Integration of genomics, transcriptomics and metabolomics identifies candidate loci underlying fruit weight in loquat

Ze Peng1,†, Chongbin Zhao1,†, Shuqing Li1, Yihan Guo1, Hongxia Xu3, Guibing Hu1, Zongli Liu1, Xiuping Chen2, Junwei Chen3, Shunquan Lin1, Wenbing Su1,2,*, Xianghui Yang1,*

1State Key Laboratory for Conservation and Utilization of Subtropical Agro-Bioresources and Key Laboratory of Innovation and Utilization of Horticultural Crop Resources in South China (Ministry of Agriculture and Rural Affairs), College of Horticulture, South China Agricultural University, Guangzhou, Guangdong 510642, China
2Fruit Research Institute, Fujian Academy of Agricultural Science, Fuzhou, Fujian 350013, China
3Institute of Horticulture, Zhejiang Academy of Agricultural Sciences, Hangzhou, Zhejiang 310021, China
†These authors have contributed equally to this work.
*Correspondence: Xianghui Yang (gzyxh@scau.edu.cn), Wenbing Su (suwenbing13@163.com)

Running title
Multi-omics insight into fruit weight in loquat

© The Author(s) 2022. Published by Oxford University Press. All rights reserved. This is an Open Access article distributed under the terms of the Creative Commons Attribution License http://creativecommons.org/licenses/by/4.0/, which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.
Abstract

Fruit weight is an integral part of fruit-quality traits and directly influences commodity values and economic returns of fruit crops. Despite its importance, the molecular mechanisms underlying fruit weight remain understudied, especially for perennial fruit tree crops such as cultivated loquat (Eriobotrya japonica Lindl.). Auxin is known to regulate fruit development, whereas its role and metabolism in fruit development remain obscure in loquat. In this study, we applied a multi-omics approach, integrating whole-genome resequencing-based quantitative trait locus (QTL) mapping with an F1 population, population genomics analysis using germplasm accessions, transcriptome analysis, and metabolic profiling to identify the genomic regions potentially associated with fruit weight in loquat. We identified three major loci associated with fruit weight, supported by both QTL mapping and comparative genomic analysis between small- and big-fruited loquat cultivars. Comparison between two genotypes with contrasting fruit weight performance through transcriptomic and metabolic profiling revealed an important role of auxin in regulating fruit development, especially at the fruit enlarging stage. The multi-omics approach identified two homologs of ETHYLENE INSENSITIVE 4 (EjEIN4) and TORNADO 1 (EjTRN1) as promising candidates controlling fruit weight. Moreover, three single nucleotide polymorphism (SNP) markers were closely associated with fruit weight. Results from this study provided insights from multiple perspectives into the genetic and metabolic controls of fruit weight in loquat. The candidate genomic regions, genes, and sequence variants will facilitate understanding the molecular basis of fruit weight and lay a foundation for future breeding and manipulation of fruit weight in loquat.

Introduction

Fruits are rich in a variety of nutrients needed by humans, such as vitamins, trace elements, and sugars, to maintain a healthy life¹. Fruit weight or size is one of the major factors influencing the quality and commodity value of fruits². Fruits with higher weight or bigger size often have a higher commodity value and generate higher economic returns. Therefore, revealing the mechanisms of fruit development underlying fruit-quality traits such as weight/size has been an important research area for many fruit tree species. Fruit weight or
size is influenced by many factors, such as genetics, environment, the interactions between the two, and cultural practices. Plant hormones have been recognized as one of the major factors affecting fruit development. Many studies have shown that auxin is an important hormone influencing fruit size, which is determined by cell number and cell size, through regulating cell division and cell expansion. A study in apple revealed that higher endogenous auxin levels may increase fruit size, and increased cell size and fruit size were observed after treatment using 1-naphthylacetic acid (NAA). Similarly, the external application of auxin promoted fruit development and increased the fruit diameter in satsuma mandarin and strawberry. Apart from auxin, many other plant hormones also play key roles in fruit growth and development, such as gibberellic acid (GA), cytokinin (CK), ethylene (ETH), abscisic acid (ABA), and brassinosteroids (BR).

Developmental regulation by endogenous auxin in plants is mediated by many processes, including the metabolism, such as biosynthesis and conjugation, transport, and signaling of auxin, the balance of which leads to auxin homeostasis. The biosynthesis of the predominant form of auxin, indole acetic acid (IAA), involves tryptophan aminotransferase of Arabidopsis/tryptophan aminotransferase-related (TAA1/TAR) and flavin monooxygenase (YUCCA). The conjugation of IAA with amino acids is mediated by proteins like Gretchen Hagen 3 (GH3) families. PIN and AUX/LAX proteins are essential to the polar transport of auxins, which is required for the tissue-specific auxin distribution. Aux/Indole-3-Acetic Acid (Aux/IAA) proteins are repressors of auxin-responsive reporter genes, and their degradation rate is modulated by auxin levels. At low auxin concentrations, the signal transduction is inhibited by Aux/IAAs that repress the transcription activities of auxin response factors (ARFs). At higher auxin concentrations, the proteasomal degradation of Aux/IAA transcriptional repressors is mediated by the transport inhibitor response 1/auxin signaling F-BOX protein (TIR1/AFB). Subsequently, ARFs are released, activating the auxin signaling. Many genes associated with hormone regulations have been identified in tomato, a model species for fruit development and ripening. Silencing of SlIAA17 in tomato led to larger fruits compared with the wild type, which was caused by increased cell size and thicker pericarp. In the same study, the cell expansion was found to be related with higher
ploidy levels, suggesting the interruption of the endoreduplication process. Silencing of \textit{SIPIN4} in tomato led to parthenocarpic fruits, suggesting a role of PIN proteins in auxin regulation of fruit set\textsuperscript{26}. ARFs are a member of the auxin signaling pathway and play an important role in fruit development by either repressing or activating the transcriptions of auxin-responsive genes\textsuperscript{27,28}. \textit{SIARF7} negatively regulates fruit set and development by modulating auxin responses\textsuperscript{29}. \textit{SIARF9} plays a negative role in cell division at early fruit development stages, and increased \textit{SIARF9} expressions in transgenic plants led to smaller fruits compared with the wild type\textsuperscript{30}. Genes controlling the levels of other endogenous hormones also play a role in fruit development. For example, in kiwifruit, the cytokinin biosynthetic gene (\textit{IPT}) increased fruit growth by promoting cell division and expansion\textsuperscript{31}.

