horticulturae

MDPI

Article

Phenotypic Characterization and DNA Fingerprinting of Tianbao Melon Using Genome-Wide SNPs

Yumeng Ren [1,†], Xiaofeng Su [1,†], Wenjing Dong [1], Minghe Hu [1], Houshun Ma [1], Qian Zhao [1], Wenhao Jiang [1], Shengkai Zhang [1], Sen Chai [1], Xiaoli Liu [1], Xiaofeng Liu [1], Kexiang Wang [2,*] and Kuipeng Xu [1,*]

[1] College of Horticulture, Qingdao Agricultural University, Qingdao 266109, China; yumeng07002@163.com (Y.R.); suxf@stu.qau.edu.cn (X.S.); dongyy52@163.com (W.D.); huminghe0314@163.com (M.H.); mahoushun@163.com (H.M.); zhaoqiann0516@163.com (Q.Z.); jiangwh0426@163.com (W.J.); zhangshengkai2718@163.com (S.Z.); chaisen@qau.edu.cn (S.C.); 201901196@qau.edu.cn (X.L.); 202101037@qau.edu.cn (X.L.)

[2] Hami Agricultural and Animal Husbandry Industry Development Investment Co., Ltd., Hami 839000, China

* Correspondence: laowang.110@163.com (K.W.); xukuipeng@qau.edu.cn (K.X.)

† These authors contributed equally to this work.

Abstract

The Tianbao melon (*Cucumis melo* subsp. *agrestis*) is a highly valued regional horticultural crop, yet its sustainable development is severely constrained by a narrow genetic base and widespread varietal admixture in the market. In this study, a panel of 32 Tianbao melon accessions was systematically evaluated by integrating field-based phenotypic assessment with genome-wide single-nucleotide polymorphism (SNP) analysis via whole-genome resequencing. Phenotypic analysis based on ten quantitative traits revealed low overall morphological variability, indicating limited discriminatory power of morphological traits alone. In contrast, 173,497 high-quality SNPs uncovered substantial hidden genetic differentiation, partitioning the accessions into four distinct genotypic groups. Notably, accessions TB-17 and TB-27, though nearly indistinguishable morphologically, exhibited clear genetic divergence in both phylogenetic and principal component analyses. Furthermore, a panel of 20 core SNPs with conserved flanking sequences was selected, generating unique molecular fingerprint profiles for all 32 accessions and achieving high discriminatory resolution (pairwise differences ranging from 10 to 13 SNPs). These findings demonstrate that the integration of phenotypic and genome-wide SNP data provides a robust framework for genetic diversity assessment and DNA fingerprinting in Tianbao melon, offering a scientific basis for cultivar identification, intellectual property protection, and precision breeding to support sustainable development of the regional melon industry.

Keywords: genetic diversity; DNA fingerprinting; Tianbao melon; genome-wide SNPs

Academic Editor: Larisa Garkava-Gustavsson

Received: 27 April 2026
Revised: 3 June 2026
Accepted: 7 June 2026
Published: 9 June 2026

1. Introduction

Melon (*Cucumis melo* L.) has undergone multiple independent domestication events, resulting in rich genetic diversity and extensive phenotypic differentiation [1–3]. As a highly diversified species, melon is classified into two major subspecies, namely *C. melo* subsp. *melo* and subsp. *agrestis*, each representing an independent domestication lineage originating from separate wild populations [3]. Thin-skinned melon (*C. melo* subsp. *agrestis*) is a regionally important horticultural crop with a long cultivation history and a broad consumer base [4,5]. It contributes significantly to farmers' income, regional agricultural development, and diverse demands of fruit markets. Among thin-skinned types, Tianbao

melon is particularly valued for its uniform fruit shape, thin rind, tender flesh, unique aroma, and high sweetness. It has emerged as a dominant cultivar in key production areas such as Laixi City, Shandong Province, and plays a central role in the regional melon industry. However, the rapid expansion of market demand has exposed critical challenges for the Tianbao melon industry. Prolonged reliance on a limited number of elite cultivars has narrowed the genetic base and reduced varietal diversity, thereby constraining long-term breeding progress [6]. Concurrently, the increasing prevalence of market adulteration, including the mixing of authentic and counterfeit Tianbao melons and uncertainty in cultivar provenance, has raised serious concerns related to intellectual property protection [7]. These problems have seriously affected breeding companies and eroded grower confidence. Accurate assessment of genetic diversity and reliable cultivar identification are therefore essential for ensuring varietal purity, safeguarding intellectual property, and promoting sustainable industrial development [8–10].

Large-scale genome resequencing studies have revealed that cultivated *agrestis* accessions harbor substantially lower nucleotide diversity and exhibit markedly extended linkage disequilibrium (LD) decay compared with cultivated *melo* accessions, reflecting a more severe domestication bottleneck [3,6]. The *agrestis* subspecies—encompassing groups such as *momordica*, *acidulus*, *conomon*, *makuwa*, and *chinensis*—is predominantly cultivated in East Asia and exhibits distinct agronomic traits including thinner flesh, lower sugar content, and unique ecological adaptations, yet has historically been underrepresented in genomic studies relative to *melo* [6]. Recent comprehensive omics efforts have begun to address this gap by generating a telomere-to-telomere (T2T) genome assembly of the *agrestis* accession 13C, a transcriptome atlas across 31 tissue types, an EMS-induced mutant library covering 97.33% of the annotated gene space, and a stable genetic transformation system. These data together provide an integrated platform specifically for functional genomics and precision breeding of this subspecies [5]. These resources substantially expand the molecular toolkit available for genetic improvement of thin-skinned melon [5].

Traditional germplasm evaluation has primarily relied on phenotypic characterization, including fruit morphology, rind thickness, and flavor-related traits. Although these descriptors are straightforward to measure, they are strongly influenced by environmental conditions and developmental stages, and frequently fail to discriminate closely related accessions, thereby limiting identification accuracy and efficiency [10–12]. Advances in high-throughput sequencing technologies have enabled widespread application of DNA-based markers for genetic diversity analysis and cultivar identification in crop species [13]. Among available marker types, single-nucleotide polymorphisms (SNPs) are particularly suited for these purposes due to their genome-wide abundance, high polymorphism rate, strong stability, and compatibility with high-throughput genotyping platforms [14,15]. SNP profiling captures genetic variation at single-base resolution, thereby overcoming the subjectivity and limited reproducibility of phenotypic assessment and providing objective insights into population structure and genetic relationships [16]. In melon, SNP-based genotyping platforms have been gradually developed, from early GoldenGate arrays to high-density genome-wide chips, enabling systematic characterization of genetic diversity across different germplasm collections [3,17]. The Melon2K SNP array is a high-quality genome-wide chip developed based on resequencing data from both *melo* and *agrestis* accessions; it enables efficient genotyping and genetic diversity analysis of melon germplasm resources, and demonstrates strong potential for accession identification and breeding [15]. Given the distinct genetic architecture of *agrestis* accessions—characterized by low nucleotide diversity, high LD, and limited inter-subspecific gene flow—subspecies-specific SNP resources and targeted genotyping strategies are particularly critical for accurate identification and classification of thin-skinned melon germplasm [3,6].

