important commercial cultivars were released; many of those were still cultivated in major tropical regions around the world. The elite characteristics of these cultivars and their success all over the world had made South Florida known as the second center of domestication (Knight and Schnell, 1994). There were 25 of 39 America accessions grouped into cluster I-1 (Supplementary Table S1; Figure 6), of which 13 accessions derived from seedlings of 'Haden' or derived from 'Haden' hybrids; the remaining accessions came from hybrids of early mango introduction in America such as 'Mulgoba,' 'Turpentine,' 'Sandesina,' and 'Bombay'. Population structure analysis showed that these American mango varieties had a mixed and higher proportion ancestry of Indian mango type (Supplementary Table S1;

these results indicated that the USA varieties were more closely related to Indian varieties than to Southeast Asian varieties, in agreement with previous analysis using 25 microsatellite loci (BJaS, 1997). Given that these varieties had a much broader genetic base and were the ancestors of the world's most popular commercial varieties (Knight and Schnell, 1994; Olano et al., 2005; Singh et al., 2016), these commercial mango varieties possess very unique genes involved with a wide range of adaptability and production stability in growing regions around the world (Mukherjee, 1997). For this reason, we may infer that the American varieties in cluster I-1 constitute a heterotic group (H1). Exceptionally, accession 'Edward' was grouped into cluster II (Southeast Asian type); the result might represent human disturbance such as cross between cluster I and cluster II or production and movement of seeds; as expected, pedigree results indicated that this accession derived from 'Haden' × 'Carabao', a hybrid between Indian types (P1) and the Southeast Asian types (P2) selected under south Florida conditions; population structure analysis showed that 'Edward' has a mixed ancestry, 54.6% Southeast Asian type ancestry and 45.4% Indian type ancestry. The results indicated that 'Edward' was on the borderline of the genetic population between P1 and P2.

During the first stage (seedling breeding) of the mango breeding program in China, a large number of original varieties were selected from seedlings of Indian or Southeast Asian type varieties (P1 or P2); however, only a small number of selected varieties have been widely planted, of which 25 of 39 Thailand mango accessions were grouped into cluster II-1; a large number (N = 21) of new mango varieties had been selected using seedling breeding from Thai mango germplasm resources which were released and promoted in China, such as varieties 'Guire No. 3,' 'Guire No. 7,' 'Guire No. 10,' 'yuanjiang_ivory,' 'haibao 1,' 'guire 71,' 'hongxiangya,' 'Jinsui,' 'Wugongci Mango,' 'Guire 11,' and 'Hwagnyu (503)' (Supplementary Table S1). Furthermore, Thailand mango accessions contain unique genetic diversity compared with other mango accessions (Warschefsky and Wettberg, 2019); from this perspective, Thailand mango germplasms may constitute another heterotic group (H2).

Varieties cultivated widely were facing common problems of degradation with longer cultivation, poorer resistance to pests and diseases, poor fruit quality, and sensitivity to abiotic stress (Schulze-Kaysers et al., 2015). It is important to expand the genetic base by introducing more exotic germplasm resources. Over the last decades, Chinese breeders had successfully bred several superior commercial mango varieties through cross breeding (Supplementary Table S1). The present results showed that mango accessions might be divided into two groups, which was consistent with the two proposed centers of domestication, and allelic admixture was observed in the genomes of commercial varieties. Interestingly, several important commercial mango varieties have been selected through cross breeding between American (H1) and Thailand mango accessions (H2) in China such as 'Jinhuang,' 'Yiwen,' 'Guixiang,' and 'Guifei' which were evenly distributed in the phylogenetic tree (Figure 6; Supplementary Table S1); of those, 'Guifei' was the most popular commercial variety in Taiwan, Hainan, and Guangxi provinces, and 'Jinhuang' was the most popular commercial variety in Taiwan, Yunnan, and Sichuan provinces in China. These results indicated that new varieties through cross breeding combine the advantages of two heterosis groups and have elevated levels of diversity.