Quantitative trait locus (QTL) mapping has been widely applied to identify genomic regions and genes associated with fruit weight and size. The first QTL controlling fruit size, \textit{FRUIT WEIGHT2.2} (\textit{FW2.2}), was identified and cloned in tomato\textsuperscript{32}. \textit{fw2.2} encodes a negative regulator of cell division, and changes in its regulation caused differences in fruit size\textsuperscript{33,34}. In addition to \textit{FW2.2}, a series of other loci controlling fruit weight were also identified, such as \textit{FW1.1}, \textit{FW2.1}, \textit{FW3.3}, \textit{FW4.1}, \textit{FW4.2}, and \textit{FW11.2}\textsuperscript{35,36}. \textit{SIKLUH}, a gene underlying \textit{FW3.2} in tomato, was found to control fruit and seed weight by regulating cell proliferation in the pericarp\textsuperscript{37,38}. \textit{FW11.3} was reported to control fruit weight by increasing the cell size\textsuperscript{39}. However, in perennial crops, QTL mapping is relatively more challenging due to a long juvenile phase and the high cost to maintain a mapping population. Currently, the QTL mapping approach has been applied in several perennial crops to identify loci underlying fruit weight/size, such as apple\textsuperscript{40,41}, pear\textsuperscript{42,43}, peach\textsuperscript{44,45}, grapevine\textsuperscript{46}, and blueberry\textsuperscript{47}. These studies revealed a complex and polygenic nature of the fruit weight/size trait, which is also influenced by environmental factors.

Transcriptome sequencing is an efficient approach to understanding transcriptional regulations associated with fruit weight/size. A study in pear compared transcriptome profiles between two accessions with contrasting fruit size at three early fruit developmental stages\textsuperscript{48}. Their results showed that the fruit size difference was mainly caused by the difference of cell
numbers, and that the small fruit size was potentially associated with the short cell division period due to early degradation of cytokinin and gibberellin\textsuperscript{48}. Another comparative transcriptome study in pear revealed that the auxin signal transduction was important for the fruit enlargement stage\textsuperscript{49}. Similarly in grapevine, according to transcriptome analysis on individuals with different berry weight performance, genes related with auxin metabolism were up-regulated in large berry genotypes, suggesting an important role of auxin in cell expansion\textsuperscript{50}.

Current knowledge on regulation mechanisms of fruit weight/size is mainly derived from annual plants, whereas relatively fewer studies are available for perennial plants. Cultivated loquat (\emph{Eriobotrya japonica} Lindl.; 2n=2x=34) is a subtropical evergreen fruit tree species native to China and belonging to the family Rosaceae\textsuperscript{51,52}. It has been cultivated for over 2000 years in China for its fleshy and tasty fruit and medicinal values, such as moisturizing the lungs, relieving cough, and reducing phlegm\textsuperscript{53}. Larger fruits are always desired in the market and have been an important goal for loquat breeding programs. Plant growth regulators, such as synthetic auxins 2,4-D\textsuperscript{54}, naphthalene acetic acid\textsuperscript{55}, and 3,5,6-TPA\textsuperscript{56}, have been applied to obtain bigger loquat fruits. Understanding the molecular mechanisms underlying fruit weight/size control will provide guidance on the manipulation of fruit weight/size, as well as facilitate breeding large fruit varieties in loquat. Currently, QTL mapping of fruit weight and transcriptome profiling directly comparing small- and big-fruited accessions have been rarely reported in loquat. Previously, the correlations between cell number/size and fruit weight were investigated in 13 loquat accessions by our group, which revealed cell size may play a predominant role in determining fruit weight\textsuperscript{57,58}. Besides, a repressor of fruit enlargement, \textit{EjBZR1}, was identified, which binds to the \textit{EjCYP90} promoter to repress its expressions and cell enlargement\textsuperscript{58}. However, the genetic basis and regulatory network controlling fruit weight remain unclear. More studies are needed to understand the genetic mechanisms controlling fruit weight in loquat.

In this study, to identify candidate genomic regions and shed light on mechanisms associated with fruit weight in loquat, we applied a combination of QTL mapping (using a
bi-parental F₁ population), population genomic analysis (using germplasm accessions), transcriptome profiling, and metabolic profiling. Results from this study reveal insights into the genetic and metabolic controls of fruit weight in loquat and provide an abundant resource for molecular breeding.

Results

Phenotypic data of the F₁ population

The phenotypic data of fruit weight of the F₁ population (130 F₁ plants) derived from the crosses between ‘Ninghaibai’ (female parent) and ‘Dafang’ (male parent), two loquat cultivars from China (Zhejiang) and Japan, respectively, were directly retrieved from our previous report⁵⁹ (Table S1). Briefly, the average fruit weight (g) of 30 mature fruits from each plant was used for analysis. The two parental plants had similar fruit weight (Table S1), but extensive variations were observed in that of the F₁ population (Fig. S1A, B). As previously reported, fruit weight showed a near-normal distribution (Fig. S1B, C), indicating its quantitative trait nature and polygenic genetic control. The phenotypic data were only recorded for one year in 2020, due to tree losses caused by cold damage in 2021.

Genotyping and high-density genetic map construction

Through whole-genome resequencing (WGRS), more than 21 Gb sequencing data (>27.00× coverage) were obtained for each parent, and an average of 8 Gb sequencing data per sample (9.57× coverage) were obtained for the F₁ individuals (Table S1). The reference genome of ‘Seventh Star’⁶⁰ was used for analysis since it was first released to the public at the initiation of this project. However, both the ‘Seventh Star’ and ‘Jiefangzhong’⁵² (released in 2021) genomes were used at the downstream candidate gene search stage. The average overall alignment rate was 97.17%. Variant calling and further filtering yielded a total of 2,184,538 single nucleotide polymorphisms (SNPs) between the two parents, with a Ti/Tv ratio of 2.48.

The above identified SNPs were used for genetic map construction with a Bin map method and HighMap software⁶¹. The SNPs were assigned to groups based on their physical
locations, and SNPs within each group were partitioned into bins after performing pair-wise linkage analysis and corrections of genotypes based on linkage phase. SNPs within the same bin were considered to have no recombination. A consensus map was constructed by using fully informative markers/bins with phasing information available for both alleles for each parent. Finally, a high-density genetic map comprised of 17 linkage groups (LGs), 3859 bins, and spanning a genetic distance of 1988.12 cM was obtained (Fig. 1A). The length of each LG ranged from 79.05 cM to 164.16 cM, with an average density of 0.52 cM/marker (Table 1). According to collinearity analysis, the Spearman correlation coefficients of all LGs were ≥0.9980 (Fig. S2), indicating high consistency between the reference genome and LGs, as well as high accuracy of recombination rate calculations.

