Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Genomic variation analysis and discovery of stable heterozygous loci in Loquat (Eriobotrya japonica)

Yuyu Tang
Southwest University
Shangjian Yang
Southwest University
Guolu Liang
Southwest University
Qiao He
Southwest University
Danlong Jing
Southwest University
Qigao Guo
Southwest University
Jiangbo Dang
dangjiangbo@126.com

Southwest University

Research Article

Keywords: Loquat, Genome, Stable heterozygous loci, InDel marker, Self-incompatibility

Posted Date: December 8th, 2025

DOI: https://doi.org/10.21203/rs.3.rs-8035482/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License. Read Full License

Additional Declarations: No competing interests reported.

Version of Record: A version of this preprint was published at BMC Plant Biology on February 16th, 2026. See the published version at https://doi.org/10.1186/s12870-026-08346-8.

Abstract

Background

The high heterozygosity of the loquat (Eriobotrya japonica Lindl.) genome has hindered the efficient application of modern breeding strategies. Although several genome assemblies have been published, comprehensive analyses of gene haplotypes, locus-specific variation, and the inheritance of heterozygous loci remain limited.

Result

To investigate genomic variation in loquat, 28 accessions were subjected to whole-genome resequencing, and the resulting reads were aligned to the reference genome for variant identification. Single nucleotide polymorphisms (SNPs), insertions and deletions (InDels), copy number variations (CNVs), and structural variants (SVs) were detected. A total of 196 heterozygous InDel loci and 182 heterozygous SNP loci were identified. Based on 48 InDel loci, 48 primer pairs were designed, and two pairs consistently amplified heterozygous genotypes across diverse loquat materials. These loci were stably heterozygous in wild, semi-wild, and cultivated accessions, as well as in progeny derived from self-pollination or hybridization, regardless of ploidy level or geographic origin. Finally, gene expression at these two loci differed significantly among various organs, including roots, stems, flowers, fruits, and leaves.

Conclusion

The consistent heterozygosity of these loci, which are located on different chromosomes, suggests a deviation from Mendelian segregation and implies the involvement of a gamete-selective or pre-fertilization recognition mechanism. These findings enhanced the understanding of the genetic mechanisms underlying the loquat genome, provided useful molecular tools for breeding, and offered new insights into loquat reproductive biology and improvement.

Introduction

Loquat (*Eriobotrya japonica* Lindl.) is an important subtropical fruit tree in China, cultivated for its leaves and fruits, which have considerable nutritional and pharmacological value. Fruit trees are characterized by high levels of heterozygosity, which helps maintain genetic diversity among fruit tree genotypes and facilitates efficient utilization of heterosis, thereby reducing trait decline caused by homozygosity. China maintains a rich collection of loquat germplasm, whose diversity sustains high genomic heterozygosity and enhances adaptability to variable environments[1]. To date, multiple versions of the loquat genome have been published (‘Qixing’, ‘Jiefangzhong’, ‘Guiye-23’, ‘Dawuxing’)[2–4]. However, all loquat genomes published to date have been collapsed consensus assemblies combining mixed haplotypes. To date, no haplotype-resolved genome of loquat has been reported, which has limited investigations into allelic variation underlying important traits at the genome-wide level. Sun et al. and Yin et al. constructed haplotype-resolved genomes for pear and honey tangerine, respectively, revealing associations between structural variation, gene co-selection, and hormonal changes affecting gene expression and fruit agronomic traits[5, 6]. Furthermore, although previous studies have revealed abundant SNP and InDel variations in the loquat genome[3, 7, 8], systematic analyses of locus-specific

variations and the inheritance patterns of heterozygous loci have not been conducted. This limitation may pose challenges to research on the loquat genome and molecular-assisted breeding.

Self-incompatibility (SI), outcrossing, and homozygote lethality are key mechanisms that generate heterozygosity, while widespread vegetative propagation contributes to the preservation of heterozygosity in fruit trees[9]. Several Rosaceae fruit trees display gametophytic self-incompatibility (GSI), a common trait in this family, including apple, plum, apricot, and sweet cherry[10]. Multiple S-RNase alleles have been identified in loquats[11–13]. However, commercially cultivated loquat varieties are frequently grown without dedicated pollinator trees. This may be attributed to S gene mutations in loquats, such as the S6-RNase mutation and the promoter deletion of S10-RNase, both of which have been linked to self-compatibility[14, 15]. Cross-pollination occurs frequently among loquats, and hybridization readily takes place between different cultivars as well as between loquats and other *Eriobotrya* species[16]. No reports of homozygous lethality have been documented in loquats.

Therefore, the genomes of 28 loquat accessions were resequenced, and SNPs, InDels, CNVs, and SVs were analyzed to elucidate genomic variation patterns in loquat. InDel heterozygous loci were identified, and codominant markers were developed. These markers were then screened across diploid and polyploid levels in natural, hybrid, and inbred populations to identify heterozygous loci with stable inheritance. Genes harboring these loci were analyzed to characterize their expression patterns. This study offers a valuable reference for research on the loquat genome and hybrid breeding programs.

Materials and methods

Plant material

Table 1 lists the 28 loquat accessions selected for whole-genome resequencing by their classifications (wild type, semi-wild type, and cultivated type). Subsequently, 10 diploid hybrid populations, 16 polyploid varieties, and 2 tetraploid × diploid progeny populations were selected for analysis. These hybrid progeny were then analyzed to investigate genetic patterns (Tables 2 and 3).In addition, different organs of the diploid cultivar ‘Huabai No. 1’, including roots, stems, flowers, fruits, and leaves, were sampled at various developmental stages. Flowers at the bud stage, young fruits, roots, stems, and leaves collected in November were used for partial gene expression analysis. All plant materials were obtained from the Loquat Germplasm Resource Garden at Southwest University and the university’s 105 Teaching and Research Base. Our research team had previously developed the hybrid populations.

Loquat leaves were collected and stored at -80°C. Genomic DNA (gDNA) was extracted using a modified CTAB method[17]. Total RNA was extracted using the RNA Polysaccharide & Polyphenol Kit (Tiangen Biotech Co., Ltd., Beijing, China). DNA quality was assessed using 1.0% agarose gel electrophoresis, and high-quality gDNA was diluted to 50 ng/μL for subsequent PCR and qPCR analyses.

Table 1
Materials for whole-genome resequencing

Group	Lines (Varieties)	Abbreviation	Quantity
wild type	Guiye 8	GZ-8	1
	Guiye 9	GZ-9	1
	Guiye 16	GZ-16	1
	Guiye 18	GZ-18	1
	Guiye 36	GZ-36	1
	Guiye 56	GZ-56	1
	Guiye 58	GZ-58	1
	Guiye 61	GZ-61	1
	Guiye 4–31	GZ-4-31	1
	Lunainai	LNN	1
	Fengduzaoshu	FDZS	1
	Guiye Lanya	GYLY	1
	Hongshicunyesheng	HSC	1
Semi-wild type	Zhuofu	ZF	1
	Shimianyuanguo	SMYG	1
	Guiye 7	GZ-7	1
Cultivated type	‘Baiyu’	BY	1
	‘Guifei’	GF	1
	‘Dawuxing’	DWX	1
	‘Bingtangzhong’	BTZ	1
	‘Huabai No.1’	HB1	1
	‘Zaozhong No.6’	ZZ6	1
	‘Ruantiaobaisha’	RTBS	1
	H30-6	H30-6	1
	‘Dayangwanbai’ (Japan)	DYWB	1
	‘Senweizaosheng’ (Japan)	SWZS	1
	‘Marc’ (Spain)	MARC	1
	‘Pelluches’ (Spain)	PELL	1

Table 2
Diploid hybrid population

Group	Lines (Varieties)	Abbreviation	Quantity
Hybrid offspring population	‘Senweizaosheng’×‘Dayangwanbai’	SWZS×DYWB	12
	‘Senweizaosheng’×‘Jinhua No.1’	SWZS×JH1	12
	‘Huabai No.1’×‘Jinhua No.1’	HB1×JH1	13
	‘Huabai No.1’×‘Pelluches’	HB1×PELL	13
Reciprocal cross offspring populations	‘Pelluches’× ‘Huabai No.1’	PELL×HB1	50
	‘Huabai No.1’×‘Pelluches’	HB1×PELL	50
	‘Ruantiaobaisha’ × ‘Guiye16’	RTBS×GZ-16	50
	‘Guiye16’× ‘Ruantiaobaisha’	GZ-16×RTBS	50
Self-crossing population	‘Huabai No.1’ Self - crossed offspring	HB1-SC	38
	‘Huabai No.2’ Self - crossed offspring	HB2-SC	38