Here, we report a systematic characterization of 32 commercially sourced Tianbao melon accessions by integrating phenotypic evaluation of ten agronomic traits with genome-wide SNP analysis from whole-genome resequencing. By combining correlation, phylogenetic, and principal component analyses, we assess the phenotypic and genetic diversity within this commercial germplasm pool, elucidate population structure and inter-accession genetic relationships, and identify cases of genotype–phenotype discordance that cannot be resolved by morphological descriptors alone. We further develop a minimal core SNP panel for high-resolution cultivar authentication and intellectual property protection, establishing a practical molecular "identity card" system for each accession. Together, these findings provide a scientific foundation and practical molecular tool for the breeding improvement, cultivar protection, and sustainable management of the Tianbao melon industry.

2. Materials and Methods

2.1. Plant Materials

A diverse panel of 32 Tianbao melon accessions was utilized in this study. Seeds were obtained from accessions commercially designated as 'Tianbao', provided by multiple companies across different production regions of China.

2.2. Experimental Design and Growth Conditions

The experiment was conducted in a solar greenhouse in Laixi, China (36.89° N, 120.52° E). The greenhouse dimensions were 20 m (length) × 8 m (width) × 5 m (height). The greenhouse was equipped with an integrated water and fertilizer system, ventilation, and shading to maintain stable conditions and was operated under the same management regime. All indoor experiments of this study were conducted at the laboratory in Qingdao Agricultural University.

Sampling was conducted during the summer growing season of 2025. Each accession was represented by 10 plants grown under identical greenhouse conditions. During the harvest period, three plants per accession were randomly selected for fruit collection and quality evaluation. Fruits with proper shape, uniform size and uniform maturity, free of pests, diseases and mechanical damage, were selected. All selected fruits were transferred to the laboratory for further analyses. To minimize sampling bias, plants used for phenotypic evaluation were randomly selected within each accession. This experimental setup was used to characterize phenotypic variation under controlled greenhouse conditions in a single growing season.

2.3. Sample Preparation and Measurement

The agronomic traits of the fruit appearance were investigated and analyzed for comparison. Fruit length (cm) from stem to blossom end and fruit width were determined using a ruler. Fruit shape index = Fruit length/Fruit width. Plant biomass traits, including fruit yield per unit area, yield per plant, and weight per melon, were measured with an electronic balance. For each accession, all the fruits of three plants were randomly selected for quality measurement, including flesh thickness, which was determined by caliper, seed cavity ratio (transverse width of seed cavity/transverse width of mature fruit), and taste and flavor. A central core of tissue (approximately 10 g) was cut from each fruit, placed on a digital refractometer Pal-1 (ATAGO Co., Ltd., Tokyo, Japan), and total soluble solids (%) was determined. Aroma (Ar) and bitterness (Bt) were evaluated by an 8-person sensory panel. Aroma was scored on a three-point scale (0 = no aroma, 1 = weak aroma, 2 = strong aroma), and bitterness was scored on a binary scale (0 = no bitterness, 1 = bitterness present).

2.4. Phenotypic Data Analysis

Phenotypic data were analyzed using Microsoft Excel 2010. Descriptive statistics of quantitative traits, including maximum, minimum, mean, standard deviation, and coefficient of variation, as well as correlation analysis, were calculated using GraphPad Prism 9 (GraphPad Software, San Diego, CA, USA). The genetic diversity index (H') was estimated based on the Shannon–Weaver diversity index [16], using the formula:

$$H' = \sum_{i=1}^{s} P_i ln P_i$$

where i represents the trait category, P_i denotes the proportion of accessions in the i-th category relative to the total number of accessions, and ln represents the natural logarithm. Qualitative traits were classified into discrete categories, and their diversity indices were calculated based on the relative frequency of each category. For quantitative traits, values were standardized prior to hierarchical cluster analysis and graphical visualization, both performed using OriginPro 2021 (OriginLab Corporation, Northampton, MA, USA). The distribution of each phenotypic trait across all accessions was visualized as boxplots using the ggplot2 package (v3.5.0) in R (v4.3.1) [18,19]. Differences in trait values between morphological clusters were assessed using the Wilcoxon rank-sum test in R, with statistical significance set at $p < 0.05$.

2.5. Variant Calling

Whole-genome resequencing reads from 32 melon accessions were aligned to the melon reference genome (Cme13C) [5] using BWA-MEM (version 0.7.17) [20]. The resulting BAM files were coordinate-sorted and indexed using SAMtools [21]. Read groups were assigned and PCR duplicates were marked using GATK4 (version 4.2.0.0) built-in toolsAddOrReplaceReadGroups and MarkDuplicates, respectively [22].

Variant calling was performed per sample using GATK4 HaplotypeCaller in GVCF mode (-ERC GVCF) [22]. The resulting per-sample GVCF files were merged using GATK4 CombineGVCFs, followed by joint genotyping with GATK4 GenotypeGVCFs to generate a population-level variant call set.

Biallelic SNPs were extracted from the variant call set using GATK4 SelectVariants. Hard filtering was applied using GATK4 VariantFiltration, retaining only sites meeting all of the following criteria: QD $\geq$ 2.0, FS $\leq$ 60.0, MQ $\geq$ 40.0, SOR $\leq$ 3.0, MQRankSum ≥ -12.5, and ReadPosRankSum ≥ -8.0. Additional filters were applied using VCFtools [23], retaining SNPs with minor allele frequency (MAF) $\geq$ 0.05, genotype missing rate $\leq$ 0.2. These filters yielded a final set of 173,497 high-quality biallelic SNPs for downstream analyses; SNP counts for each filtering step are provided in Table S2.

2.6. Population Genomic and Fingerprint Analysis

A Neighbor-Joining (NJ) phylogenetic tree was constructed from genome-wide SNP data using MEGA7 [24], with genetic distances calculated under the p-distance model and branch support assessed using 1000 bootstrap replicates. The resulting tree was visualized using FigTree v1.4.4 (http://tree.bio.ed.ac.uk/software/figtree, accessed on 3 February 2026).

Prior to PCA, linkage disequilibrium (LD) pruning was performed using PLINK [25] with a sliding-window approach, retaining SNPs with pairwise $r^2 < 0.2$. PCA was then conducted on the LD-pruned SNP dataset using PLINK [25], and the resulting principal component scores were visualized in R [19].