**Figure 1.** Distribution of markers in the high-density genetic map and the quantitative trait loci associated with fruit weight. (A) The X-axis corresponds to each linkage group; the Y-axis represents the genetic distance (cM). (B) The X-axis represents each linkage group; the blue line corresponds to the LOD score of the left Y-axis; the red line corresponds to the phenotypic variation explained (%) of the right Y-axis. The horizontal grey line indicates the threshold of LOD with 99.5% confidence interval (LOD=6.2).
Table 1. Summary statistics of 17 linkage groups.

| Linkage group | No. of loci (bins) | Map distance (cM) | Average distance between loci (cM) | Maximum gap (cM) | Gaps < 5 cM (%) |
|---------------|--------------------|-------------------|-----------------------------------|-----------------|-----------------|
| 1             | 210                | 137.25            | 0.65                              | 12.13           | 99.52           |
| 2             | 241                | 92.32             | 0.38                              | 2.36            | 100.00          |
| 3             | 191                | 119.65            | 0.63                              | 7.90            | 99.47           |
| 4             | 343                | 132.16            | 0.39                              | 4.84            | 100.00          |
| 5             | 203                | 122.06            | 0.60                              | 10.20           | 98.02           |
| 6             | 349                | 164.16            | 0.47                              | 7.01            | 99.43           |
| 7             | 164                | 94.24             | 0.57                              | 4.84            | 100.00          |
| 8             | 127                | 79.05             | 0.62                              | 6.13            | 98.41           |
| 9             | 209                | 97.35             | 0.47                              | 10.68           | 99.52           |
| 10            | 282                | 143.59            | 0.51                              | 9.73            | 98.22           |
| 11            | 165                | 101.00            | 0.61                              | 7.90            | 98.17           |
| 12            | 292                | 136.56            | 0.47                              | 7.90            | 99.31           |
| 13            | 207                | 79.46             | 0.38                              | 4.84            | 100.00          |
| 14            | 179                | 116.79            | 0.65                              | 5.27            | 99.44           |
| 15            | 288                | 143.41            | 0.50                              | 5.70            | 99.65           |
| 16            | 238                | 116.24            | 0.49                              | 6.57            | 99.16           |
| 17            | 171                | 112.83            | 0.66                              | 8.35            | 99.41           |
| Total         | 3859               | 1988.12           | -                                 | -               | -               |

QTL identification for fruit weight

Quantitative trait loci (QTLs) were identified for fruit weight based on the high-density linkage map and the phenotypic data. In total, 17 QTLs were detected on eight LGs, corresponding to eight chromosomes based on the ‘Seventh Star’ genome, including Chr3, Chr4, Chr8, Chr12, Chr14, Chr15, Chr16, and Chr17 (Fig. 1B, Table 2). The large number of QTLs reflected the complex and polygenic nature of fruit weight. Among these QTLs, we identified three major QTLs on Chr8, Chr12, and Chr15, which had a phenotypic variation explained (PVE) value ranging between 20.0%~49.7%. Sequences surrounding the flanking markers of the QTLs were extracted from the ‘Seventh Star’ genome and compared to the ‘Jiefangzhong’ genome to identify the corresponding coordinates. Both the physical sizes and the numbers of annotated genes of the corresponding genomic regions of these 17 QTLs were highly similar between these two reference genomes (Table S2). The physical sizes of the 17 QTLs ranged from 0.09 to 23.16 Mb covering a total of 6,113 genes in the ‘Seventh Star’ genome, while the physical sizes ranged from 0.09 to 23.29 Mb covering 6249 genes in the ‘Jiefangzhong’ genome.
Table 2. Summary information of quantitative trait loci associated with fruit weight.

| Linkage group | Start (cM) | End (cM) | No. of bins | LOD range          | PVE range (%) |
|---------------|------------|----------|-------------|--------------------|---------------|
| 3             | 67.518     | 105.182  | 79          | 6.22~10.86         | 19.8~31.9     |
| 3             | 65.587     | 66.36    | 6           | 6.41~6.87          | 20.3~21.6     |
| 3             | 60.544     | 61.319   | 2           | 6.21~6.38          | 19.7~20.2     |
| 4             | 8.272      | 29.533   | 20          | 6.38~11.66         | 20.2~33.8     |
| 4             | 51.836     | 76.221   | 143         | 6.68~10.29         | 21.1~30.6     |
| 8             | 0          | 28.415   | 6           | 6.63~18.25         | 20.9~47.6     |
| 12            | 14.231     | 74.782   | 124         | 6.31~19.4          | 20.0~49.7     |
| 14            | 83.081     | 88.899   | 12          | 6.94~8.1           | 21.8~24.9     |
| 15            | 80.128     | 143.41   | 164         | 6.3~15.07          | 20.0~41.4     |
| 15            | 25.859     | 32.088   | 7           | 6.61~6.71          | 20.9~21.1     |
| 16            | 31.388     | 32.163   | 5           | 6.81~7.83          | 21.4~24.2     |
| 16            | 72.154     | 83.911   | 9           | 6.71~9.29          | 21.1~28.1     |
| 16            | 87.021     | 90.502   | 14          | 6.35~6.98          | 20.1~21.9     |
| 16            | 106.093    | 110.748  | 11          | 7.41~9.78          | 23.1~29.3     |
| 17            | 86.08      | 87.624   | 7           | 6.53~7.11          | 20.7~22.3     |
| 17            | 21.725     | 22.111   | 2           | 9.46~9.48          | 28.5~28.5     |
| 17            | 47.753     | 52.401   | 14          | 6.23~7.33          | 19.8~22.9     |