Table 3
Ploidy of polyploid materials and hybrid populations

Group	Lines (Varieties)	Abbreviation	Quantity
triploid	A313	A313	1
	A322	A322	1
	D25	D25	1
	D27	D27	1
	‘Seedless Guoyu’	LGY	1
	‘Seedless Fuzao’	LFZ	1
	‘Seedless Zaoyu’	LZY	1
	‘Huayu Seedless No.1’	LHY1	1
tetraploid	A479	A479	1
	B431	B431	1
	B432	B432	1
	B433	B433	1
	B456	B456	1
	H424	H424	1
	K376-1	K376-1	1
	K474	K474	1
tetraploid×Diploid offspring	B431 × ‘Pelluches’ progeny	B432×PELL	37
	B431 × ‘Guifei’ progeny	B432×GF	37

Whole genome library construction and sequencing

Genomic DNA extracted from each sample was used to construct one sequencing library with an average insert size of 400 bp. Libraries were sequenced on the Illumina NovaSeq 6000 platform (Illumina, San Diego, CA, USA) using a paired-end mode (2 × 150 bp) according to the manufacturer’s standard protocols. The quality of raw sequencing reads was initially assessed using FastQC (v0.11.9) to evaluate base composition, GC content, and sequence quality metrics. Adapter sequences and low-quality reads were filtered out using fastp (v0.20.0). The “MarkDuplicates” tool from the Picard software package was used to remove duplicates. The IndelRealigner command in the GATK suite was employed to realign all reads near insertions/deletions (InDels), enhancing the accuracy of SNP prediction.

Variant calling, SNPs, InDels, CNVs, and SVs

High-quality reads were aligned to the reference genome using BWA-MEM (v0.7.12-r1039) with default parameters[18]. The resulting SAM files were converted to sorted BAM files using Picard (v1.107), and duplicates were removed with MarkDuplicates. To minimize mapping errors near InDels, realignment was performed using RealignerTargetCreator and IndelRealigner in GATK (v3.8). Mapping statistics.

SNPs and InDels were called using UnifiedGenotyper in GATK with stand_call_conf = 30 and stand_emit_conf = 10, followed by filtering with stringent criteria to ensure accuracy. Population-level variants were jointly called for trait-associated samples and annotated using ANNOVAR[19]. CNVs were detected using CNVnator (v0.2.7) with bin_size = 100, retaining regions with read depth ≤ 0.05 or ≥ 1.8. Structural variants, including deletions, insertions, inversions, and intra- or inter-chromosomal translocations, were identified using BreakDancer (v1.3.7).

Screening of heterozygous loci

The whole-genome resequencing data of 28 loquat accessions (Table 1) were analyzed to identify InDel and SNP loci that were heterozygous in all accessions. Heterozygous loci were required to meet the following criteria: (1) the sequencing depth at the locus was greater than two-thirds of the average genome-wide depth for the corresponding accession; (2) a chi-square test showed no significant deviation from a 1:1 allelic depth ratio ($p > 0.05$). A total of 196 heterozygous InDel loci were identified. Due to the difficulty of designing PCR-detectable primers based on SNPs, this study focused on developing PCR primers from InDel loci for validation and further analysis of selected heterozygous loci.

Primers were designed using SnapGene v6.0.2. For each heterozygous locus, 200 bp upstream and 200 bp downstream sequences were extracted. Forward primers were designed within the 1–200 bp upstream region, while reverse primers were designed within the 210–400 bp downstream region. Primer lengths ranged from 19 to 22 bp, with annealing temperatures of 52–63°C. A total of 48 primer pairs were designed based on 48 heterozygous InDel loci.

The designed primers were then screened according to the following criteria: (1) the InDel locus remained heterozygous (producing two bands); (2) the amplified fragment size ranged from 80 to 400 bp. Primers were synthesized by Qingke Biotechnology Co., Ltd. (Chongqing, China) and initially tested using genomic DNA from 11 diploid parental accessions, which corresponded to the cultivated materials listed in Table 1 (excluding H30-6). Primers that produced clear amplification and two distinct bands were considered suitable for further experiments targeting heterozygous InDel loci.

Primer testing and screening of heterozygous loci

PCR amplification was performed in a 20 μL reaction system containing 1 μL genomic DNA (50 ng/μL), 0.5 μL forward primer (10 μM), 0.5 μL reverse primer (10 μM), 0.3 μL 10.5 dNTP Mix (2.5 mmol/L), 2 μL 10× buffer (containing 15 mmol/L Mg^{2+}), 0.2 μL rTaq DNA polymerase (5 U/μL), and 15.5 μL ddH_2O. Amplification was carried out using a qTOWER³ real-time PCR system under the following conditions: initial denaturation at 95°C for 10 min, followed by 35 cycles of 95°C for 30 s, annealing at 50–62°C for 30 s, and extension at 72°C for 30 s, with a final extension at 72°C for 10 min and a final hold at 16°C.

The PCR products were separated on 8% non-denaturing polyacrylamide gels, and gel images were documented. From the 48 primer pairs tested, six primer pairs produced clear amplification with two distinct bands, corresponding to heterozygous loci, and were selected for subsequent experiments.

Quantitative analysis of heterozygous loci in polyploid loquat

First, conventional PCR was used to genotype two heterozygous loci, previously identified as stably heterozygous in diploid loquat, in triploid and tetraploid accessions.

Second, following the method reported by Wang Haiyan[20], qPCR was used to analyze the number of different alleles. Primer design: Based on the insertion-deletion fragments obtained by comparing the heterozygous locus sequences with the genome, common primers and long-fragment-specific primers were designed for alleles A and a, respectively. The primers were validated using the diploid loquat materials ‘Huabai No.1’ and ‘Jinhua No.1’ as standard samples, and the validated primers were used for subsequent polyploid genotyping. The qPCR system was consistent with that described in the following section, and the primers were designed using Primer 5 software.

Calculation method: The relative copy number of allele A is b1, the relative copy number of alleles A and a is b3, and the relative copy number of allele a is b2. b2 = b3 - b1. The genotype of the material can be determined using the A/(A + a) value, where b = b1/(b1 + b2). The genotype of triploid and tetraploid loquat materials is calculated based on the proportion of the b value, allowing for the genotyping of different stable heterozygous loci in these materials.

RT-qPCR (real-time fluorescence quantitative PCR) analysis

Gene expression was analyzed in the roots, stems, flowers, fruits, and leaves of the diploid variety ‘Huabai 1’. Total RNA was extracted and cDNA synthesized according to the instructions of the Tiangen RNA extraction kit. Quantitative PCR (qPCR) was performed using a qTOWER³ real-time PCR system (Analytik Jena, Germany) with NovoStart® SYBR Green SuperMix (Nearshore ProteinTech Co., Ltd.). The allele-specific primers used for qPCR are listed in Supplementary Table S2 and were designed using Primer5 software.

The 20 μL qPCR reaction system contained 2 μL cDNA, 2 μL primers, 6 μL ddH_2O, and 10 μL 2× SYBR Green I. The amplification program consisted of an initial denaturation at 94°C for 5 min, followed by 40 cycles of 94°C for 15 s and 50–60°C for 20 s, with a subsequent melting curve analysis from 60 to 95°C. Relative gene expression levels were calculated using the $2^{\wedge-\Delta\Delta Ct}$ method, with EjActin (GenBank accession: JN004223) as the internal reference. Each sample was analyzed in three biological replicates.

Results

Genome-wide variant discovery and distribution

The sequence alignment results and sequencing depths for 28 loquat materials are shown in Table S3, with an average alignment rate of 99.08% and average sequencing depths ranging from 14.89× to 26.44×. After filtering Illumina reads and performing alignment, we called SNPs and InDels using GATK Unified Genotyper.