Conclusions

The exchange between the two groups (P1 and P2) elevated levels of diversity of mango accessions; the results of their genetic diversity and pedigree indicated the blood relationship of the main commercial varieties in China. Our results indicated that the proposed heterotic group existed in P1 and P2, respectively; hybridization occurred widely between P1 and P2, and most of accessions (80.99%) were of mixed ancestry, perhaps including multiple hybridization events and regional selection. The present work may have direct implications of a new strategy for mango breeding and germplasm management, which merits further investigation.

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found below: Bioproject accession number: PRJNA1037345.

Author contributions

QL: Funding acquisition, Writing – original draft. HP: Writing – original draft, Data curation. XH: Data curation, Formal analysis, Writing – original draft. SW: Conceptualization, Data curation, Writing – original draft. YH: Data curation, Formal analysis, Investigation, Writing – original draft. HX: Writing – original draft. GX: Conceptualization, Data curation, Formal analysis, Investigation, Writing – original draft. RY: Writing – original draft. DL: Data curation, Investigation, Resources, Writing – original draft. ZY: Supervision, Writing – review & editing.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2024.1328126/full#supplementary-material

References

Alemu, A., Feyissa, T., Letta, T., and Abeyo, B. (2020). Genetic diversity and population structure analysis based on the high density SNP markers in Ethiopian durum wheat (*Triticum turgidum* ssp. durum). *BMC Genet.* 21, 18. doi: 10.1186/s12863-020-0825-x

Alexander, D. H., Novembre, J., and Lange, K. (2009). Fast model-based estimation of ancestry in unrelated individuals. *Genome Res.* 19, 1655–1664.

BJaS, R. J. (1997). "The mango: botany, production and uses," in *The mango: botany, production and uses.* Ed. L. RE (CAB International, Wallingford), 27.

Boeven, P. H. G., Longin, C. F. H., and Würschum, T. (2016). A unified framework for hybrid breeding and the establishment of heterotic groups in wheat. *Theor. Appl. Genet.* 129, 1231–1245. doi: 10.1007/s00122-016-2699-x

Bompard, J. (2009). "Taxonomy and systematics," in *The mango: Botany, production and uses.* Ed. R. E. Litz (CAB International, Wallingford), 19–41.

Bowers, J. (1999). Historical genetics: the parentage of chardonnay, gamay, and other wine grapes of northeastern France. *Science* 285, 1562–1565. doi: 10.1126/science.285.5433.1562

Bruce, R. W., Torkamaneh, D., Grainger, C., Eskandari, M., and Rajcan, I. (2019). Genome-wide genetic diversity is maintained through decades of soybean breeding in Canada. *Theor. Appl. Genet.* 132, 3089–3100. doi: 10.1007/s00122-019-03408-y

Chang, Y. S., He, P., Wang, H. B., Li, H. F., Wang, S., and Li, L. G. (2019). Application of high throughput sequencing to evaluate the genetic diversity among wild apple species indigenous to Shandong, China, and introduced cultivars. *Plant Mol. Biol. Rep.* 37, 63–73. doi: 10.1007/s11105-019-01138-5

Crespan, M., Calo, A., Giannetto, S., Sparacio, A., Storchi, P., and Costacurta, A. (2008). 'Sangiovese' and 'Garganega' are two key varieties of the Italian grapevine assortment evolution. *Vitis* 47, 97–104.

Cui, J. J., Yang, Y., Luo, S. B., Wang, L., Huang, R. K., Wen, Q. F., et al. (2020). Whole-genome sequencing provides insights into the genetic diversity and domestication of bitter gourd (*Momordica* spp). *Hortic. Res.* 7, 85. doi: 10.1038/s41438-020-0305-5

de Hoon, M. J., Imoto, S., Nolan, J., and Miyano, S. (2004). Open source clustering software. *Bioinformatics* 20, 1453.