Population genomics of small- and big-fruited cultivars

To gain additional evidence for identification of the genomic regions potentially associated with fruit weight, we also obtained WGRS data for a total of 20 commonly known small- or big-fruited loquat cultivars (Table S3). WGRS data were generated in this study for 15 genomes, with an average coverage of 10.85×, while we also utilized published WGRS data for five genomes with an average coverage of 8.45×. Variant calling and filtering identified a total of 3,697,866 SNPs, which were utilized for the following analyses. Phylogenetic analysis revealed that small- or big-fruited cultivars were not specific to a branch or a cluster (Fig. 2A, B). Although this set of germplasm accessions were limited in number, they represented diverse genetic backgrounds or clusters. Calculation of $F_{st}$ using a sliding window approach identified genomic regions genetically highly differentiated between the small- and big-fruited groups. By focusing on genomic regions above the top 1% $F_{st}$ line ($F_{st}$=0.31), we found a large proportion of Chr1 highly differentiated between the two groups. Moreover, we noticed a region of Chr8 containing the highest peak above the top 0.1%
$F_{st}$ line ($F_{st}=0.42$). Interestingly and importantly, this highly differentiated region between the two groups on Chr8 overlapped with a major QTL associated with fruit weight on the same chromosome (Fig. 2D). Similarly, the highest $F_{st}$ peaks on Chr12 and Chr15 also overlapped with the two major QTLs. In summary, the overlapped regions on Chr8, Chr12, and Chr15, with evidence from both the bi-parental QTL mapping approach and $F_{st}$ analysis utilizing germplasm accessions, may play an essential role in determining fruit weight of loquat, and they can be prioritized for candidate gene search.
Figure 2. Genetic relationship and population-differentiation of small- and big-fruited cultivars, and comparison of QTL regions and top $F_{st}$ regions potentially associated with fruit weight. (A) and (B) The black color indicates small-fruited cultivars; the red color indicates big-fruited cultivars. The phylogenetic tree was constructed using PHYLIP with 100 bootstraps and visualized using MEGAX software. Principal component analysis (PCA) was carried out using Plink. (C) Distribution of $F_{st}$ comparing small- and big-fruited cultivars. Each dot represents a sliding window of 10 Kb, within which the average $F_{st}$ value was calculated. The blue line represents the top 1% $F_{st}$ line ($F_{st}=0.31$); the brown line represents...
the top 0.1% $F_{st}$ line ($F_{st}$=0.42). (D) illustrates the comparisons for Chr8, Chr12, and Chr15, which contain three major QTLs for fruit weight.

Comparative transcriptome analysis of two sister lines with contrasting fruit weight

To investigate transcriptional regulations related with fruit development and facilitate candidate gene searching, we obtained transcriptome profiles at five different developmental stages for two sister lines ZP44 and ZP65 (Fig. 3A) derived from the cross between Zaozhong No. 6 (female parent) and Peluches (male parent). ZP44 had extremely small fruits (11~15 g), while ZP65 had extremely large fruits (up to 82.69 g). These two genotypes had relatively bigger fruit weight difference compared with those observed in the F1 population. Their fruit development processes were observed by measuring the fruit diameter (transverse diameter) due to its high correlation coefficient (0.964) with fruit weight. Data were recorded at 0, 7, 14, 28, 42, 56, 63, 77, 84, and 91 days past anthesis (DPA) (Fig. 3B). The fruit size difference between the two genotypes was small at the earlier stages. However, the fruit diameter of ZP44 reached a plateau at 56 DPA, whereas the fruits of ZP65 were still enlarging and reached a plateau at 77 DPA. Five representative stages were selected for transcriptome sequencing, including 0 (S1), 7 (S2), 28 (S4), 56 (S6), and 77 (S8) DPA. The phenotypic observations revealed that S6 is a key developmental stage, at which the fruit development of ZP44 and ZP65 diverged and became apparently different.
Figure 3. Observation of fruit development and differentially expressed genes identified from the RNA-seq experiment. (A) Observations of fruit development of ZP44 and ZP65 at 10 time points, including 0 days past anthesis (0D, S1), 7D (S2), 14D, 28D (S4), 42D, 56D (S6), 63D, 77D (S8), 84D, and 91D. (B) The comparison of fruit diameter (transverse diameter) between ZP44 and ZP65. The values are the mean fruit diameter of 15 fruits. The error bar is for standard error. (C) Principal component analysis using transcriptome profiles of all 30 replicates. (D) Summary of differentially expressed genes (DEGs) comparing ZP44 and ZP65 at each stage. (E) The comparison of DEGs at each stage.

An average of 7 Gb transcriptome sequencing data per sample were obtained for a total
of 30 samples (2 genotypes × 5 stages × 3 biological replicates) (Table S4). To utilize and improve the gene annotation of the ‘Jiefangzhong’ reference genome generated by our group, reads were mapped to the ‘Jiefangzhong’ genome with an average overall mapping rate of 95.01%. Consistent with phenotypic observations, the largest difference between ZP44 and ZP65 was observed for stage S6 according to the principal component analysis of the 30 transcriptome profiles (Fig. 3C). The differentially expressed genes (DEGs) between ZP44 and ZP65 at each stage were identified and compared (Table S5). Similarly, S6 had the largest number of DEGs with 3,267 up-regulated genes and 2,623 down-regulated genes (Fig. 3D), and it had the largest number of stage-unique DEGs (2,579 for ZP44_S6 vs ZP65_S6) (Fig. 3E). Therefore, both phenotypic observations and transcriptome profiling suggested some unique changes happened at the S6 stage, which may be associated with the larger fruit weight difference between the two genotypes. Gene ontology (GO) enrichment analysis was performed for the DEGs between ZP44 and ZP65 at each stage. Strikingly, eight hormone-related GO terms were significantly enriched for the DEGs at S6, including “regulation of hormone levels”, “hormone metabolic process”, “cellular response to auxin stimulus”, “auxin-activated signaling pathway”, “response to gibberellin”, “hormone catabolic process”, “response to auxin”, and “hormone biosynthetic process” (Table S6). Interestingly, the three auxin-related GO terms were only enriched specifically at the S6 stage, suggesting auxin-related genes may play an important role at this key stage when the fruit was enlarging. In addition to auxin and gibberellin, we also noticed the enrichment of salicylic acid-related GO terms at the S4 stage, indicating its role at the early fruit development stage. The pathway enrichment analysis identified numerous sugar-related pathways significantly enriched at the S6 stage (Table S7), suggesting the difference of sugar accumulations corresponding to the different fruit size of the two genotypes.

To further identify the different expression patterns during fruit development between ZP44 and ZP65, the DEGs were clustered into groups using the K-means method (Table S8, Fig. 4). A total of 16 sub-classes or types were identified, among which eight sub-classes showed obviously different expression patterns between the two genotypes, including sub-classes 2, 4, 5, 7, 8, 12, 13, and 14. Sub-classes 4 and 8 contained 434 and 197 genes that
were differentially expressed at all stages between the two genotypes, respectively. Genes belonging to sub-classes 2, 7, and 13 showed different expression patterns at the S6 stage between the two genotypes. By comparing the genomic locations of the DEGs with the QTLs and top $F_{st}$ regions, we found 311 DEGs located only in QTLs, 16 DEGs located only in top 1% $F_{st}$ regions, while eight DEGs located in the overlapped regions. These genes with different expression patterns will facilitate following candidate gene searching.