A minimal core SNP panel was identified using MinimalMarker (v1.04) software with the parameters -kM -m1 -b1 [26]. This tool employs a greedy iterative algorithm that, at each step, selects the SNP maximally resolving the greatest number of previously unresolved

accession pairs, continuing until all accessions are uniquely discriminated. Specifically, -kM designates 'M' as the symbol for missing genotype data, -m1 activates the fast search mode, and -b1 enables bitwise operations to accelerate computation. This analysis was applied to all 32 accessions using the genome-wide SNP dataset, and the resulting fingerprint profiles were visualized using a custom R script.

3. Results

3.1. Phenotypic Diversity in Tianbao Melon

A panel of 32 Tianbao melon accessions, sourced from multiple commercial seed companies across different production regions of China, was evaluated for ten agronomic traits under standardized conditions. These traits comprised fruit length (FL), fruit width (FW), fruit shape index (FSI), single fruit weight (SFW), yield per plant (YPP), flesh thickness (FT), seed cavity ratio (SCR), soluble solids content (SSC), aroma (Ar) and bitterness (Bt) (Figure S1). Coefficients of variation (CV) ranged from 2.37% (SCR) to 11.07% (YPP), with most traits exhibiting CV < 10% (Table 1), indicating limited phenotypic variation and high morphological uniformity across accessions. The Shannon–Weaver diversity index (H′) ranged from 1.701 (SCR) to 2.001 (FSI), reflecting moderately low phenotypic diversity within this commercial germplasm pool. Together, these results indicate that phenotypic descriptors alone provide insufficient resolution to discriminate among closely related Tianbao melon accessions.

Table 1. Coefficient of variation for field traits in melon.

Index	Mean [1]	SD [2]	Max [3]	Min [4]	SEM [5]	CV/% [6]	H′ [7]
FL	9.07	0.42	9.81	8.17	0.07	4.69	1.887
FW	10.42	0.40	11.47	9.83	0.07	3.80	1.825
FSI	0.87	0.04	0.95	0.80	0.01	4.39	2.001
FT	1.91	0.14	2.16	1.68	0.02	7.28	1.852
SCR	0.62	0.01	0.66	0.60	0.00	2.37	1.701
SFW	0.43	0.04	0.54	0.36	0.01	9.71	1.899
YPP	1.62	0.18	2.07	1.21	0.03	11.07	1.869
SSC	11.52	0.81	13.10	10.10	0.14	7.06	1.706
Ar							1.066
Bt							0.594

[1] Arithmetic mean; [2] standard deviation; [3] maximum; [4] minimum; [5] standard error of the mean; [6] coefficient of variation; [7] Shannon–Weaver diversity index.

3.2. Correlation Analysis and Clustering Patterns Based on Phenotypic Traits

Pearson correlation analysis revealed strong interrelationships among traits (Figure 1A). Fruit width showed the strongest positive correlation with single fruit weight ($r = 0.83$, $p < 0.001$), followed by fruit length with single fruit weight ($r = 0.65$, $p < 0.001$) and yield per plant ($r = 0.65$, $p < 0.005$). Pulp thickness was positively correlated with fruit length ($r = 0.48$, $p < 0.001$) and width ($r > 0.50$, $p < 0.05$) but exhibited a strong negative correlation with seed cavity ratio ($r = -0.81$, $p < 0.001$). Negative correlations were observed between fruit size traits (length, width, weight) and soluble solids content ($r = -0.30$ to 0.00); however, none of these associations reached statistical significance (all $p > 0.05$). Given the limited sample size ($n = 32$), the present results cannot confirm a stable inverse relationship between fruit size and sugar accumulation, and further validation with a larger germplasm panel is warranted. Seed cavity ratio was positively correlated with SSC ($r = 0.32$, $p < 0.1$). These results highlight a classic yield-quality trade-off driven by assimilate partitioning, underscoring the challenge of simultaneously improving yield and sugar content based solely on phenotypic selection.