Dunteman, G. H. (1989). *Principal Components Analysis. Number 69.* (Sage). doi: 10.4135/9781412985475

Evanno, G., Regnaut, S., and Goudet, J. (2005). Detecting the number of clusters of individuals using the software structure: A simulation study. *Mol. Ecol.* 14, 2611–2620. doi: 10.1111/j.1365-294X.2005.02553.x

Excoffier, L., and Lischer, H. L. (2010). Arlequin suite ver 3.5: a new series of programs to perform. population genetics analyses under Linux and Windows. *Mol. Ecol. Resour.* 10, 564–567. doi: 10.1111/j.1755-0998.2010.02847.x

Fang, H., Liu, H., Ma, R., Liu, Y., Li, J., Yu, X., et al. (2020). Genome-wide assessment of population structure and genetic diversity of Chinese Lou onion using specific length amplified fragment (SLAF) sequencing. *PloS One* 15, e0231753. doi: 10.1371/journal.pone.0231753

Gao, A., Chen, Y., Crane, J. H., Zhu, M., Huang, J., and Luo, H. (2011). Status and analysis on mango production in China, advances in biomedical engineering. *International Conference on Agricultural and Biosystems Engineering* 1-2, 472–476.

He, X. H., Li, Y. R., Guo, Y. Z., Tang, Z. P., and Li, R. B. (2005). Genetic analysis of 23 mango cultivar collection in Guangxi province revealed by ISSR. *Mol. Plant Breed* 3, 829–834.

Huang, G. D., Zhao, Y., Li, R. W., Zhou, J. N., and Chen, Y. S. (2013). Research status and strategy of mango germplasm resources and cultivars breeding in Guangxi province. *China Trop. Agric.* 4, 47–49.

Jones, H., Gosman, N., Horsnell, R., Rose, G. A., Everest, L. A., Bentley, A. R., et al. (2013). Strategy for exploiting exotic germplasm using genetic, morphological, and environmental diversity: the Aegilops tauschii Coss. *Example Theor. Appl. Genet.* 126, 1793–1808. doi: 10.1007/s00122-013-2093-x

Kashkush, K., Fang, J., Eli, T., Hillel, J., and Lavi, U. (2001). Cultivar identification and genetic map of mango (*Mangifera indica*). *Euphytica* 1221, 129–136. doi: 10.1023/A:1012646331258

Knight, R. J., Campbell, R. J., and Maguire, I. (2009). "Important mango cultivars and their descriptions," in *The mango: botany, production and uses*. Ed. R. E. Litz (CAB International, Wallingford, UK), 42–66.

Knight, R. J., and Schnell, R. J. (1994). Mango introduction in Florida and the 'Haden' cultivar's significance to the modern industry. *Economic Bot.* 48, 139–145. doi: 10.1007/BF02908201

Kuhn, D. N., Dillonb, N., Ballyb, I., Grohc, A., Rahamanc, J., Warschefskyc, E., et al. (2019). Estimation of genetic diversity and relatedness in a mango germplasm collection using SNP markers and a simplified visual analysis method. *Scientia Hortic.* 252, 156–168. doi: 10.1016/j.scienta.2019.03.037

Lal, S., Singh, A. K., Singh, S. K., Srivastav, M., Singh, B. P., Sharma, N., et al. (2017). Association analysis for pomological traits in mango (*Mangifera indica* L.) by genic-SSR markers. *Trees Struct. Funct.* 31, 1391–1409. doi: 10.1007/s00468-017-1554-2

Li, G. H., Chen, H. C., Liu, J. L., Luo, W. L., Xie, D. S., Luo, S. B., et al. (2019). A high-density genetic map developed by specific-locus amplified fragment (SLAF) sequencing and identification of a locus controlling anthocyanin pigmentation in stalk of Zicaitai (*Brassica rapa* L. ssp. chinensis var. purpurea). *BMC Genomics* 20, 343. doi: 10.1186/s12864-019-5693-2

Li, H., and Durbin, R. (2009). Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics* 25, 1754–1760. doi: 10.1093/bioinformatics/btp324

Liang, Q., Song, K., Lu, M., Dai, T., Yang, J., Wan, J., et al. (2022). Transcriptome and metabolome analyses reveal the involvement of multiple pathways in flowering intensity in mango. *Front. Plant Sci.* 13, 933923. doi: 10.3389/fpls.2022.933923

Litz, R. E. (2009). *The Mango: Botany, Production and Uses* (U.S.A: CABI), pp.10.