Figure 4. K-means clustering of differentially expressed genes. The number in black font refers to the number of genes in each sub-class; the number in red font represents the number of differentially expressed genes between ZP44 and ZP65 for the ‘star’ labeled time point.

Auxin metabolite profiling

Considering a possibly important role of auxin metabolism in fruit development, the auxin metabolic status was investigated at the same five stages for ZP44 and ZP65 (Table S9). A total of 17 types of auxin metabolites were detected in these samples. The clustering based on metabolite levels revealed a general trend that the levels of most detected metabolites were higher for the first three stages than that of the latter two stages (Fig. 5A). The levels of many metabolites significantly decreased when comparing S1 vs S6 or S1 vs S8 within each
genotype (Table S10, S11). This result suggested that auxin signaling and metabolism may be more active at earlier fruit developmental stages than the mature fruit stage. Interestingly, we found significant differences in metabolite levels for several metabolites between the two genotypes, including L-tryptophan (TRP), 3-Indoleacetonitrile (IAN), N-(3-Indolylacetyl)-L-alanine (IAA-Ala), and 2-oxindole-3-acetic acid (OxIAA) (Fig. 5B-E). Importantly, the precursor of auxin, TRP, still maintained high levels in ZP65 at stages S4, S6, and S8, indicating high auxin levels, whereas its levels decreased sharply in ZP44. Therefore, consistent with findings from transcriptome analysis, the auxin metabolite profiling validated the higher auxin levels in ZP65 than that in ZP44, especially at the fruit enlarging stage. These results suggested that auxin metabolism shall be one of the key factors determining the fruit weight/size difference between ZP44 and ZP65.
Figure 5. The clustering of auxin metabolites and significantly different metabolites between ZP44 and ZP65. (A) A heat map showing auxin metabolites levels. Data were processed using the z-score normalization method. (B)-(E) Comparisons of metabolite levels for L-tryptophan (TRP), 3-Indoleacetonitrile (IAN), N-(3-Indolylacetyl)-L-alanine (IAA-Ala), and 2-oxindole-3-acetic acid (OxIAA).

Candidate genes and markers potentially associated with fruit weight

To search for candidate genes potentially associated with fruit weight, the following criteria were considered: 1) whether located in QTLs; 2) whether located within or adjacent to top 1% $F_{st}$ regions; 3) whether differentially expressed between ZP44 and ZP65; 4)
whether homologous to known genes associated with fruit weight or related functions (such as cell proliferation, organ size, etc.); 5) whether contain or close to markers highly associated with fruit weight (first assessed based on predicted genotypes from GATK, then validated using Sanger sequencing); and 6) whether contain polymorphisms in functional domains between small- and big-fruited genotypes. Due to the potential different genetic backgrounds of the plant materials used for QTL analysis, $F_{st}$ analysis, and RNA-seq, a candidate gene may not necessarily meet all the criteria. After mining, we identified two candidate genes with relatively more evidence, including a homolog of ETHYLENE INSENSITIVE 4 (EjEIN4) and a homolog of TORNADO 1 (EjTRN1) (Table 3). Three other genes were also summarized, although with relatively less evidence, including three homologs of Auxin Response Factor 5 (EjARF5), S-adenosylmethionine decarboxylase proenzyme 1 (EjSAMDC1), and Growth-Regulating Factor 1 (EjGRF1) (Table 3). For visualization, the QTLs, top 1% $F_{st}$ regions, the DEGs located in these regions, candidate genes, and markers highly associated with fruit weight were plotted to a map for both reference genomes (Fig. 6, Table S12, S13).

Table 3. Summary of evidence for candidate genes potentially associated with fruit weight in loquat.

| Evidence                                                                 | EjEIN4 | EjTRN1 | EjARF5 | EjGRF1 | EjSAMDC1 |
|-------------------------------------------------------------------------|--------|--------|--------|--------|----------|
| Within QTLs                                                             | Yes    | No     | Yes    | Yes    | Yes      |
| Within or adjacent to top 1% $F_{st}$ regions                           | Yes    | Yes    | Yes    | Yes    | Yes      |
| Differentially expressed between ZP44 & ZP65                            | No     | Yes    | No     | Yes    | Yes      |
| Homologous to known genes associated with fruit weight or related functions | Yes    | Yes    | Yes    | Yes    | Yes      |
| Contain markers highly associated with fruit weight                     | Yes    | Yes    | Yes    | Close* | NA       |
| Contain polymorphisms in functional domains between small- and big-fruited genotypes | Yes    | Yes    | No     | NA     | NA       |

*Marker 1 is 127,731 bp away from EjGRF1.
Figure 6. Distribution of QTL regions, top 1% $F_{st}$ regions, and candidate genes potentially associated with fruit weight in two reference genomes of loquat. The explanations for color codes of regions and genes are at the bottom of the figure. The connecting lines between the two genomes indicate the same chromosomes.

In total, three SNP markers highly associated with fruit weight, including Marker 1, 2 (2’), and 3, were identified using a validation panel comprised of the extremely small- and big-fruited F$_1$ individuals, as well as the commonly known small- and big-fruited loquat cultivars/accessions (Fig. 7A-C). Marker 1 was identified at 395 bp upstream of EjEIN4. Marker 2 and 2’ were located at 1,524 bp upstream of EjARF5 and within its 8$^{th}$ intron, respectively. Marker 3 was located at the coding region of EjTRN1. Alleles of Marker 1 and 2 (2’) were segregating in the F$_1$ population, whereas the genotype of Marker 3 was fixed for the F$_1$ population and only showed high association with fruit weight for the natural germplasm accessions. By utilizing the predicted genotypes of Marker 1 and 2 from GATK, the fruit weight for the F$_1$ individuals with different genotype combinations were compared.
Marker 2 was used for the comparison due to more available data for the minor class of genotype combinations than that of Marker 2’. The results showed that the combination ‘G/G and A/A’ had lower fruit weight, ‘G/C and A/C’ had relatively higher fruit weight, while those with only one ‘big-fruited’ genotype (either ‘G/C’ and ‘A/A’ or ‘G/G’ and ‘A/C’) were somewhere in the middle (Fig. 7A). Above results showed high correlations between fruit weight and the identified markers.