After filtering, a total of 33,887,591 SNPs and 2,938,761 InDels were detected across the genome. Most variants were distributed in intergenic (71.38% of SNPs and 65.35% of InDels) and intronic regions (9.87% and 13.33%, respectively), while only a small proportion occurred in exonic regions (4.86% of SNPs and 1.92% of InDels). Among exonic variants, 2.88% of SNPs were nonsynonymous, and 0.73% of InDels caused frameshift deletions. These findings suggest that the majority of genomic variations are located in noncoding regions, with a small subset potentially affecting gene function.

The numbers of InDel and SNP loci in each loquat sample are summarized in Tables S6 and S7, respectively. InDels were predominantly 1–24 bp in length, with single-base insertions or deletions being most common. SNPs were mainly single-base substitutions. We calculated the proportions of heterozygous SNP and InDel loci relative to the reference genome size in different loquat accessions (Fig. 1a,b). The types and numbers of CNV and SV variants were quantified (Fig. 1b,d). Cultivated accessions showed the highest numbers of deletions and inter-chromosomal translocations, whereas wild accessions had more duplications, deletions, inversions, and intra-chromosomal translocations. Semi-wild accessions displayed intermediate variation, generally resembling the cultivated type. Overall, more CNVs were observed in cultivated varieties, suggesting genomic simplification during domestication, whereas wild varieties retained greater structural diversity.

Identification of universally heterozygous loci

Among the 28 resequenced accessions, 74 InDel loci and 128 SNP loci were fully heterozygous, representing 0.087% of all InDels and 0.0116% of all SNPs, respectively (Table S7 and S6). In total, 74 InDel loci were consistently heterozygous across all samples. Based on the genome resequencing and InDel identification approach described in the Materials and Methods, 48 loci were selected for primer development. The primer sequences are listed in Table S1, and their chromosomal distribution is shown in Fig. 2a.

PCR screening was conducted using the genomes of 11 resequenced loquat accessions as templates and primers designed for 48 InDel loci. All 48 primer pairs produced amplification bands, although some showed excessive non-specific amplification. After additional screening and primer redesign, two pairs that generated clear double bands were obtained (Fig. 2b). These loci (Table 4) were located at LG10:9288618 and LG11:7010931 on different chromosomes.

These two stable heterozygous loci were further validated using genomic DNA from the remaining diploid materials (16 wild and 1 cultivated accession). Because the Guiye 56 and wild accessions in Hongshi Village had already died before the later sampling, only 15 samples were used for validation with the primers (LG10:9288618). As shown in Fig. 2c, all diploid materials exhibited double bands, confirming them as heterozygous loci. Overall, two InDel loci located on independent chromosomes were identified as consistently heterozygous across diverse loquat accessions, providing strong evidence for loci with stable inheritance independent of genetic background or geographical origin.

Table 4
Table of Stable Heterozygous Site Primers

Gene name	Chromosome	Location	Forward primer sequence (5'→3')	Reverse primer sequence (5'→3')
EVM0016038	LG10	9288618	AAATGAAACTCTTAGTCGGTC	CAGATGCACTACTTACTGG
EVM0009023	LG11	7010931	ACTATTCGATGCACAATATCCG	CAAAGGCTCACCTCATGGA

Genetic patterns of stable heterozygous loci in diploid organisms and analysis of related gene expression

The results showed that in the hybrid progeny (Fig. 3a, b), reciprocal cross populations (Fig. 3c, d), and self-pollinated progeny (Fig. 3e, f), the amplification products of these 2 primer pairs in the genomic DNA of the materials all yielded two bands. Therefore, these two loci (Table 4) are heterozygous in diploid hybrid progeny, reciprocal cross populations, and self-pollinated progeny. These two heterozygous loci exhibit stable heterozygous inheritance across different populations and are not influenced by the mode of inheritance (reciprocal cross, self-pollination, etc.).

Analysis of the stability and inheritance of heterozygous loci at the polyploid level

The two heterozygous loci identified as stably inherited in diploid populations (Table 4) were further investigated to explore their inheritance patterns in polyploid loquat. Triploid, tetraploid, and tetraploid × diploid progenies were selected for this analysis (Table 3).

The results showed that both primer pairs amplified two clear bands in the genomic DNA of triploid and tetraploid accessions (Fig. 4a, b) as well as in self-cross progenies (Fig. 4c,d), indicating heterozygosity. These findings suggest that the two loci remain stably heterozygous in polyploid loquats, with no alteration caused by changes in ploidy level (triploid vs. tetraploid). Furthermore, the banding patterns of polyploid hybrids were consistent with those of diploid hybrids, confirming the stable inheritance of these loci in polyploid progenies.

Design and screening of primers for a gene typing system based on qPCR

Using the genomic DNA of the diploid cultivars ‘Jinhua No.1’ and ‘Huabai No.1’ as templates, primers were validated. Primers showing high specificity, a single peak in the melting curve, and the ability to distinguish heterozygous loci were selected for subsequent genotyping of polyploid materials. The qPCR genotyping system could only be applied to loci with relatively large InDel fragment differences. Only one locus yielded suitable primers for amplification, and the resulting sequence is shown in Table S10. This stable heterozygous locus (LG10:9288618) consistently exhibited the Aa genotype across all tested individuals, where “A” and “a” correspond to the short and long fragments, respectively. As shown in Table 5, the locus-specific primers designed for this site can be effectively utilized for subsequent qPCR-based genotyping of triploid and tetraploid loquat materials. Detailed result analysis is in Table S11.

Table 5
Genotyping of diploid loquat accessions using the locus-specific primer LG10:9288618.

Variety (lines)	Ploidy	Parameter b	Genotype
Huabai No.1	2x	0.48 ± 0.015	Aa
Jinhua No.1	2x	0.5 ± 0.014	Aa

Statistical analysis of genotyping results at heterozygous loci in polyploidy

The qPCR genotyping of the stable heterozygous locus (LG10:9288618) showed that, in both triploid and tetraploid loquats, the qPCR genotyping (Table 7) fully agreed with the marker-based gel electrophoresis results. No homozygous genotypes were detected, and all samples maintained heterozygosity, although allelic dosage ratios varied among genotypes. These results further confirm that this locus consistently exhibits stable heterozygosity across diploid and polyploid loquat populations. Detailed result analysis is in Table S12. In addition, results of triploid genotyping showed that the heterozygous locus LG010: 9288618 was not linked tightly. Because if the heterozygous loci were linked tightly, there was no AAa genotype; all genotypes were AAAaaa.

Table 7
Heterozygous 3(LG010: 9288618)Genotyping Statistics

Variety (lines)	Ploidy	Parameter b	Genotype
A313	3x	0.46 ± 0.012	AAa
A322	3x	0.46 ± 0.032	AAa
D25	3x	0.48 ± 0.025	AAa
D27	3x	0.44 ± 0.077	AAa
‘Seedless Guoyu’	3x	0.46 ± 0.049	AAa
‘Seedless Zaoyu’	3x	0.45 ± 0.010	AAa
‘Huayu Seedless No.1’	3x	0.52 ± 0.100	AAa
‘Seedless Fuzhao’	3x	0.44 ± 0.009	AAa
A479	4x	0.47 ± 0.043	AAaa
B431	4x	0.47 ± 0.061	AAaa
B432	4x	0.49 ± 0.047	AAaa
B433	4x	0.48 ± 0.02	AAaa
B456	4x	0.46 ± 0.018	AAaa
H424	4x	0.41 ± 0.077	AAaa
K376-1	4x	0.43 ± 0.078	AAaa
K474	4x	0.45 ± 0.032	AAaa

Organ-specific expression of locus-harboring genes

To further explore the functions of these heterozygous loci, we preliminarily measured the expression levels of the genes at these two loci in different organs (roots, stems, flowers, fruits, and leaves) of the diploid loquat ‘Huabai No.1’ (Fig. 5).

Site (LG010:9288618) is located in the gene EVM0016038, which exhibits the highest expression levels in leaves, significantly higher than in other tissues ($p < 0.05$). To explore possible biological roles, we quantified expression levels of the two locus-harboring genes in roots, stems, leaves, flowers, and fruits of diploid loquat 'Huabai No.1' (Fig. 5). EVM0016038 displayed pronounced leaf-enriched expression, whereas EVM0009023 was root-biased ($p < 0.05$). Sequence homology analysis revealed that EVM0016038 corresponds to Arabidopsis AT4G28980 (CDKF1), a cyclin-dependent kinase-activating kinase, while EVM0009023 corresponds to AT2G04630 (RPB6B), a subunit of RNA polymerase II. The expression levels of genes at different loci vary across different organs, and their functions and expression patterns require further investigation.