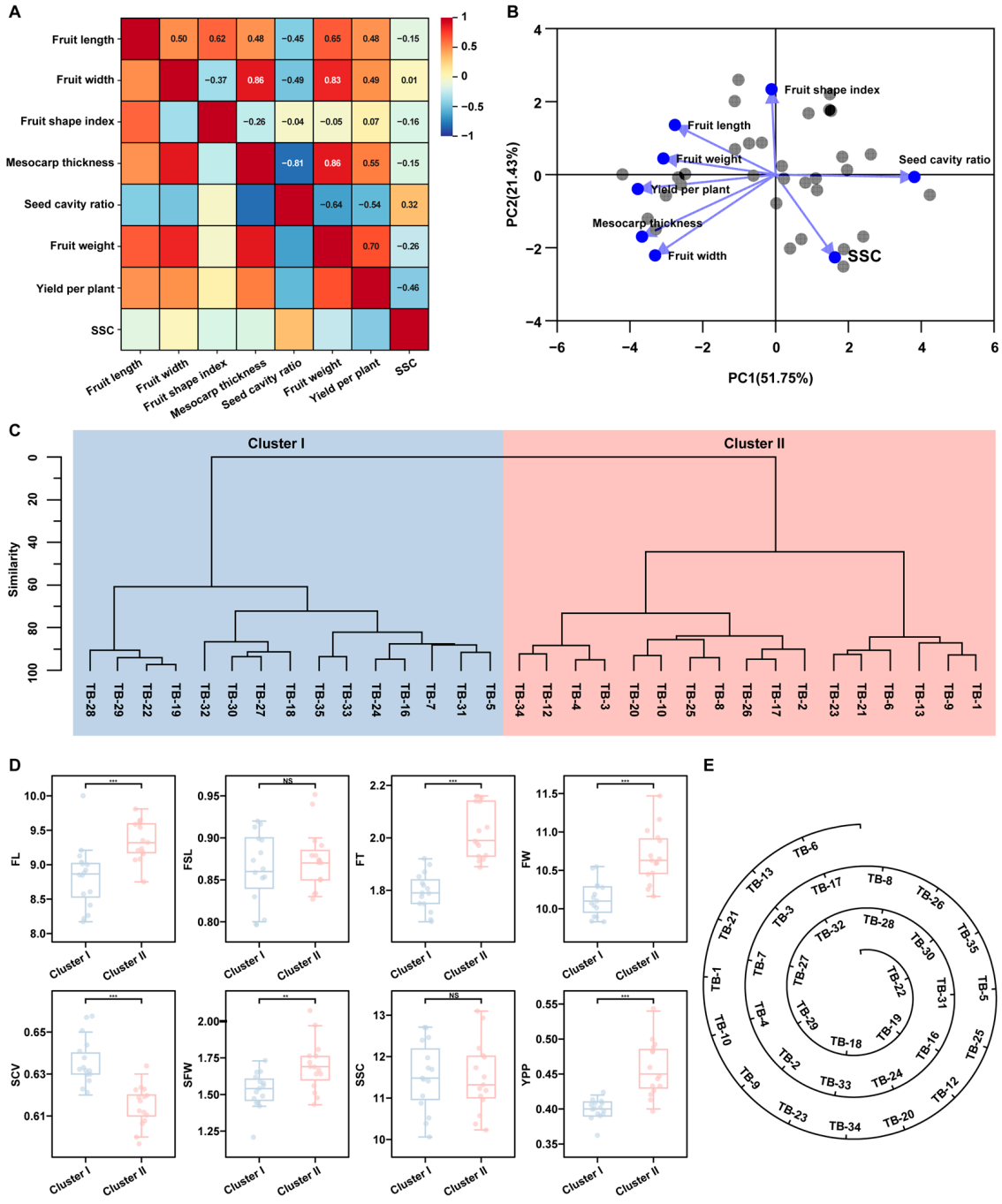

Figure 1. (**A**) Pearson correlation matrix among eight phenotypic traits of 32 melon accessions. (**B**) PCA biplot showing accessions (dots) and trait vectors (arrows). The percentages of variance explained by the first and second principal components (PC1 and PC2) are 51.75% and 21.43%, respectively. (**C**) Hierarchical clustering dendrogram classifying the 32 accessions into two distinct clusters, namely Cluster I (blue) and Cluster II (red), based on integrated physiological traits. (**D**) Boxplot comparisons of eight agronomic traits between Cluster I and Cluster II identified in (**C**). Significant differences between the two clusters were assessed using the Wilcoxon rank-sum test (Wilcoxon W statistic). Significance levels are indicated as ** ($p < 0.01$), and *** ($p < 0.001$); NS (not significant). (**E**) Multi-trait genotype ideotype distance index (MGIDI) visualization of melon phenotypes ranked based on performance.

To resolve the complex multivariate structure of phenotypic variation, principal component analysis (PCA) was performed (Figure 1B). The first three principal components (PCs) with eigenvalues > 0.9 explained 86.83% of the total phenotypic variance (Figure 1B,

Tables S3 and S4). PC1 (51.75% variance) was dominated by yield-related traits (negative loadings for YPP, SFW, FW, FT and positive loadings for SCR), primarily reflecting fruit size and yield potential. PC2 (21.43% variance) was strongly associated with fruit shape index, capturing morphological variation independent of yield. PC3 was primarily associated with SSC, reflecting fruit quality.

Hierarchical clustering via Ward's method classified the 32 accessions into two distinct morphological clusters (Figure 1C). Cluster I (17 accessions) was characterized by superior fruit dimensions and yield, exhibiting significantly higher FL, FW, FT, SFW, and YPP compared with Cluster II ($p < 0.001$) (Table S5). By contrast, no significant differences were observed in FSI ($p = 0.12$) or SSC ($p = 0.48$) between clusters. These results indicate that although fruit size differs, the core morphological identity and sweetness of Tianbao melon remain highly uniform, further underscoring the limitations of morphological traits for accession discrimination (Figure 1D). Cultivar scores derived from principal component values were used to generate a comprehensive performance index (Figure 1E). Accessions with higher scores exhibited greater overall agronomic performance; TB-22, TB-19, TB-18, TB-29, and TB-27 ranked highest, all belonging to Cluster II, consistent with the hierarchical clustering results.

3.3. Genome-Wide SNP Identification and Distribution

To overcome the limitations of phenotypic evaluation, whole-genome resequencing of the 32 accessions using the Cme13C melon reference genome (average mapping depth ~5.39×, ranging from 4.89× to 7.10×) yielded 173,497 high-confidence biallelic SNPs after stringent quality filtering (see Section 2.5). Sliding-window analysis (1 Mb window size) revealed an uneven distribution of SNPs across the chromosomes. High-density SNP hotspots were prominently localized within the 16–25 Mb region of Chromosome 1 and the 15–18 Mb region of Chromosome 5 (Figure 2A). These hypervariable genomic regions represent critical loci underlying genetic differentiation among Tianbao melon germplasm resources.

3.4. Genetic Similarity and Population Structure Based on SNPs

Pairwise genetic similarity coefficients, calculated across all 173,497 SNPs, ranged from 0.85 to 0.96 (Figure 2B), with the majority of values exceeding 0.90, quantitatively confirming the narrow genetic base within the commercial Tianbao melon pool. To further characterize genetic variation within and between the two principal clades, nucleotide diversity (π) and per-accession heterozygosity were calculated for GA and GB (TB-17 and TB-27 were excluded as single-accession groups). Nucleotide diversity was 0.0503 for GA and 0.0567 for GB, with a pairwise FST of 0.017 between the two groups (Table S7). Mean per-accession heterozygosity was 0.1148 for GA and 0.1192 for GB; individual heterozygosity values ranged from 0.074 to 0.221 across all 32 accessions, with TB-27 exhibiting the highest value (0.221; Table S8). To elucidate population structure, NJ phylogenetic analysis (1000 bootstrap replicates) and PCA were performed on the genome-wide SNP dataset (Figure 2C,D; Table S6), partitioning the 32 accessions into four genomic groups: two principal clades (GA and GB) and two singleton lineages represented by individual accessions TB-17 and TB-27, each forming an independent branch. Genomic classification exhibited marked discordance with morphological clustering. Most notably, accessions TB-17 and TB-27, which are phenotypically near-indistinguishable and were grouped together in morphological analyses, displayed relatively low genetic similarity (0.85–0.90) and were resolved as entirely independent lineages in both NJ phylogeny and PCA (Figure 2B–D). This pronounced genotype–phenotype discordance suggests that key Tianbao-type morphological traits may be governed by a small number of loci subject to convergent phenotypic selection

across genetically divergent backgrounds, and that environmental plasticity may further mask underlying genomic differentiation. These findings underscore the inadequacy of morphological descriptors for accurate cultivar authentication and highlight the power of genome-wide SNP data to resolve hidden genetic divergence among phenotypically uniform accessions.