Figure 7. Single nucleotide polymorphism markers closely associated with fruit weight. (A) The distributions of fruit weight for F1 individuals with predicted genotypes from GATK for Marker 1 and 2. The predicted genotypes of Marker 2 were used for the comparison due to more available data for the minor class of genotype combinations than that of Marker 2’. The
genotypes of Dafang were ‘C/G’ at Marker 1 and ‘A/C’ at Marker 2. The genotypes of Ninghaibai were ‘G/G’ at Marker 1 and ‘A/A’ at Marker 2. The number of plants in each column corresponds to the genotype combination in that column. (B) Three genotypes at the four SNP loci based on Sanger sequencing from the validation panel. (C) The genotypes obtained from Sanger sequencing for the validation panel at Marker 1, 2, 2’, and 3.

To investigate the polymorphisms in the gene regions, we obtained the full-length sequences for the three genes tagged with markers highly associated with fruit weight, including \( \text{EjEIN4}, \text{EjTRN1}, \) and \( \text{EjARF5} \). The whole gene region of \( \text{EjTRN1} \) was covered deeply by uniquely mapped Illumina reads for ZP44 and ZP65 (both are homozygous for the whole gene region). Therefore, the sequences were manually retrieved by visualizing the alignment file. For \( \text{EjEIN4} \) and \( \text{EjARF5} \), we aligned the 3rd generation PacBio HiFi reads of Puye (available in-house from our genome projects) to the Jiefangzhong reference genome, and manually retrieved the gene sequences by visualizing the alignments. The accession Puye was used due to its homozygous ‘small-fruited’ genotypes at Marker 1 and Marker 2, as well as due to the availability of HiFi reads data. Strikingly, \( \text{EjTRN1} \) turned out to be a highly promising candidate gene. In Arabidopsis, \( \text{TRN1} \), an auxin-signaling pathway gene, was involved in auxin transport, cell proliferation, and plant growth\(^62\). \( \text{EjTRN1} \) was differentially expressed between ZP44 and ZP65 (Fig. 8A). Consistently, we identified a large 21-bp deletion in the promoter of ZP44, which was also validated using PCR and gel electrophoresis (Fig. 8B). By comparing the predicted protein sequences between ZP44 and ZP65, we found that the Marker 3, a SNP marker, led to an amino acid change (M/T) within its leucine-rich repeat (LRR) domain, making it a functional marker, while this amino acid change was also highly associated with fruit weight (Fig. 7C). After genotyping the validation panel at the Indel locus, we found that only ZP44 and Tongpi, two small-fruit ed accessions, had the homozygous ‘deletion’ genotype, although a few other accessions were also carrying the ‘deletion’ (Fig. S3). Consequently, it turned out to be Marker 3 that was highly correlated with fruit weight instead of the Indel locus, although it is also likely that this Indel may also contribute to the small-fruit phenotype by influencing the expressions of \( \text{EjTRN1} \). In Arabidopsis, \( \text{EIN4} \), an ethylene receptor gene, was reported to interact with \( \text{SAUR76}, \text{SAUR77}, \) and \( \text{SAUR78} \), which were associated with plant growth and cell expansion\(^63\). Interestingly, by comparing the predicted protein sequences of \( \text{EjEIN4} \) between Jiefangzhong...
and Puye, two homozygous ‘big-fruited’ and ‘small-fruited’ genotypes at Marker 1, we identified at least one amino acid difference within each of its domains (Fig. S4). Based on these results, *EjTRN1* and *EjEIN4* were considered as two promising candidates associated with fruit weight in loquat.

![Graph and Diagram](image)

**Figure 8.** The differential expression patterns and polymorphisms of *EjTRN1* between ZP44 and ZP65. (A) The relative expressions of *EjTRN1* at five developmental stages of ZP44 and ZP65. (B) An insertion/deletion identified at the promoter of *EjTRN1* with PCR validation. (C) The comparison of protein sequences of *EjTRN1* between ZP44 and ZP65. The functional domains were predicted using the conserved domain search tool of NCBI.

In tomato, *SlARF5* was reported to regulate fruit development and influence fruit size.
and weight. For *EjARF5*, we finally obtained the sequences of two haplotype alleles of Puye that only differ at one amino acid towards the end of the predicted protein (Fig. S5). By comparing these protein sequences (the small-fruited genotype) with that of Jiefangzhong (the big-fruited genotype), we did identify three amino acid differences and one Indel (two amino acids insertion/deletion) within its coding region (Fig. S5). However, these polymorphisms were not located within its predicted functional domains. It might be possible that the causal polymorphism(s) may be at other sites. However, it might also be possible that it was another unknown gene in proximity with Marker 2 (2') instead of *EjARF5* underlying the fruit weight variation associated with this locus. More future experiments can be designed to further validate its function.

In addition to above genes, *EjSAMDC1* and *EjGRF1* were differentially expressed between ZP44 and ZP65 (Fig. S6A, B). In apple, *MdSAMDC1* was reported to be involved in fruit development and cell growth. Strikingly, the expression of *EjSAMDC1* was known to be down-regulated by auxin in Arabidopsis. By investigating the expression patterns, we found that the expression of the *EjSAMDC1* homolog was significantly much higher in ZP44 than that in ZP65 for the S6 stage (Fig. S3). It was likely that the auxin levels were much higher in ZP65 at the S6 stage, therefore repressing the expression of *EjSAMDC1*, in contrast to the low auxin levels in ZP44 at the S6 stage, where the expression of *EjSAMDC1* was not repressed. *EjGRF1* was also adjacent to Marker 1 (127,731 bp away) highly associated with fruit weight. In Arabidopsis, *GRF1* was reported to play a role in the regulation of cell expansion and influence the organ sizes. Therefore, these two genes may also be worthy of future in-depth investigations. Collectively, the multi-omics approach identified several candidate genes potentially associated with fruit weight in loquat.
**Discussion**

Fruit weight is an integral part of fruit-quality traits for fruit tree crops such as loquat, which directly influences the commercial value and consumer acceptance. Despite its importance, the associated markers and underlying genetic mechanisms, which can greatly facilitate breeding cultivars with optimal fruit weight performance, remain understudied in loquat. This was largely caused by a major obstacle specifically for perennial trees that usually need to undergo a juvenile phase of many years before entering the reproduction stage and bearing fruits. Plant hormones, especially auxin, have been reported to play important roles in regulating fruit development\(^{69,70}\). However, the role of auxin and its metabolism remain largely unexplored in loquat. To address these questions, we applied a multi-omics approach, integrating WGRS, transcriptome sequencing, and metabolite profiling, to identify genomic regions potentially associated with fruit weight. Our results revealed insights into the genetic and metabolic controls of fruit weight, and provided an abundant resource of genomic regions, genes, and sequence variants, which will enable future in-depth investigations and facilitate loquat breeding.