Liu, X. Y., Xiao, Y. L., Zi, Y., Yan, J., Li, C. H., Du, C. X., et al. (2023). Differential effects of low and high temperature stress on pollen germination and tube length of mango (*Mangifera indica* L.) genotypes. *Sci. Rep.* 13, 611. doi: 10.1038/s41598-023-27917-5

Lora, J., and Hormaza, J. I. (2018). Pollen wall development in mango (Mangifera indica L., Anacardiaceae). *Plant Reprod.* 31, 385–397. doi: 10.1007/s00497-018-0342-5

Luo, C., He, X. H., Chen, H., Hu, Y., and Ou, S. J. (2012). Genetic relationship and diversity of Mangifera indica L.: revealed through SCoT analysis. *Genet. Resour. Crop Evol.* 59, 1505–1515. doi: 10.1007/s10722-011-9779-1

Luo, C., He, X. H., Chen, H., Ou, S. J., and Gao, M. P. (2010). Analysis of diversity and relationships among mango cultivars using Start Codon Targeted (SCoT) markers. *Biochem. System. Ecol.* 38), 1176–1184. doi: 10.1016/j.bse.2010.11.004

Luo, C., Shu, B., Yao, Q., Wu, H., Xu, W., and Wang, S. (2016). 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.* 7, 1310. doi: 10.3389/fpls.2016.01310

Lyu, Y. Z., Dong, X. Y., Huang, L. B., Zheng, J. W., He, X. D., Sun, H. N., et al. (2020). SLAF-seq uncovers the genetic diversity and adaptation of Chinese elm (*Ulmus parvifolia*) in eastern China. *Forests* 11, 80. doi: 10.3390/f11010080

Ma, X., Luo, X., Wei, Y., Bai, T., Shi, J., Zheng, B., et al. (2021). Chromosome-scale genome and comparative transcriptomic analysis reveal transcriptional regulators of β-carotene biosynthesis in mango. *Front. Plant Sci.* 12, 749108. doi: 10.3389/fpls.2021.749108

Mango Genome Consortium, Bally, I. S. E., Bombarely, A., Chambers, A. H., Cohen, Y., Dillon, N. L., et al. (2021). The 'Tommy Atkins' mango genome reveals candidate genes for fruit quality. *BMC Plant Biol.* 21, 108. doi: 10.1186/s12870-021-02858-1

McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A., et al. (2010). The Genome Analysis Toolkit: a map reduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* 20, 1297–1303. doi: 10.1101/gr.107524.110

Mukherjee, S. K. (1949). The mango and its wild relatives. *Sci. Culture* 26, 5–9.

Mukherjee, S. K. (1997). *The mango-botany, production and uses*. Ed. R. E. Litz (Wallingford, UK: CABI Publishing), 669.

Mukherjee, S., and Litz, R. (2009). "Introduction: botany and importance," in *The mango: Botany, production and uses*, *2nd ed*. Ed. R. E. Litz (Wallingford: CAB international), 1–18.

Nadeem, M. A., Nawaz, M. A., Shahid, M. Q., Doǧan, Y., Comertpay, G., Yildiz, M., et al. (2017). DNA molecular markers in plant breeding: current status and recent advancements in genomic selection and genome editing. *Biotechnol. Biotechnol. Equip.* 32 (2), 261–285. doi: 10.1080/13102818.2017.1400401

Nishio, S., Yamada, M., Takada, N., Kato, H., Onoue, N., Sawamura, Y., et al. (2014). Environmental variance and broad-sense heritability of nut traits in Japanese chestnut breeding. *Hortscience* 49, 696–700. doi: 10.21273/HORTSCI.49.6.696

Niu, Y., Gao, C., and Liu, J. (2021). Comparative analysis of the complete plastid genomes of *Mangifera* species and gene transfer between plastid and mitochondrial genomes. *Peer J.* 9, e10774. doi: 10.7717/peerj.10774

Olano, C., Schnell, R. J., Quintanilla, W., and Campbell, R. J. (2005). Pedigree analysis of Florida mango cultivars. *Proc. Fla. State Hortic. Soc* 118, 192–197.