Discussion

In this study, two InDel loci exhibiting consistent heterozygosity were identified and validated across wild, semi-wild, and cultivated loquat (Eriobotrya japonica) accessions of varying ploidy levels and geographic origins. The persistence of heterozygous genotypes in all materials, including self-pollinated and hybrid progeny, reveals a striking deviation from Mendelian segregation, as previously reported in other perennial out-crossing species (e.g. citrus)[21]. Such stability implies that specific genetic or reproductive mechanisms maintain allelic balance, providing a new perspective on the inheritance dynamics of perennial fruit trees.

Investigation into the mechanisms underlying stable heterozygous loci in loquat

Departures from Mendelian segregation can originate either before fertilization or after zygote formation. In flowering plants, pre-zygotic regulation includes chemotropic guidance and the controlled reception of the pollen tube by the embryo sac. Synergid cells secrete cysteine-rich attractants (LURE peptides) that guide conspecific pollen tubes to the micropyle, establishing species-specific recognition before fertilization[22]. Subsequently, the receptor-like kinase FERONIA mediates pectin- and nitric-oxide-dependent pollen-tube reception, terminating growth of the incoming tube and preventing multiple pollen entries[23]. These pre-fertilization mechanisms, well-characterized in model plants such as Arabidopsis, act immediately before gamete fusion and therefore provide a mechanistic precedent for a compatibility bias that could, in principle, favor heteroallelic combinations and maintain stable heterozygosity. Unilateral cross-incompatibility systems, such as the Zea mays Ga2 locus can alter gamete transmission before fertilization. The Ga2 gene forms a unilateral pre-zygotic barrier in maize, where pollen carrying compatible alleles successfully penetrates silks while incompatible pollen is arrested, thereby distorting segregation ratios at the population level[24]. Similarly, a comparable pre-fertilization compatibility bias could underlie the stable heterozygosity observed in loquat. In Rosaceae, the S-RNase-based GSI system prevents self-pollen from achieving fertilization through pistil-expressed S-RNases that selectively degrade self-pollen RNA. At the same time, compatible pollen overcomes the restriction imposed by non-self S-RNases through S-locus F-box proteins (SLFs)[25, 26]. Acting entirely before fertilization, this molecular recognition system enforces allelic diversity and promotes outcrossing. Analogously, although direct molecular evidence in loquat is currently lacking, if similar allele-dependent interactions occur in loquat, a comparable pre-zygotic bias at guidance or reception could favor heteroallelic combinations and sustain universal heterozygosity.

Alternatively, post-fertilization explanations for segregation distortion in plants include meiotic drive systems such as Ab10 in maize and structural-variant-linked transmission bias, as well as balanced-lethal architectures documented in *Oenothera* that maintain permanent heterozygosity by suppressing recombination and enforcing large linked chromosomal blocks[27–30]. Another potential post-fertilization mechanism is homozygous lethality, which can enrich heterozygous site frequencies by preventing the survival of individuals carrying homozygous alleles at specific loci. Although extensive research has examined single-allele homozygous lethality[31, 32], our study observed only heterozygous loci: no homozygous band patterns for either haplotype were detected, with exclusively heterozygous patterns present. This indicates that the stable heterozygosity observed in loquat does not conform to conventional homozygous lethality.

Although such mechanisms are capable of generating segregation distortion, they are unlikely to account for the loquat pattern. Moreover, no evidence of zygotic or endosperm lethality was detected among accessions, further excluding a post-fertilization cause. Sustained heterozygosity across multiple ploidy levels and genetically diverse accessions would necessitate persistent genotype-specific lethality or the existence of tightly linked balanced-lethal complexes, neither of which is supported by our data and both of which would be expected to produce distinctive linkage and structural signatures[27, 28].

Taken together, mechanisms acting before fertilization provide the most coherent framework for our data. The stable recovery of heterozygotes without detectable homozygotes across diploid, triploid, and tetraploid backgrounds is most consistent with a pre-zygotic compatibility bias operating at pollen-tube guidance or reception.

This interpretation is conceptually aligned with LURE-mediated attraction and FERONIA-dependent reception, which establish recognition steps immediately before fertilization, and it is distinct from the canonical S-locus GSI pathway in Rosaceae[22, 23, 26]. The core regulatory locus of GSI remains the single S site. In loquat, this locus, including the S gene, has been mapped to linkage group (LG) 13[12, 33]. Unlike the S-RNase-linked loci on LG13, our heterozygous loci are located on LG10 and LG11, indicating that they are physically unlinked to the S locus and likely follow a distinct genetic mechanism.

Future experiments resolving the stage of bias by time-resolved pollen-tube imaging, allele-marked competitive pollination, and staged genotyping of gametes, zygotes, and embryos will be decisive. In addition to hypothesizing potential recognition-based mechanisms, we examined the expression of genes located at these stable heterozygous loci to explore their possible biological roles.

Expression analysis of genes at stable heterozygous loci

The leaf-enriched gene EVM0016038 shares homology with Arabidopsis AT4G28980, encoding the cyclin-dependent kinase-activating kinase CDKF1 (CAK1AT). CDKF1 functions as a master regulator that phosphorylates and stabilizes CDKD proteins, ensuring proper transcriptional and mitotic progression during cell proliferation and gametogenesis[34, 35]. Given its pivotal role in coordinating CDK–CAK signaling, maintaining heterozygosity at this locus may buffer key checkpoints linking cell-cycle control and transcriptional activation.

The root-enriched gene EVM0009023 is homologous to Arabidopsis AT2G04630, which encodes the RPB6B subunit of RNA polymerase II. This protein was identified as an interactor of the receptor-like kinase MRLK in

Arabidopsis, and both proteins are coexpressed in shoot and root apical meristems, where phosphorylation-dependent nuclear translocation of target transcription factors occurs[36]. This result is consistent with our finding that this gene exhibits high expression levels in roots. Together, their expression patterns suggest that CAK–RLK–Pol II signaling modules may integrate gametic compatibility signals with transcriptional control, potentially contributing to the persistence of heterozygosity in loquat.

Our findings support a model in which stable heterozygosity at two InDel loci is maintained through a pre-fertilization recognition mechanism distinct from the canonical S-RNase-mediated self-incompatibility system. In this model, heteroallelic combinations are preferentially fertilized, contributing to the persistence of heterozygosity and potentially enhancing heterosis and reproductive compatibility in loquat. Although the mechanism maintaining stable heterozygosity remains to be fully elucidated, its potential value lies in maintaining heterosis, improving reproductive compatibility, and guiding hybrid breeding strategies in loquat and other perennial fruit trees.

Future studies should determine the developmental stage at which segregation distortion arises through time-resolved genotyping and competitive pollination assays, and verify the universality of these loci across ploidy levels using large-scale population and linkage analyses. These efforts will clarify whether stable heterozygosity in loquat is maintained by pollen competition or other pre-zygotic recognition mechanisms.

Conclusion

This study identified and validated two heterozygous InDel markers located on different chromosomes, which consistently exhibited stable heterozygosity across loquat accessions of different geographic origins (China, Japan, and Spain) and ploidy levels (diploid, triploid, and tetraploid). The consistent deviation from Mendelian segregation suggests that specific genetic or reproductive factors may contribute to maintaining heterozygosity in loquat, possibly involving pre-fertilization recognition or gamete selection processes. Although the underlying mechanism remains to be clarified, these loci provide a useful basis for further genetic and molecular investigations and may offer references for studies on loquat reproduction and breeding.

Abbreviations

InDel
insertion/deletion
SNP
single nucleotide polymorphism
CNV
copy number variant
SV
structural variant
GSI
gametophytic self-incompatibility
SI
self-incompatibility
SSI

sporophytic self-incompatibility
PAGE
polyacrylamide gel electrophoresis

Declarations

Authors' contributions

JBD,QGG,YYT and SJY designed the experiments. YYT and SJY selected the material. YYT and SJY performed the experiments. YYT, SJY, and JBD analyzed the data. YYT and SJY wrote the paper. JBD, QGG, and GLL commented on the manuscript and revised the text and structure. QH,DLJ,QGG and GLL provided project management and acquired the project funding. All authors read and approved the final manuscript.