High-density genetic maps are an important genomic resource and a prerequisite for QTL mapping and map-based cloning\(^{40,45,71,72}\). However, few studies on genetic maps were available in loquat. With WGRS data (~10× coverage), we were able to genotype the F\(_1\) population in a genome-wide manner. Due to the large number of co-segregating SNPs, we applied a Bin method, which has been applied previously\(^{61,73,74}\), to group SNPs without recombination into the same bin, and created a consensus map. Previously, a high-density genetic map was constructed for bronze loquat, comprised of 960 loci, and spanning 1707.4 cM with a density of 1.78 cM/marker\(^{72}\). The genetic map in current study spanned 1988.12 cM, covering all 17 chromosomes with a density of 0.52 cM/marker, which is considerably higher than that in previous studies based on amplified fragment length polymorphisms (AFLPs), simple sequence repeats (SSRs), or SNPs using RAD-seq\(^{72,75,76}\).

The high-density genetic map and phenotypic data enabled the identification of 17 QTLs potentially associated with fruit weight in loquat. The number of QTLs for loquat fruit weight
seems to be higher than that of several other crops, such as blueberry (seven QTLs)
47, grapevine (six QTLs)
46, peach (four QTLs)
45, and apple (two QTLs)
42. This larger number may be explained by several possible reasons. Firstly, it is likely due to the complex and polygenic nature of this trait, or due to the nature of this crop. The different genetic backgrounds of parental genotypes used in these studies may also explain the QTL number differences, since fewer loci may be identified if the parental genotypes are fixed at some loci that have no segregation in the population. Secondly, it might be attributed to the limited number of F1 individuals, which may only reveal a rough genomic region, or due to the limitation of one year of phenotypic data. However, we did observe three major QTLs explaining 20.0%~49.7% of the phenotypic variation. It seems that compared with QTLs with minor effects
45,47, major effect QTLs may be more stable across different years or environments, such as the example in grapevine
46. The major fruit weight QTL with 61% PVE in that study was stable across different studies
77,78. The QTLs identified in the current study can be further validated in other populations or compared with other studies in the future to identify stable QTLs.

Considering the large number of QTLs and relatively large corresponding genomic regions, we also utilized WGRS data of commonly known small- or big-fruited loquat cultivars or accessions from the natural population preserved in the germplasm, whose fruit weight has been tested for many years. Corroborating our QTL results, the top Fst peak on Chr8 was highly consistent with the QTL peak on the same chromosome. Therefore, the Fst analysis based on loquat germplasm accessions with diverse genetic backgrounds supports the existence of a major locus associated with fruit weight on the lower part of Chr8. It is highly possible that the alleles controlling fruit weight at this locus are prevalent in loquat germplasm. Similarly, the Fst peaks on Chr12 and Chr15 and located within the two major QTLs also support the existence of two other major loci associated with fruit weight. We believe the three loci on Chr8, Chr12, and Chr15 play an important role in controlling fruit weight in loquat, as evidenced by both QTL-mapping and Fst analysis. Apart from these three major loci, we also observed many genomic regions highly differentiated between small- and big-fruited loquat accessions on other chromosomes, which is consistent with the large
number of QTLs identified in the bi-parental QTL mapping approach. Collectively, the overlapped regions between QTLs and the top $F_{st}$ regions are good candidate genomic regions potentially controlling fruit weight in loquat.

However, the results from QTL analysis, $F_{st}$ analysis, and RNA-seq analysis should be interpreted with caution, since different materials were used in these analyses, except that the two materials used for RNA-seq (ZP44 and ZP65) are offspring of two big-fruited accessions (Zaozhong No. 6 and Peluches) included in the $F_{st}$ analysis. We would propose that the overlapping results between the QTL analysis and $F_{st}$ analysis can be a coincidence while also may have significant relevance and implications. The germplasm accessions used for $F_{st}$ analysis represent diverse genetic backgrounds, covering likely most of the gene pool involving fruit weight/size. In comparison, the two parental cultivars used for the QTL analysis would represent relatively narrower genetic backgrounds, since it is likely that some genes/alleles controlling fruit weight/size may already be fixed (not segregating) in the F$_1$ population. From this perspective, the overlapping results could be a coincidence since it is unknown in terms of which genes from the whole gene pool are represented or segregating in the F$_1$ population. An example would be Marker 3 (located within gene EjTRN1), whose genotypes/alleles are fixed or not segregating in the F$_1$ population, but highly associated with fruit weight in the germplasm validation panel. From another perspective, the genomic regions with high $F_{st}$ values indicate that they are highly differentiated between the small- and big-fruited groups. These $F_{st}$ peaks were captured because the SNP alleles in those regions are quite different between the two groups. Therefore, the overlapping results could cover genes/alleles that were not only segregating in the F$_1$ population, but also common/universal in the loquat germplasm. In other words, this could cover genes/alleles that were recombined after human selection and breeding, during which the big-fruited alleles may have accumulated in the big-fruited group, while the small-fruited alleles may have accumulated in the small-fruited group. Both Marker 1 and Marker 2 (located in both QTLs and top $F_{st}$ regions) can serve as the examples, since they are highly associated with fruit weight for both the F$_1$ individuals and the germplasm accessions in the validation panel.
Since the two materials, ZP44 and ZP65, used for RNA-seq are different with that of the QTL and Fst analysis, the genes underlying fruit weight/size in the F1 population, or in the germplasm accessions may not be necessarily differentially expressed between ZP44 and ZP65. Therefore, when searching for candidate genes, we not only focused on the DEGs, but also used a series of other criteria, such as whether they are tagged with markers highly associated with fruit weight, whether homologous to known genes associated with fruit weight, etc. Consistent with the fact that ZP44 and ZP65 were derived from the cross between two big-fruited cultivars, it is highly likely that they are of the ‘big-fruited’ genotype for many loci, including those tagged with Marker 1 and Marker 2. Supported by plenty of evidence, EjTRN1 (tagged with Marker 3) can be one of the promising candidate genes that could explain the extreme small fruit weight of ZP44.