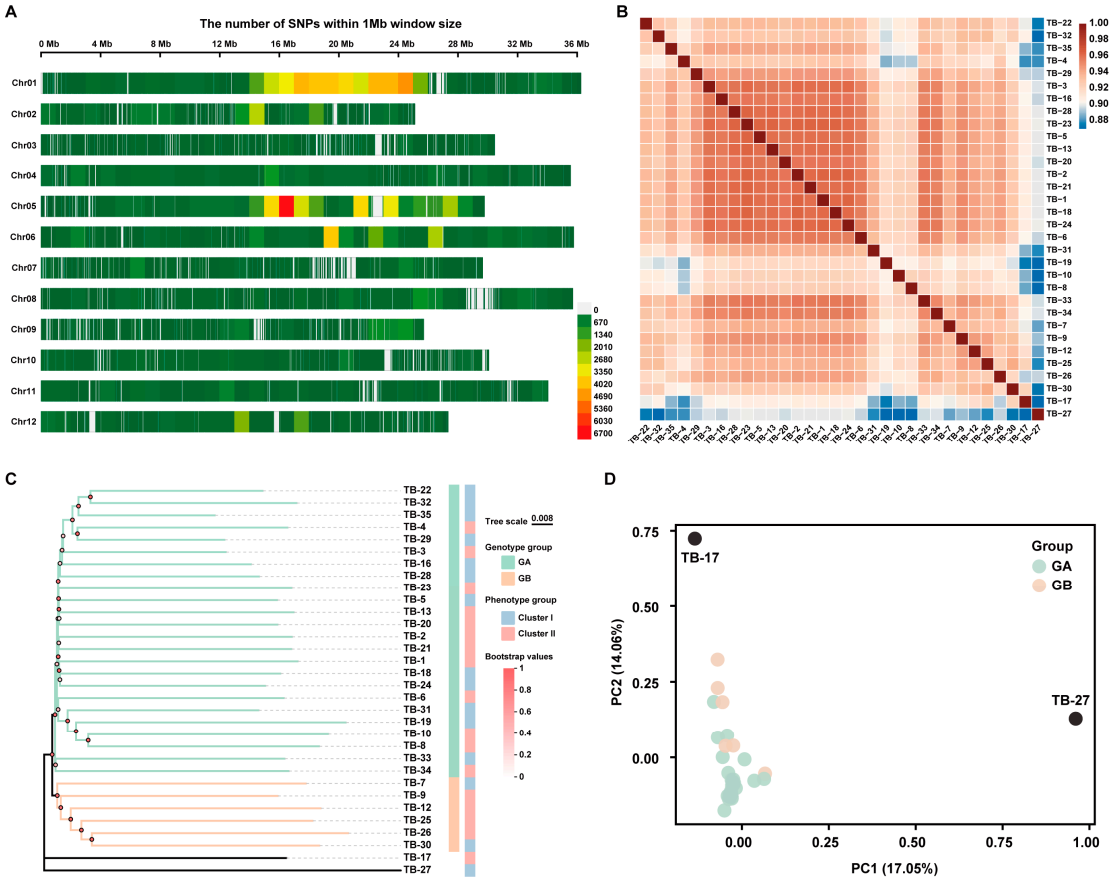

Figure 2. (**A**) Genome-wide distribution pattern of SNPs across the 12 chromosomes. The horizontal axis represents the physical position on the chromosome. (**B**) Pairwise genetic similarity coefficients among 32 accessions were calculated from genome-wide SNP data and visualized as a heatmap. Color scale ranges from blue (low similarity) to red (high similarity). (**C**) SNP-based phylogenetic tree of 32 Tianbao melon accessions. Branch lengths are proportional to genetic distance. Genotypic groupings (Group A, light green; Group B, light orange) and phenotypic cluster assignments (Cluster I, light blue; Cluster II, light red) are indicated by background shading. Notably, TB-17 and TB-27 were resolved as independent singleton lineages. (**D**) Principal component analysis of 32 accessions based on SNPs.

3.5. Development of Core SNP Fingerprints

To facilitate rapid and cost-effective cultivar identification for the melon industry, we aimed to develop a streamlined molecular tool. By applying stringent filtering criteria to the genome-wide variant dataset, a highly informative panel of 20 core SNP loci was identified (Table S9). These loci, widely distributed across the genome, exhibit sufficient polymorphism to capture the genetic uniqueness of each line. Using this 20-SNP panel, a unique genotypic barcode was successfully constructed for all 32 accessions (Figure 3). Pairwise comparisons revealed that most accession pairs differed at 8–15 of the 20 loci, with a mean pairwise difference of 11 SNPs (Figure S3). This robust discriminatory power indicates that the selected core SNPs provide adequate resolution to authenticate all tested

Tianbao-type melons, offering an efficient technical solution for germplasm protection, addressing market adulteration, and supporting intellectual property protection.

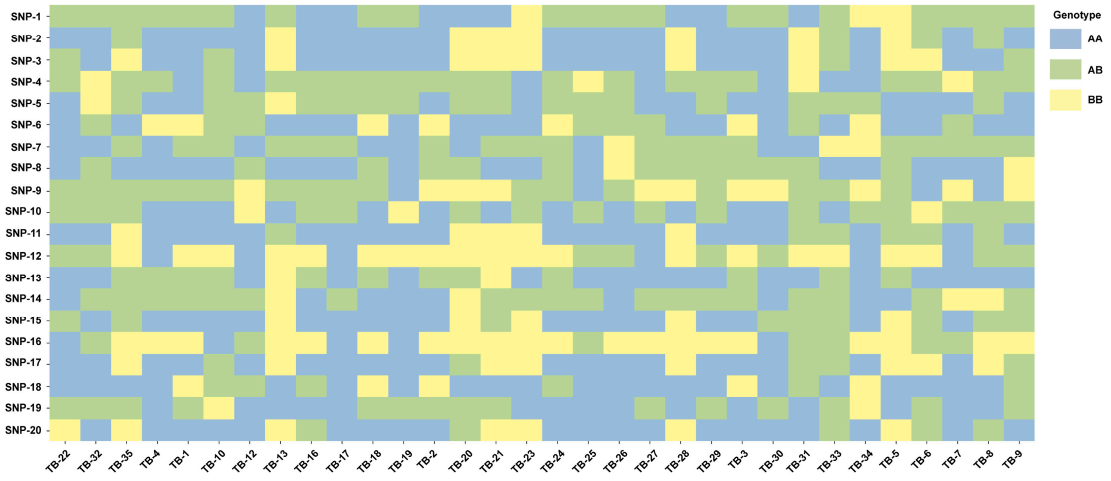

Figure 3. Core SNP-derived fingerprint profile of representative accessions. Different colors correspond to the three genotypes: AA (blue), AB (green), and BB (yellow).

4. Discussion

Systematic phenotypic characterization of 32 commercially sourced Tianbao melon accessions revealed limited overall morphological variation, with coefficients of variation below 10% for most traits and Shannon–Weaver diversity indices in the moderate-to-low range. This low phenotypic diversity is consistent with the documented history of intensive directional selection in commercial Tianbao melon breeding, where prolonged focus on a narrow set of market-favored traits—including uniform fruit shape, thin rind, and high sweetness—has progressively narrowed the genetic base and homogenized varietal characteristics [27–29]. Among measured phenotypic traits, fruit size-related parameters (fruit length, fruit width, single fruit weight) exhibited the highest coefficients of variation and thus represent the principal morphological axes for accession differentiation, consistent with the broader finding that fruit morphology is the primary source of phenotypic diversity in melon populations [30]. By contrast, appearance traits such as fruit skin base color and flesh color showed high stability and consistency across accessions, further indicating that the Tianbao melon showed relatively stable appearance-related traits under the present growing conditions in terms of appearance and morphology under long-term artificial selection. The limited discriminatory power of these traits confirms that it is difficult to effectively distinguish closely related accessions by traditional morphological observation alone, consistent with the findings of Coelho de Amorin et al. [31]. This study was conducted in a single greenhouse during a single growing season; therefore, the present phenotypic results should be interpreted as controlled-environment observations, and multi-season, multi-location trials will be needed to further assess trait stability and genotype × environment interactions.