Acknowledgements

Not applicable

Funding

This work was financially supported by the National Key R&D Program of China (No.2023YFD1600800), the National Natural Science Foundation of China (No. 32470386), National Natural Science Foundation of China (No. 32171820), the Chongqing Science and Technology Commission (cstc2024ycjh-bgzxm0202 and cstc2021jscx-gksbX0010) and the Characteristic Fruit Industry and Technology System Innovation Team of Chongqing Agriculture and Rural Affairs Commission [No. 2020 (3) 01 and No. CQMAITS20230505].

Data availability

The sequencing datasets generated and analyzed in this study have been deposited in the National Genomics Data Center (NGDC). They are available under BioProject accession PRJCA052068 and Genome Sequence Archive (GSA) accession CRA034324. The data can be accessed via the following link: https://ngdc.cncb.ac.cn/gsa/s/D6xiC7qr.

Supplementary Information

Supplementary Table S1-S12

Supplementary Figure S1-S17. Full-length, uncropped electrophoresis gel images corresponding to all cropped gels shown in the main figures.

Ethics approval and consent to participate

Not applicable.

Consent for publication

All authors agree to publish the final manuscript.

Conflict of interest

The authors declare that they have no conflict of interest.

References

1. Wan N-F, Fu L, Dainese M, Hu Y-Q, Pødenphant Kiær L, Isbell F, et al. Plant genetic diversity affects multiple trophic levels and trophic interactions. Nat Commun. 2022;13:7312. https://doi.org/10.1038/s41467-022-35087-7.
2. Jiang S, An H, Xu F, Zhang X. Chromosome-level genome assembly and annotation of the loquat (*Eriobotrya japonica*) genome. GigaScience. 2020;9:giaa015. https://doi.org/10.1093/gigascience/giaa015.
3. Jing D, Liu X, He Q, Dang J, Hu R, Xia Y, et al. Genome assembly of wild loquat (*Eriobotrya japonica*) and resequencing provide new insights into the genomic evolution and fruit domestication in loquat. Horticulture Research. 2023;10:uhac265. https://doi.org/10.1093/hr/uhac265.
4. Wang Y. A draft genome, resequencing, and metabolomes reveal the genetic background and molecular basis of the nutritional and medicinal properties of loquat (Eriobotrya japonica (Thunb.) Lindl). Hortic Res. 2021;8:231. https://doi.org/10.1038/s41438-021-00657-1.
5. Sun M, Cao B, Li K, Li J, Liu J, Xue C, et al. Haplotype-resolved, gap-free genome assemblies provide insights into the divergence between Asian and European pears. Nat Genet. 2025;57:2040–51. https://doi.org/10.1038/s41588-025-02273-4.
6. Yin M, Song X, He C, Li X, Li M, Li J, et al. The haplotype-resolved genome assembly of an ancient citrus variety provides insights into the domestication history and fruit trait formation of loose-skin mandarins. Genome Biol. 2025;26:61. https://doi.org/10.1186/s13059-025-03535-4.
7. Peng Z, Zhao C, Li S, Guo Y, Xu H, Hu G, et al. Integration of genomics, transcriptomics and metabolomics identifies candidate loci underlying fruit weight in loquat. Horticulture Research. 2022;9:uhac037. https://doi.org/10.1093/hr/uhac037.
8. Li J. Current Status and Advances in Loquat Genomics: From Genome Mapping to Molecular Breeding. tgmb. 2024. https://doi.org/10.5376/tgmb.2024.14.0028.
9. Zhang Q, Chen W, Sun L, Zhao F, Huang B, Yang W, et al. The genome of Prunus mume. Nat Commun. 2012;3:1318. https://doi.org/10.1038/ncomms2290.
10. Wu J, Gu C, Khan MA, Wu J, Gao Y, Wang C, et al. Molecular Determinants and Mechanisms of Gametophytic Self-Incompatibility in Fruit Trees of Rosaceae. Critical Reviews in Plant Sciences. 2013;32:53–68. https://doi.org/10.1080/07352689.2012.715986.
11. Nyska R, Raz A, Baras Z, Shafir S, Goldway M, Schneider D. Self-compatibility in loquat (Eriobotrya japonica Lindl.) is possibly due to S6-RNase mutation. Scientia Horticulturae. 2013;161:43–8. https://doi.org/10.1016/j.scienta.2013.05.027.
12. Carrera L, Sanzol J, Herrero M, Hormaza JI. Genomic characterization of self-incompatibility ribonucleases (S-RNases) in loquat (Eriobotrya japonica Lindl.) (Rosaceae, Pyrinae). Mol Breeding. 2009;23:539–51. https://doi.org/10.1007/s11032-008-9254-7.
13. Yang Q, Fu Y, Wang Y, Liu L, Li X, Peng S. Identification of 21 novel S-RNase alleles and determination of S-genotypes in 66 loquat (Eriobotrya) accessions. Mol Breeding. 2018;38:61. https://doi.org/10.1007/s11032-018-0800-7.

14. Niska R, Goldway M, Schneider D. S6-RNase Is a Marker for Self-compatibility in Loquat (Eriobotrya japonica Lindl.). horts. 2010;45:1146–9. https://doi.org/10.21273/HORTSCI.45.8.1146.
15. Wu C, Zhang Y, Wang H, Yan M, Wang S. Self-compatibility of 'Zaohuang' loquat is attributed to the lack of a 52 bp fragment in the S-RNase promoter. Scientia Horticulturae. 2022;301:111093. https://doi.org/10.1016/j.scienta.2022.111093.
16. Li G, Zhang Z, Yang X, Qiao Y, He X, Gao Y, et al. Inter-specific and Inter-generic Hybridization Compatibility of Eriobotrya Species (Loquat) and Related Genera. Horticultural Plant Journal. 2016;2:315–22. https://doi.org/10.1016/j.hpj.2017.02.001.
17. Wen G, Dang J, Xie Z, Wang J, Jiang P, Guo Q, et al. Molecular karyotypes of loquat (Eriobotrya japonica) aneuploids can be detected by using SSR markers combined with quantitative PCR irrespective of heterozygosity. Plant Methods. 2020;16:22. https://doi.org/10.1186/s13007-020-00568-7.
18. Li H, Durbin R. Fast and accurate short read alignment with Burrows–Wheeler transform. Bioinformatics. 2009;25:1754–60. https://doi.org/10.1093/bioinformatics/btp324.
19. Wang K, Li M, Hakonarson H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. Nucleic Acids Research. 2010;38:e164–e164. https://doi.org/10.1093/nar/gkq603.
20. Wang H, Dang J, Wu D, Xie Z, Yan S, Luo J, et al. Genotyping of polyploid plants using quantitative PCR: application in the breeding of white-fleshed triploid loquats (Eriobotrya japonica). Plant Methods. 2021;17:93. https://doi.org/10.1186/s13007-021-00792-9.
21. Ollitrault P, Ahmed D, Costantino G, Evrard J-C, Cardi C, Mournet P, et al. Segregation Distortion for Male Parents in High Density Genetic Maps from Reciprocal Crosses between Two Self-Incompatible Cultivars Confirms a Gametophytic System for Self-Incompatibility in Citrus. Agriculture. 2021;11:379. https://doi.org/10.3390/agriculture11050379.
22. Okuda S, Tsutsui H, Shiina K, Sprunck S, Takeuchi H, Yui R, et al. Defensin-like polypeptide LUREs are pollen tube attractants secreted from synergid cells. Nature. 2009;458:357–61. https://doi.org/10.1038/nature07882.
23. Duan Q, Liu M-CJ, Kita D, Jordan SS, Yeh F-LJ, Yvon R, et al. FERONIA controls pectin- and nitric oxide-mediated male–female interaction. Nature. 2020;579:561–6. https://doi.org/10.1038/s41586-020-2106-2.
24. Kermicle JL, Evans MMS. The Zea mays Sexual Compatibility Gene ga2: Naturally Occurring Alleles, Their Distribution, and Role in Reproductive Isolation. Journal of Heredity. 2010;101:737–49. https://doi.org/10.1093/jhered/esq090.
25. Takayama S, Isogai A. SELF-INCOMPATIBILITY IN PLANTS. Annu Rev Plant Biol. 2005;56:467–89. https://doi.org/10.1146/annurev.arplant.56.032604.144249.
26. Sassa H. Molecular mechanism of the S-RNase-based gametophytic self-incompatibility in fruit trees of Rosaceae. Breed Sci. 2016;66:116–21. https://doi.org/10.1270/jsbbs.66.116.
27. Golczyk H, Massouh A, Greiner S. Translocations of Chromosome End-Segments and Facultative Heterochromatin Promote Meiotic Ring Formation in Evening Primroses. The Plant Cell. 2014;26:1280–93. https://doi.org/10.1105/tpc.114.122655.
28. Steiner E. NEW ASPECTS OF THE BALANCED LETHAL MECHANISM IN OENOTHERA. Genetics. 1956;41:486–500. https://doi.org/10.1093/genetics/41.4.486.