Popenoe, W. (1920). *The mango. Manual of tropical and subtropical fruits* (New York, NY, USA: Macmillan)), 79–145.

Pritchard, J. K., Stephens, M., and Donnelly, P. (2000). Inference of population structure using multilocus genotype data. *Genetics* 155, 945–959. doi: 10.1093/genetics/155.2.945

Rajwana, I. A., Tabbasam, N., Malik, A. U., Malik, S. A., Rahman, M., and Zafar, Y. (2008). Assessment of genetic diversity among mango (*Mangifera indica* L.) genotypes using RAPD markers. *Scientia Hortic.* 117, 297–301. doi: 10.1016/j.scienta.2008.04.009

Saitou, N., and Nei, M. (1987). The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.* 4, 406–425.

Samal, K. C., Jena, R. C., Swain, S. S., Das, B. K., and Chand, P. K. (2012). Evaluation of genetic diversity among commercial cultivars, hybrids and local mango (*Mangifera indica* L.) genotypes of India using cumulative RAPD and ISSR markers. *Euphytica* 185, 195–213. doi: 10.1007/s10681-011-0522-y

Schnell, R. J., Olano, C. T., Meerow, A. W., Campbell, R. J., and Kuhm, D. N. (2006). Mango genetic diversity analysis and pedigree inferences for Florida cultivars using microsatellite. *Am. Soc. Hortic. Sci.* 13, 11.

Schnell, R. J., Ronning, C. M., and Knight, R. J. (1995). Identification of cultivars and validation of genetic relationships in *Mangifera indica* L. using RAPD markers. *Theor. Appl. Genet.* 90, 269–274. doi: 10.1007/BF00222213

Schulze-Kaysers, N., Feuereisen, M. M., and Schieber, A. (2015). Phenolic compounds in edible species of the anacardiaceae family a review. *RSC Adv.* 5, 14. doi: 10.1039/C5RA11746A

Shamili, M., Fatahi, R., and Hormaza, J. I. (2012). Characterization and evaluation of genetic diversity of Iranian mango (*Mangifera indica* L., Anacardiaceae) genotypes using microsatellites. *Scientia Hortic.* 148, 230–234. doi: 10.1016/j.scienta.2012.09.031

Shen, C., Jin, X., Zhu, D., and Lin, Z. X. (2017). Uncovering SNP and indel variations of tetraploid cottons by SLAF-seq. *BMC Genomics* 18, 247. doi: 10.1186/s12864-017-3643-4

Sherman, A., Rubinstein, M., Eshed, R., Benita, M., Ish-Shalom, M., Sharabi-Schwage, M., et al. (2015). Mango (*Mangifera indica* L.) germplasm diversity based on single nucleotide polymorphisms derived from the transcriptome. *BMC Plant Biol.* 15, 277. doi: 10.1186/s12870-015-0663-6

Singh, S., Karihaloo, J. L., and Gaikwad, A. B. (2007). DNA fingerprinting of some mango (*Mangifera indica* L) cultivars using anchored-ISSR markers. *Plant Biochem. Biotechnol.* 16, 113–117. doi: 10.1007/BF03321984

Singh, N. K., Mahato, A. K., Jayaswal, P. K., Singh, A., and Singh, S. (2016). Origin, diversity and genome sequence of mango (*Mangifera indica* L.). *Indian J. History Sci.* 51 (2), 355–368. doi: 10.16943/ijhs/2016/v51i2.2/48449