Multivariate phenotypic analyses—including correlation analysis, principal component analysis, and cluster analysis—consistently revealed both the dominant contribution of fruit morphology to population differentiation and the fundamental limitations of phenotypic approaches for resolving closely related accessions. Correlation analysis revealed significant or extremely significant positive associations among fruit size-related traits, indicating that longitudinal diameter, transverse diameter, and single fruit weight exhibit coordinated variation during fruit development, consistent with the general developmental biology of melon fruit morphology [17]. Principal component analysis further confirmed

that fruit size and shape indices contributed most to population-level differentiation; however, accession distributions in PCA scatter plots were relatively concentrated due to the narrow overall range of variation, precluding the formation of clear, well-resolved independent groups. Phenotypic cluster analysis could classify accessions to a certain extent based on morphological similarity, yet cluster boundaries were ambiguous: some phenotypically similar accessions were not grouped together, while some accessions with obvious phenotypic differences clustered in adjacent branches, resulting in weak stability and consistency of clustering outcomes. The fundamental cause of these inconsistencies is that phenotypic traits are inherently susceptible to environmental conditions, cultivation practices, and growth stage, producing strong genotype-by-environment interaction effects in trait expression; furthermore, phenotypes represent the phenotypic outcome of gene expression and cannot fully reflect true genetic differences at the genomic level. For a germplasm pool with a narrow genetic background and pronounced morphological homogeneity such as the Tianbao melon population, relying solely on phenotypic approaches for germplasm identification, kinship analysis, and breeding parent selection carries a considerable risk of misclassification and is insufficient to meet the practical needs of precise accession identification, lineage traceability, and intellectual property protection [27,32]. As demonstrated by Lamine et al., molecular analysis is independent of environmental effects and can provide additional and precise information for the assessment of genetic diversity that phenotypic approaches cannot offer [33].

It should be noted that the diversity metrics reported here characterize the commercial Tianbao melon pool specifically—a panel of accessions sourced from multiple independent seed companies across different production regions of China—and should not be extrapolated to the broader genetic diversity of *C. melo* subsp. *agrestis*. Within this commercial pool, the narrow range of pairwise genetic similarity values (0.85–0.96) observed across independently sourced accessions from geographically dispersed suppliers reflects a genuine biological property of the Tianbao germplasm, consistent with the documented pattern of narrow genetic backgrounds in intensively selected commercial cultivar groups [34]. To overcome the limitations of phenotypic evaluation, we performed genome-wide SNP analysis, which revealed discernible genotypic variation and a structured population stratification among the 32 accessions despite their high phenotypic uniformity. The low FST (0.017) between GA and GB is consistent with the narrow commercial genetic background of this germplasm pool. Most pairwise genetic similarity values exceeded 0.90, quantitatively confirming the limited genetic base of the commercial Tianbao melon pool, consistent with findings in other intensively selected horticultural cultivar groups [35]. NJ phylogenetic analysis and PCA resolved the accessions into four genomic groups, revealing a population structure that is not apparent from morphological evaluation alone [17]. The pronounced SNP density hotspots identified on Chromosome 1 (16–25 Mb) and Chromosome 5 (15–18 Mb) are of particular interest: these hypervariable genomic intervals likely reflect different selection histories among commercial Tianbao lineages and may harbor loci governing key agronomic traits, representing high-priority targets for fine-mapping and functional characterization in future studies [3,6].

The most striking finding of this study is the profound genotype–phenotype discordance observed for accessions TB-17 and TB-27. Despite being phenotypically near-indistinguishable—and grouped together under morphological clustering—these two accessions displayed low pairwise genetic similarity (0.85–0.90) and were resolved as distinct singleton lineages in both NJ phylogeny and PCA. This discordance most plausibly reflects convergent phenotypic selection: independent breeding lineages may have converged on the same Tianbao-type morphological ideotype through selection on a small number of key loci, while retaining divergent genomic backgrounds at the majority of neutral sites [36].

Environmental plasticity may further contribute by masking genomic differentiation at the phenotypic level, particularly for traits sensitive to environmental conditions [37]. This finding has direct practical implications: accessions that appear identical by morphological inspection may originate from genetically distinct lineages, rendering phenotypic evaluation unreliable for accession rights protection, provenance tracing, and authentication in market dispute resolution [38,39]. These results thus underscore the necessity of integrating molecular marker technologies to achieve stable, reliable, and legally defensible genetic characterization of commercial melon germplasm.

5. Conclusions

This study provides a comprehensive assessment of genetic diversity and cultivar identification in Tianbao melon by integrating phenotypic evaluation with genome-wide SNP analysis. The results reveal that, despite limited morphological variation among accessions, extensive hidden genetic differentiation exists, as evidenced by high-density SNP data that resolve the population into distinct genetic groups. This phenotypic–genotypic discordance underscores the limitations of morphology-based classification and highlights the importance of genome-wide variation in accurately capturing genetic diversity. This study further demonstrates that a set of core SNP markers can effectively distinguish all accessions, enabling high-resolution DNA fingerprinting and accurate cultivar identification. These findings suggest that the genetic architecture of Tianbao melon is driven by underlying genomic variation that is not readily detectable through phenotypic traits. Our work establishes an efficient framework for genetic resource management, varietal purity assessment, and intellectual property protection. It also provides practical tools for precision breeding aimed at improving cultivar integrity and sustainability. The present findings are based on a panel of 32 commercially sourced accessions; future investigations incorporating a broader range of regional germplasm will be valuable for further consolidating the identified population structure and expanding the applicability of the core SNP fingerprinting panel. Additionally, linking the identified genomic variation with functional traits and breeding applications remains an important direction for future work.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/horticulturae12060714/s1. Figure S1: Frequency distribution of traits in 32 Tianbao melon accessions; Figure S2: Phenotypic characteristics of Tianbao melon fruits; Figure S3: Distribution histogram of pairwise sample differences based on SNP fingerprinting. Table S1. Accession list. Table S2. Summary of SNP filtering steps. Table S3. Eigenvalues of the principal components and their contribution rates and cumulative contribution rates. Table S4. Load matrix of principal components. Table S5. Comparison of eight quality traits across two major groups. Table S6. Eigenvalues, contribution rates, and cumulative contribution rates of principal components based on genotypes. Table S7. Summary of population genetic statistics. Table S8. Heterozygosity of 32 Tianbao accessions. Table S9. Information on the core 20 SNPs.