29. Kanizay LB, Pyhäjärvi T, Lowry EG, Hufford MB, Peterson DG, Ross-Ibarra J, et al. Diversity and abundance of the abnormal chromosome 10 meiotic drive complex in Zea mays. Heredity. 2013;110:570–7. https://doi.org/10.1038/hdy.2013.2.
30. Dawe RK. The maize abnormal chromosome 10 meiotic drive haplotype: a review. Chromosome Res. 2022;30:205–16. https://doi.org/10.1007/s10577-022-09693-6.
31. White JK, Gerdin A-K, Karp NA, Ryder E, Buljan M, Bussell JN, et al. Genome-wide Generation and Systematic Phenotyping of Knockout Mice Reveals New Roles for Many Genes. Cell. 2013;154:452–64. https://doi.org/10.1016/j.cell.2013.06.022.
32. The International Mouse Phenotyping Consortium, Dickinson ME, Flenniken AM, Ji X, Teboul L, Wong MD, et al. High-throughput discovery of novel developmental phenotypes. Nature. 2016;537:508–14. https://doi.org/10.1038/nature19356.
33. Lv S, Qiao X, Zhang W, Li Q, Wang P, Zhang S, et al. The Origin and Evolution of RNase T2 Family and Gametophytic Self-incompatibility System in Plants. Genome Biology and Evolution. 2022;14:evac093. https://doi.org/10.1093/gbe/evac093.
34. Takatsuka H, Ohno R, Umeda M. The Arabidopsis cyclin-dependent kinase-activating kinase CDKF;1 is a major regulator of cell proliferation and cell expansion but is dispensable for CDKA activation. The Plant Journal. 2009;59:475–87. https://doi.org/10.1111/j.1365-313X.2009.03884.x.
35. Takatsuka H, Umeda-Hara C, Umeda M. Cyclin-dependent kinase-activating kinases CDKD;1 and CDKD;3 are essential for preserving mitotic activity in *Arabidopsis thaliana*. The Plant Journal. 2015;82:1004–17. https://doi.org/10.1111/tpj.12872.
36. Fujita H, Takemura M, Tani E, Nemoto K, Yokota A, Kohchi T. An Arabidopsis MADS-Box Protein, AGL24, is Specifically Bound to and Phosphorylated by Meristematic Receptor-Like Kinase (MRLK). Plant and Cell Physiology. 2003;44:735–42. https://doi.org/10.1093/pcp/pcg092.

Figures

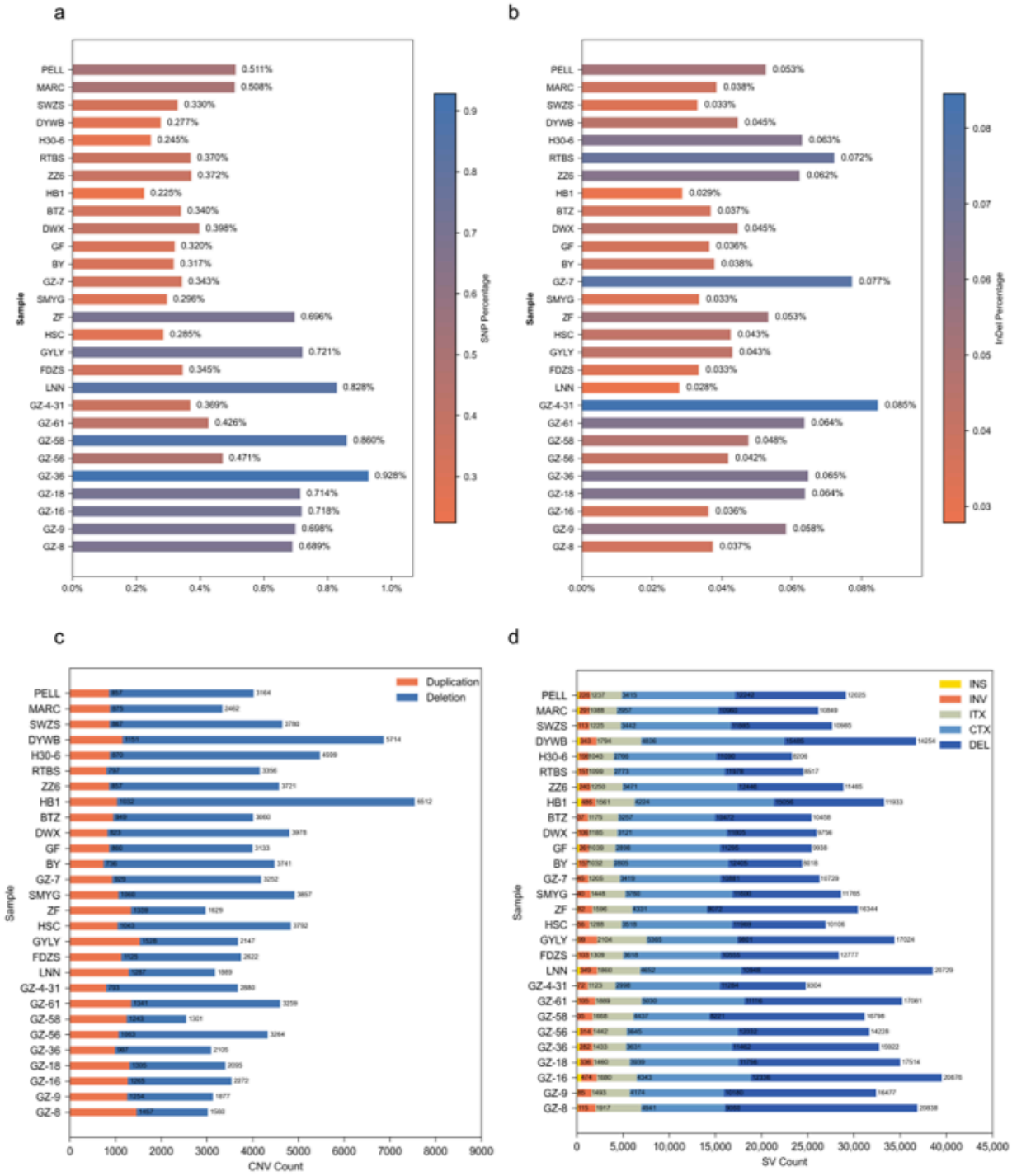


Figure 1

Loquat accessions used in this study, categorized by type: wild-type (GZ-8, GZ-9, GZ-16, GZ-18, GZ-36, GZ-56, GZ-58, GZ-61, GZ-4-31, LNN, FDZS, GYLY, HSC), semi-wild-type (ZF, SMYG, GZ-7), and cultivated-type (BY, GF, DWX, BTZ, HB1, ZZ6, RTBS, H30-6, DYWB, SWZS, MARC, PELL). (a) Proportion of heterozygous SNP loci relative to genome size in different loquat accessions. (b) Proportion of heterozygous InDel loci relative to genome size in different loquat accessions. (c)The stacked bar chart displays the counts of duplication (orange) and deletion

(blue) copy number variations (CNVs) detected across 28 loquat accessions. Each bar represents one accession, with the two stacked colors indicating the proportion of duplication and deletion events in the total CNV count per accession. The numbers on top of each bar denote the exact counts. The X-axis shows the total number of CNV events, and the Y-axis lists the accession names. (d) The stacked bar chart shows the counts of five types of structural variants (SV) detected in 28 loquat accessions, including insertions (INS, yellow), inversions (INV, orange), intra-chromosomal translocations (ITX, gray), inter-chromosomal translocations (CTX, light blue), and deletions (DEL, dark blue). Each bar represents one accession, with stacked segments indicating the total counts of different SV types. The numbers at the top of the bars denote the exact counts. The X-axis represents the number of SV events, and the Y-axis lists the accession names.