Sun, X., Liu, D., Zhang, X., Li, W., Liu, H., Hong, W., et al. (2013). SLAF-seq: an efficient method of large-scale *de novo* SNP discovery and genotyping using high-throughput sequencing. *PloS One* 8, e58700. doi: 10.1371/journal.pone.0058700

Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M., and Kumar, S. (2011). MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol. Biol. Evol.* 28, 2731–2739. doi: 10.1093/molbev/msr121

Wang, P., Luo, Y. F., Huang, J. F., Gao, S., Zhu, G., Dang, Z., et al. (2020). The genome evolution and domestication of tropical fruit mango. *Genome Biol.* 21, 60. doi: 10.1186/s13059-020-01959-8

Warschefsky, E. J., and Wettberg, E. J. B. (2019). Population genomic analysis of mango (*Mangifera indica*) suggests a complex history of domestication. *New Phytol.* 222, 2023–2037. doi: 10.1111/nph.15731

Wei, Q., Wang, W., Hu, T., Hu, H., Wang, J., and Bao, C. (2020). Construction of a SNP-based genetic map using SLAF-seq and QTL analysis of morphological traits in eggplant. *Front. Genet.* 11, 178. doi: 10.3389/fgene.2020.00178

Wu, Y. S., Vicente, F. S., Kaijian Huang, K., Dhliwayo, T., Costich, D. E., Semagn, K., et al. (2016). Molecular characterization of CIMMYT maize inbred lines with genotyping–by–sequencing SNPs. *Theor. Appl. Genet.* 129, 753–765. doi: 10.1007/s00122-016-2664-8

Xu, S. P. (1992). Mango germplasm in Hainan Island, agricultural press of China, 273–279 (In Chinese).

Yamanaka, N., Hasran, M., Xu, D. H., Tsunematsu, H., Salma Idris, H., and Ban, Y. (2006). Genetic relationship and diversity of four *mangifera* species revealed through AFLP analysis. *Genet. Resour. Crop Evol.* 53, 949–954. doi: 10.1007/s10722-004-6695-7

Yang, H. B. (2019). "Analysis of the historical significance of Zheng He's voyages to the western seas," in *2019 9th International Conference on Education and Social Science*

(ICESS 2019) Copyright ©. (UK: Francis Academic Press), 253–255. doi: 10.25236/icess.2019.049

Yang, Q., Jin, H., Yu, X., Fu, X., Zhi, H., and Yuan, F. (2020). Rapid identification of soybean resistance genes to soybean mosaic virus by SLAF-seq bulked segregant analysis. *Plant Mol. Biol. Rep.*. doi: 10.1007/s11105-020-01227-w

Zeng, B., Yan, H. D., Liu, X. C., Zang, W. J., Zhang, A. L., Zhou, S. F., et al. (2017). Genome-wide association study of rust traits in orchard grass using SLAF-seq Technology. *Hereditas* 154, 5. doi: 10.1186/s41065-017-0027-3

Zhang, Q., Wei, X., Liu, N., Zhang, Y., Xu, M., Zhang, Y., et al. (2020). Construction of an SNP-based high-density genetic map for Japanese plum in a Chinese population using specific length fragment sequencing. *Tree Genet. Genomes* 16, 18. doi: 10.1007/s11295-019-1385-y

Zhang, R., Xu, G., Li, J., Yan, J., Li, H., and Yang, X. (2018). Patterns of genomic variation in Chinese maize inbred lines and implications for genetic improvement. *Theor. Appl. Genet.* 131, 1207–1221. doi: 10.1007/s00122-018-3072-z

Zhou, L., He, X. H., Yu, H. X., Chen, M. Y., Fan, Y., Zhang, X. J., et al. (2020). Evaluation of the genetic diversity of mango (*Mangifera indica* L.) seedling germplasm resources and their potential parents with start codon targeted (SCoT) markers. *Genet. Resour. Crop Evol.* 67, 41–58. doi: 10.1007/s10722-019-00865-8