Author Contributions: Conceptualization, K.W. and K.X.; methodology, Y.R. and M.H.; software, M.H., X.S., Q.Z. and S.Z.; validation, Y.R. and W.D.; resources, S.C., X.L. (Xiaoli Liu) and X.L. (Xiaofeng Liu); data curation, H.M. and W.J.; writing—original draft preparation, Y.R., M.H., W.D. and X.S.; writing—review and editing, K.W. and K.X.; supervision, S.C.; funding acquisition, K.W., S.C. and X.L. (Xiaofeng Liu). All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Qingdao Science and Technology for People's Livelihood Demonstration Special Project, grant number: 25-1-5-xdny-25-nsh (founder: S.C.); the second batch of "Tianchi Talents" introduction program of Xinjiang Uygur Autonomous Region in 2023 (founder: K.W.); and the "Taishan Scholar" Foundation of the People's Government of Shandong Province, grant number: tsqn202408233 (founder: Xiaofeng Liu).

Data Availability Statement: Data are contained within the article or Supplementary Materials.

Conflicts of Interest: Author Kexiang Wang was employed by the company Hami Agricultural and Animal Husbandry Industry Development Investment Co. The remaining 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.

References

1. Chen, X.; Li, H.; Dong, Y.; Xu, Y.; Xu, K.; Zhang, Q.; Yao, Z.; Yu, Q.; Zhang, H.; Zhang, Z. A wild melon reference genome provides novel insights into the domestication of a key gene responsible for melon fruit acidity. *Theor. Appl. Genet.* **2024**, *137*, 144. [CrossRef]
2. Tanaka, K.; Shigita, G.; Dung, T.P.; Nhi, P.T.P.; Takahashi, M.; Monden, Y.; Nishida, H.; Ishikawa, R.; Kato, K. Polyphyletic domestication and inter-lineage hybridization magnified genetic diversity of cultivated melon, *Cucumis melo* L. *Breed. Sci.* **2025**, *75*, 168–178. [CrossRef]
3. Zhao, G.; Lian, Q.; Zhang, Z.; Fu, Q.; He, Y.; Ma, S.; Ruggieri, V.; Monforte, A.J.; Wang, P.; Julca, I.; et al. A comprehensive genome variation map of melon identifies multiple domestication events and loci influencing agronomic traits. *Nat. Genet.* **2019**, *51*, 1607–1615. [CrossRef] [PubMed]
4. Nuñez-Palenius, H.G.; Gomez-Lim, M.; Ochoa-Alejo, N.; Grumet, R.; Lester, G.; Cantliffe, D.J. Melon fruits: Genetic diversity, physiology, and biotechnology features. *Crit. Rev. Biotechnol.* **2008**, *28*, 13–55. [CrossRef]
5. Xu, Y.; Liu, B.; Li, Y.; Chen, X.; Yan, C.; Liu, Y.; Wang, H.; Wang, J.; Dong, W.; Deng, S.; et al. A comprehensive omics resource and genetic tools for genetic research and precision breeding of *Cucumis melo* ssp. *agrestis*. *Plant Cell* **2025**, *37*, koaf272. [CrossRef] [PubMed]
6. Liu, S.; Gao, P.; Zhu, Q.; Zhu, Z.; Liu, H.; Wang, X.; Weng, Y.; Gao, M.; Luan, F. Resequencing of 297 melon accessions reveals the genomic history of improvement and loci related to fruit traits in melon. *Plant Biotechnol. J.* **2020**, *18*, 2545–2558. [CrossRef]
7. Tu, K.; Wen, S.; Cheng, Y.; Xu, Y.; Pan, T.; Hou, H.; Gu, R.; Wang, J.; Wang, F.; Sun, Q. A model for genuineness detection in genetically and phenotypically similar maize variety seeds based on hyperspectral imaging and machine learning. *Plant Methods* **2022**, *18*, 81. [CrossRef]
8. Guo, Z.; Yang, Q.; Huang, F.; Zheng, H.; Sang, Z.; Xu, Y.; Zhang, C.; Wu, K.; Tao, J.; Prasanna, B.M.; et al. Development of high-resolution multiple-SNP arrays for genetic analyses and molecular breeding through genotyping by target sequencing and liquid chip. *Plant Commun.* **2021**, *2*, 100230. [CrossRef] [PubMed]
9. Kuang, M.; Wei, S.-J.; Wang, Y.-Q.; Zhou, D.-Y.; Ma, L.; Fang, D.; Yang, W.-H.; Ma, Z.-Y. Development of a core set of SNP markers for the identification of upland cotton cultivars in China. *J. Integr. Agric.* **2016**, *15*, 954–962. [CrossRef]
10. Wang, Y.; Wu, X.; Li, Y.; Feng, Z.; Mu, Z.; Wang, J.; Wu, X.; Wang, B.; Lu, Z.; Li, G. Identification and Validation of a Core Single-Nucleotide Polymorphism Marker Set for Genetic Diversity Assessment, Fingerprinting Identification, and Core Collection Development in Bottle Gourd. *Front. Plant Sci.* **2021**, *12*, 747940. [CrossRef]
11. Fu, Y.-B. Understanding crop genetic diversity under modern plant breeding. *Theor. Appl. Genet.* **2015**, *128*, 2131–2142. [CrossRef]
12. Miao, L.; Wang, X.; Yu, C.; Ye, C.; Yan, Y.; Wang, H. What factors control plant height? *J. Integr. Agric.* **2024**, *23*, 1803–1824. [CrossRef]
13. Rasheed, A.; Hao, Y.; Xia, X.; Khan, A.; Xu, Y.; Varshney, R.K.; He, Z. Crop Breeding Chips and Genotyping Platforms: Progress, Challenges, and Perspectives. *Mol. Plant* **2017**, *10*, 1047–1064. [CrossRef]
14. Wang, L.; Xun, H.; Aktar, S.; Zhang, R.; Wu, L.; Ni, D.; Wei, K.; Wang, L. Development of SNP Markers for Original Analysis and Germplasm Identification in *Camellia sinensis*. *Plants* **2023**, *12*, 162. [CrossRef]
15. Yu, Q.; Li, S.; Su, X.; Chen, X.; Dong, Y.; Yao, Z.; Jiang, N.; Chai, S.; Zhang, Z.; Xu, K. Melon2K array: A versatile 2K liquid SNP chip for melon genetics and breeding. *Hortic. Plant J.* **2025**, *11*, 314–322. [CrossRef]
16. Ousmael, K.; Whetten, R.W.; Xu, J.; Nielsen, U.B.; Lamour, K.; Hansen, O.K. Identification and high-throughput genotyping of single nucleotide polymorphism markers in a non-model conifer (*Abies nordmanniana* (Steven) Spach). *Sci. Rep.* **2023**, *13*, 22488. [CrossRef] [PubMed]
17. Esteras, C.; Formisano, G.; Roig, C.; Díaz, A.; Blanca, J.; Garcia-Mas, J.; Gómez-Guillamón, M.L.; López-Sesé, A.I.; Lázaro, A.; Monforte, A.J.; et al. SNP genotyping in melons: Genetic variation, population structure, and linkage disequilibrium. *Theor. Appl. Genet.* **2013**, *126*, 1285–1303. [CrossRef] [PubMed]
18. Wickham, H. ggplot2. In *Wiley Interdisciplinary Reviews: Computational Statistics*; ACM: New York, NY, USA, 2011; Volume 3, pp. 180–185.
19. Wickham, H.; Bryan, J.; Chang, W.; McGowan, L.; François, R.; Grolemund, G.; Lemon, J.; Kassambara, A.; Mundt, F.; Revelle, W. R: A language and environment for statistical computing. *J. Open Source Softw.* **2020**, *8*, 1–20.