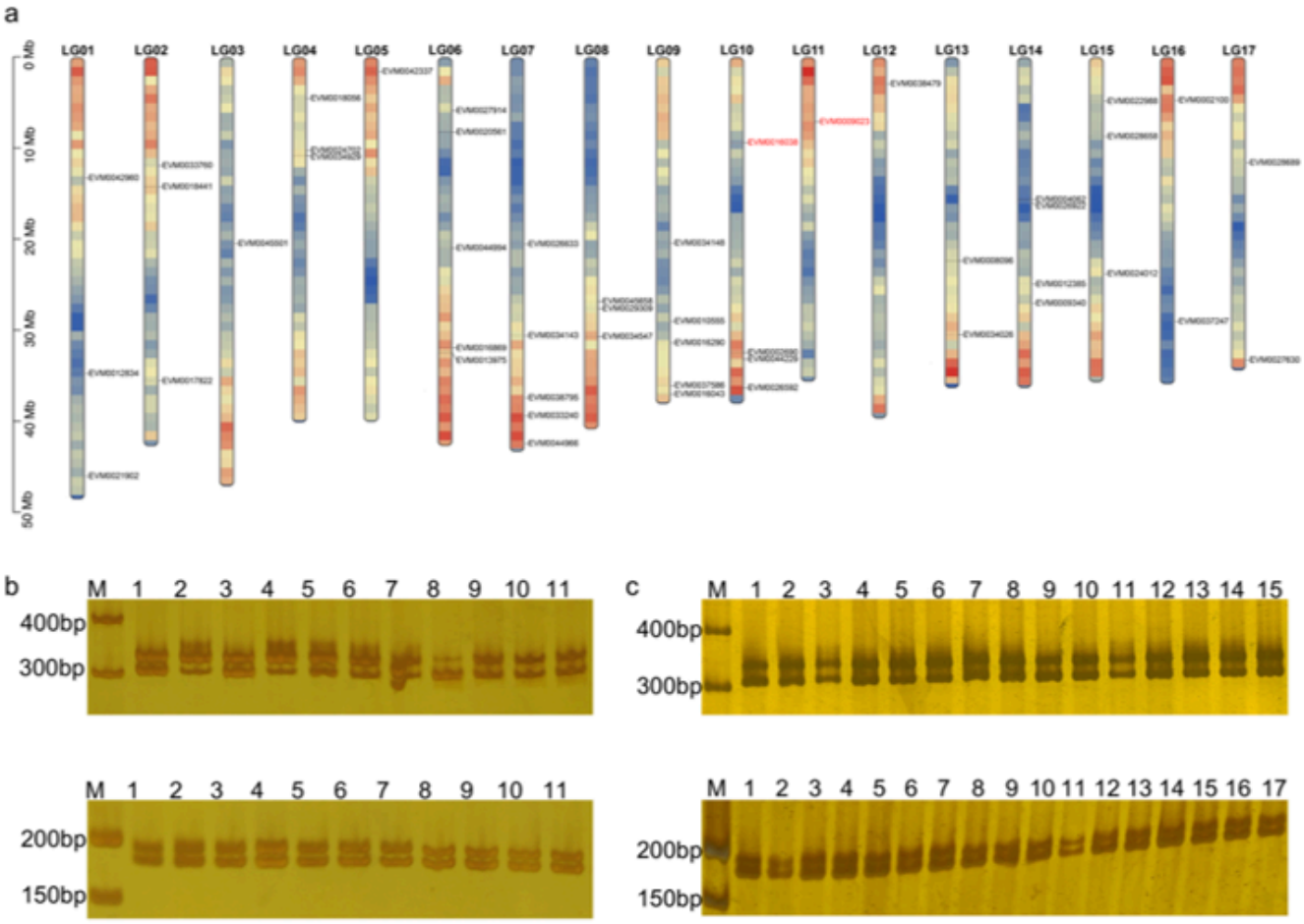


Figure 2

Distribution, screening, and validation of heterozygous loci in loquat. (a) Distribution of 48 heterozygous marker loci across 17 chromosomes. Red indicates the selected heterozygous markers. (b) PAGE analysis of primer screening using 11 loquat accessions (1–11: ZZ6, DYWB, BY, GF, RTBS, PELL, BTZ, MARC, HB1, SWZS, DWX). (c) PAGE validation of selected InDel primers using 15 and 17 loquat accessions, respectively (1–15: FDZS, GZ-7, GZ-8, GZ-9, GZ-16, GZ-18, GZ-36, GZ-4-31, GZ-58, GZ-61, GYLY, LNN, SMYG, ZF, H30-6; 1–17: FDZS, HSC, GZ-56, GZ-7, GZ-8, GZ-9, GZ-16, GZ-18, GZ-36, GZ-4-31, GZ-58, GZ-61, GYLY, LNN, SMYG, ZF, H30-6). Figures (b) and (c) correspond to loci LG10:9288618 and LG11:7010931, from top to bottom. The marker used is DL 500 DNA marker.