20. Li, H.; Durbin, R. Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics* **2009**, *25*, 1754–1760. [CrossRef] [PubMed]
21. Danecek, P.; Bonfield, J.K.; Liddle, J.; Marshall, J.; Li, H. Twelve years of SAMtools and BCFtools. *GigaScience* **2021**, *10*, giab008. [CrossRef]
22. McKenna, A.; Hanna, M.; Banks, E.; Sivachenko, A.; Cibulskis, K.; Kernytsky, A.; Garimella, K.; Altshuler, D.; Gabriel, S.; Daly, M. The Genome Analysis Toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* **2010**, *20*, 1297. [CrossRef]
23. Danecek, P.; Auton, A.; Abecasis, G.; Albers, C.A.; Banks, E.; DePristo, M.A.; Handsaker, R.E.; Lunter, G.; Marth, G.T.; Sherry, S.T.; et al. The variant call format and VCFtools. *Bioinformatics* **2011**, *27*, 2156–2158. [CrossRef]
24. Kumar, S.; Stecher, G.; Tamura, K. MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. *Mol. Biol. Evol.* **2016**, *33*, 1870–1874. [CrossRef] [PubMed]
25. Purcell, S.; Neale, B.; Todd-Brown, K.; Thomas, L.; Ferreira, M.A.R.; Bender, D.; Maller, J.; Sklar, P.; de Bakker, P.I.W.; Daly, M.J.; et al. PLINK: A Tool Set for Whole-Genome Association and Population-Based Linkage Analyses. *Am. J. Hum. Genet.* **2007**, *81*, 559–575. [CrossRef]
26. Fujii, H.; Ogata, T.; Shimada, T.; Endo, T.; Iketani, H.; Shimizu, T.; Yamamoto, T.; Omura, M. Minimal marker: An algorithm and computer program for the identification of minimal sets of discriminating DNA markers for efficient variety identification. *J. Bioinform. Comput. Biol.* **2013**, *11*, 1250022. [CrossRef]
27. Luan, F.; Delannay, I.; Staub, J.E. Chinese melon (*Cucumis melo* L.) diversity analyses provide strategies for germplasm curation, genetic improvement, and evidentiary support of domestication patterns. *Euphytica* **2008**, *164*, 445–461. [CrossRef]
28. Mohammadi, S.A.; Prasanna, B.M. Analysis of Genetic Diversity in Crop Plants—Salient Statistical Tools and Considerations. *Crop Sci.* **2003**, *43*, 1235–1248. [CrossRef]
29. Ren, D.; Liao, J.; Zhang, K.; Zhang, J.; Qu, J.; Ma, G.; Li, J. Genetic Analysis and Fingerprint Construction for Thick-Skinned Melon (*Cucumis melo* subsp. *melo*) Based on InDel Markers. *Plants* **2025**, *14*, 3782. [CrossRef]
30. Monforte, A.J.; Diaz, A.; Caño-Delgado, A.; van der Knaap, E. The genetic basis of fruit morphology in horticultural crops: Lessons from tomato and melon. *J. Exp. Bot.* **2014**, *65*, 4625–4637. [CrossRef]
31. Clisneide, C.; Queiroz, M.; Iana, P.; Ronaldo, S.; de Souza Santos, S.; Barbosa, G. Morphological diversity and identification of accessions of melon. *Afr. J. Agric. Res.* **2016**, *11*, 3622–3632. [CrossRef]
32. Jolliffe, I.T.; Cadima, J. Principal component analysis: A review and recent developments. *Philos. Trans. A Math. Phys. Eng. Sci.* **2016**, *374*, 20150202. [CrossRef] [PubMed]
33. Lamine, M.; Mliki, A. Elucidating genetic diversity among sour orange rootstocks: A comparative study of the efficiency of RAPD and SSR markers. *Appl. Biochem. Biotechnol.* **2015**, *175*, 2996–3013. [CrossRef]
34. Nguyen, N.N.; Kwon, Y.-S.; Park, J.-R.; Sim, S.-C. Development of a Core Set of SSR Markers for Cultivar Identification and Seed Purity Tests in Oriental Melon (*Cucumis melo* L. var. *makuwa*). *Hortic. Sci. Technol.* **2019**, *37*, 119–129. [CrossRef]
35. Zhang, J.; Yang, J.; Lv, Y.; Zhang, X.; Xia, C.; Zhao, H.; Wen, C. Genetic diversity analysis and variety identification using SSR and SNP markers in melon. *BMC Plant Biol.* **2023**, *23*, 39. [CrossRef] [PubMed]
36. Chen, W.; Chen, L.; Zhang, X.; Yang, N.; Guo, J.; Wang, M.; Ji, S.; Zhao, X.; Yin, P.; Cai, L. Convergent selection of a WD40 protein that enhances grain yield in maize and rice. *Science* **2022**, *375*, eabg7985. [CrossRef]
37. Villellas, J.; Ehrlén, J.; Crone, E.E.; Csergő, A.M.; Garcia, M.B.; Laine, A.L.; Roach, D.A.; Salguero-Gómez, R.; Wardle, G.M.; Childs, D.Z. Phenotypic plasticity masks range-wide genetic differentiation for vegetative but not reproductive traits in a short-lived plant. *Ecol. Lett.* **2021**, *24*, 2378–2393. [CrossRef] [PubMed]
38. Zhang, J.; Ren, J.; Yang, J.; Fu, S.; Zhang, X.; Xia, C.; Zhao, H.; Yang, K.; Wen, C. Evaluation of SNP fingerprinting for variety identification of tomato by DUS testing. *Agric. Commun.* **2023**, *1*, 100006. [CrossRef]
39. Li, X.; Hu, Z.; Yu, W.; Xie, H.; Wang, X.; Huang, P.; Zhang, X.; Yang, J.; Li, Y.; Zhao, W. Advances and challenges in plant molecular marker technologies and their applications in the artificial intelligence empowered era. *Front. Plant Sci.* **2025**, *16*, 1757949. [CrossRef]