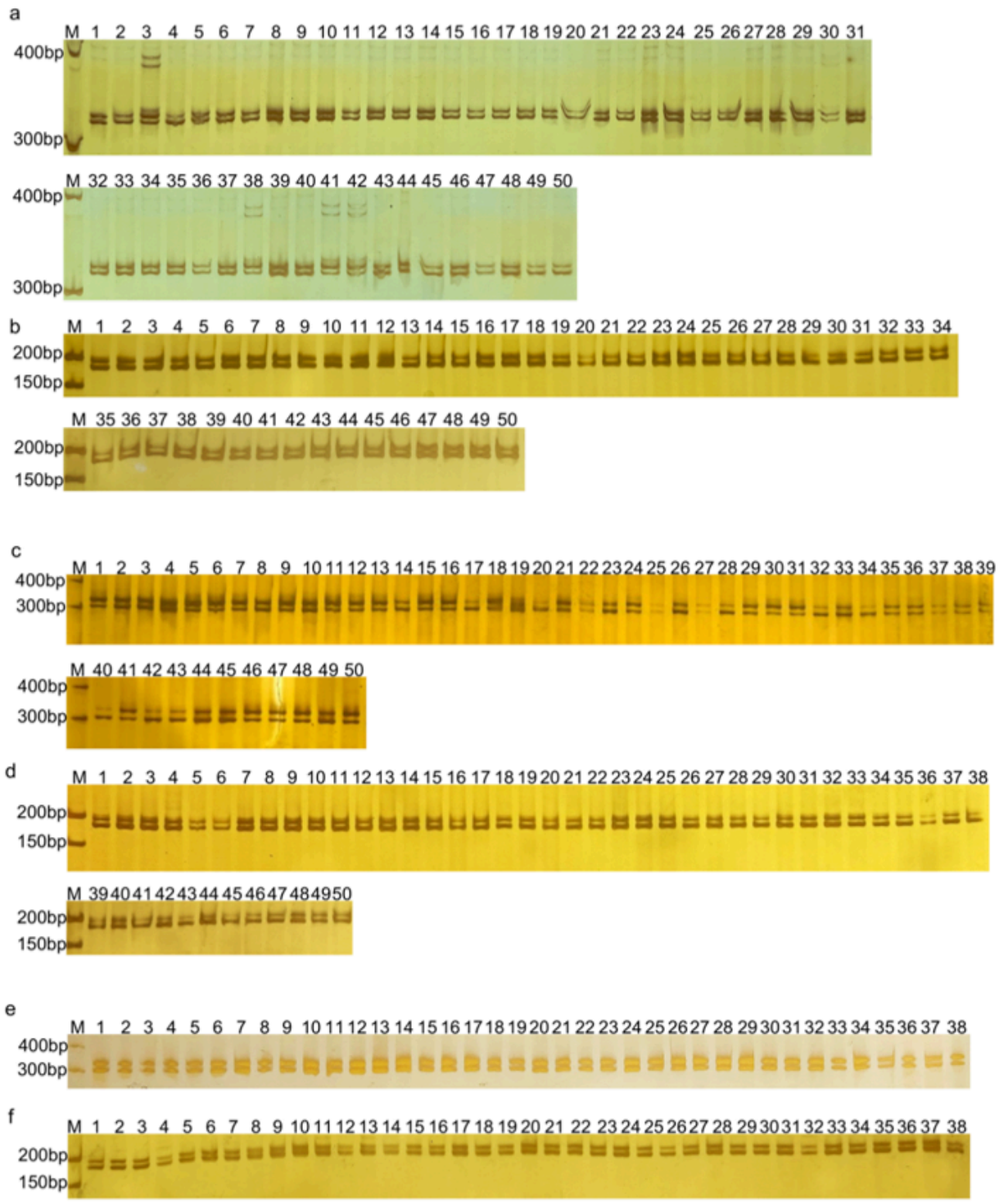


Figure 3

Partial polyacrylamide gel electrophoresis (PAGE) results of heterozygous loci in loquat. (a, c, e) Results for locus LG10:9288618; (b, d, f) results for locus LG11:7010931. (a, b) Hybrid progeny: samples 1–12 (SWZS×DYWB), 13–24 (HB1×JH1), 25–37 (HB1×PELL), and 38–50 (SWZS×JH1). (c, d) Reciprocal crosses: the first 1-50 samples (GZ-16×RTBS) and the last 1-50 samples (RTBS×GZ-16). (e, f) Self-pollinated progeny: samples 1–38 (HB1-SC) and 1–38 (HB2-SC). The Marker is the DL 500 DNA Marker.

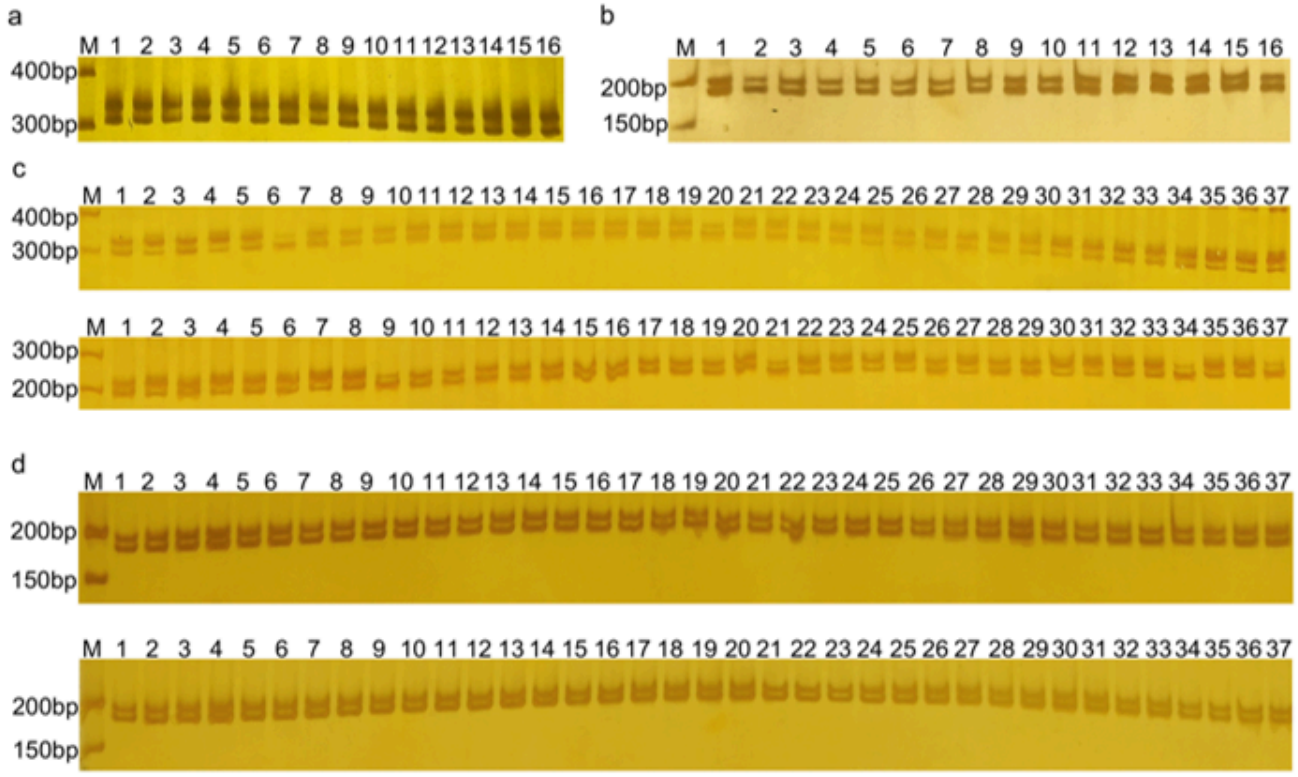


Figure 4

Gel electrophoresis (PAGE) of heterozygous loci in polyploid loquat progenies using two locus-specific primers. (a, c) Results for locus LG10:9288618; (b, d) results for locus LG11:7010931.(a, b) 1–16 is A313, A322, D25, B433, B456, A479, B431, B432, D27, LGY, LFZ, LZY, LHY1, H424, K376-1, and K474. (c, d) The first 1-37 samples (B431×PELL) and the last 1-37 samples (B431×GF). The marker is the DL 500 DNA marker.

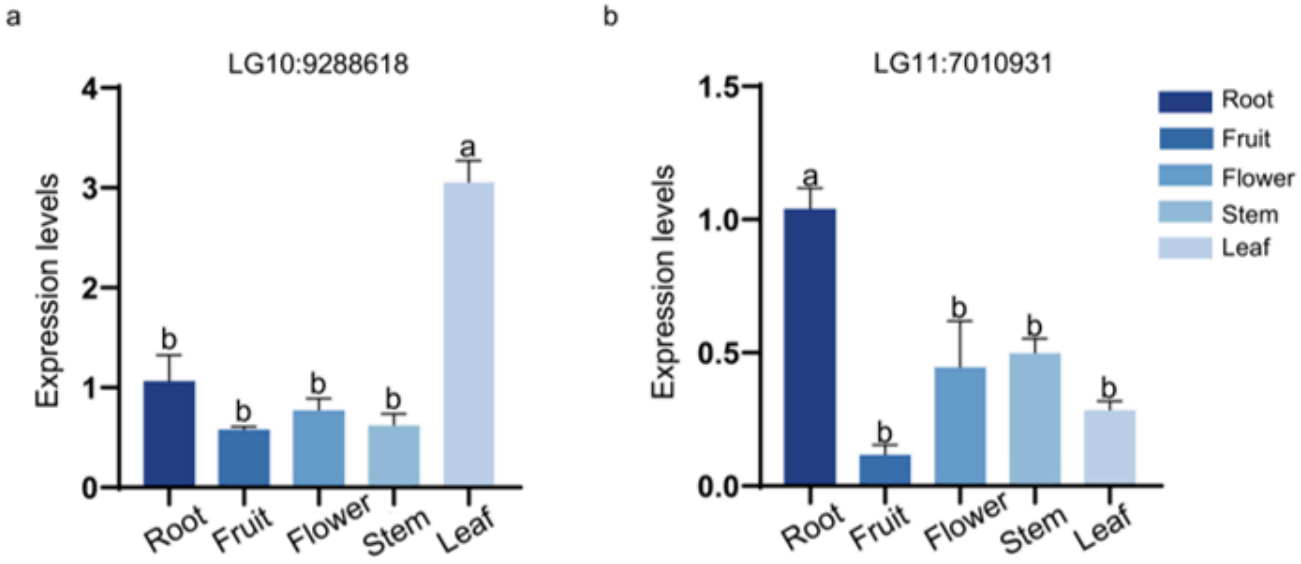


Figure 5

Expression levels of two heterozygous locus-associated genes in different tissues of loquat (‘Huabai No.1’). (a) Locus (LG010:9288618); (b) Locus (LG11:7010931). Gene expression in root, fruit, flower, stem, and leaf tissues was quantified using qPCR. Different letters above the bars indicate significant differences among tissues ($p < 0.05$)

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SupplementaryTableS1S12xlsx.xlsx
- SupplementaryFigureS1S17.pdf