An important role of auxin in regulating fruit weight has been revealed from both transcriptomic and metabolic perspectives. To facilitate candidate gene search, we applied comparative transcriptome analysis for two sister lines, ZP44 and ZP65, which were derived from a cross between two commonly known big-fruited cultivars, Zaozhong No. 6 and Peluches. It is interesting that the cross between two big-fruited cultivars led to ZP44, an extremely small-fruited line. Therefore, genes or alleles from Zaozhong No. 6 and Peluches may have recombined in ZP44 that likely accumulated alleles leading to small fruits. Both phenotypic observations and transcriptome profiling of these two contrasting lines suggested the key divergence leading to fruit size difference was at the S6 stage, when the fruits are quickly expanding. Interestingly, both functional enrichment analysis and metabolomics analysis revealed a key role of auxin at this important stage. Auxin is an important plant hormone closely involved in the initiation of fruit development and determining fruit size by regulating cell division and expansion7,29. Its vital role in fruit development is implied in the fact that the highest auxin levels are usually observed in developing fruits compared with that in other parts of a plant79. In accordance with this, we also observed relatively higher levels of auxin metabolites at earlier fruit development stages compared with that of mature fruit stages. Moreover, a few studies have already reported the effect of auxin application on fruit development in loquat56,80. One study identified that the application of synthetic auxin 2,4-DP
increased the fruit size of loquat, also leading to earlier maturation and harvest time\textsuperscript{80}. Similarly, another study also revealed that the treatment of synthetic auxin 3,5,6-TPA led to larger fruit size by decreasing cell turgor pressure\textsuperscript{56}. In the current study, we identified several auxin-related candidate genes, such as \textit{EjTRN1}, \textit{EjARF5}, and \textit{EjSAMDC1} (Table S13). This result implies that variations of auxin-related genes may be recombined through breeding and selection, leading to cultivars with various fruit weight and size.

The role of auxin in fruit development is achieved through regulations of auxin-responsive genes by gene families such as \textit{ARFs}, \textit{Aux/IAAs}, \textit{GH3}, and \textit{Small Auxin Up RNAs} (\textit{SAURs})\textsuperscript{81-84}. In addition, auxin transport also plays a role. In Arabidopsis, different auxin transport patterns were observed in \textit{trn1} and \textit{trn2} mutants compared with that of the wild type, which could lead to twisted growth by changing auxin distributions\textsuperscript{62,85}. In the current study, the expression of \textit{EjTRN1} seemed decreasing to lower levels significantly earlier in ZP44 than that in ZP65 during fruit development, implying that auxin transport and distribution may be impacted in ZP44 compared with that in ZP65. This may contribute to the shorter duration of auxin supply in ZP44, as supported by the expression patterns of \textit{EjSAMDC1} as well as the results from metabolomics. In addition to auxin, ethylene can also play an important role in loquat fruit development since the candidate gene \textit{EjEIN4} is an ethylene receptor gene. The ethylene and auxin signaling genes may act together in regulating the fruit development. Therefore, the metabolic status of ethylene may be worthy of future research.

**Conclusion**

In this study, combining bi-parental QTL mapping and population genomics analysis of small- and big-fruited accessions, we identified three major loci on Chr8, Chr12, and Chr15 potentially associated with fruit weight, as well as other genomic regions. We have developed three SNP markers associated with fruit weight, which can be further applied in breeding loquat cultivars with higher fruit weight. Our results revealed that the auxin signaling pathway may be under selection over the breeding history of loquat, leading to cultivars with various fruit weight performance. Results from this study provided insights into fruit weight
using a multi-omics approach, with perspectives from genomics, transcriptomics, and metabolomics. Overall, the genomic regions, candidate genes, and other results provide abundant and critical information for understanding the genetic basis of fruit weight in this fruit tree crop, as well as lay a solid foundation for future breeding and manipulation of fruit weight/size in loquat.

Materials and methods

Plant materials and phenotyping

An F1 population was obtained from the crosses between two loquat cultivars, Ninghaibai (female parent) and Dafang (male parent), which were planted at the Yangdu Scientific Research and Innovation Base of Zhejiang Academy of Agricultural Sciences (30.4383°N, 120.4164°E) in 2012. A total of 130 F1 plants that could stably form flowers and bear fruits since 2017 were selected for investigation of fruit weight at mature fruit stage in May 2020. The plants were grown under natural conditions and regular management, without flower or fruit thinning, nor bagging. For each tree, 30 loquat fruits were randomly collected for measurement of fruit weight. The details of phenotyping were described in our previous report59. A total of 14 commonly known small-fruited and big-fruited cultivars with fruit weight information publicly available (Table S3) and two sister lines: ZP44 (extremely small-fruited) and ZP65 (extremely big-fruited), which were derived from a cross between Zaozhong No. 6 × Peluches), were grown at the loquat germplasm resource preservation garden (South China Agricultural University, Guangzhou, China). Fruits of ZP44 and ZP65 were collected at 0, 7, 14, 28, 42, 56, 63, 77, 84, and 91 days past anthesis (DPA) with three biological replicates. Fruits collected at 0, 7, 28, 56, and 77 DPA were subjected to RNA-sequencing using the Illumina NovaSeq 6000 platform (150 bp paired-end reads). The RNA samples were extracted using the EASYspin Plus Plant RNA Extraction Kit (Aidlab, China).

DNA samples were extracted using young leaves and the M5 CTAB Plant gDNA Extraction kit (Mei5 Biotechnology Co., Ltd, Beijing, China). DNA samples of the F1 population and the two parental cultivars were subjected to whole-genome resequencing
using the Illumina NovaSeq 6000 (150 bp paired-end) platform. DNA samples of germplasm accessions were subjected to whole-genome resequencing using the Illumina HiSeq 2000 (125 bp paired-end) platform.

**Mapping and variant calling**

Clean reads generated from whole-genome resequencing were mapped to the ‘Seventh Star’ reference genome\(^6^0\), which was first released to the public at the initiation of this project in 2020, using BWA-mem (BWA 0.7.11)\(^8^6\). Variants were called using GATK 4.2.0\(^8^7\) following the Best Practice pipeline (https://www.broadinstitute.org/gatk/guide/). The identified SNPs were filtered based on the following criteria: 1) Bi-allelic SNPs; 2) Quality > 30; 3) Depth ≥ 4; 4) missing rate < 0.25; 5) minor allele frequency > 0.05.

**Linkage map construction and QTL identification**

For linkage map construction, a Bin strategy (patent application number: 201810689069.8) developed by Beijing Biomarker Company (Beijing, China) was applied with HighMap software\(^6^1\). Briefly, the polymorphic SNPs with a homozygous genotype for Ninghaibai and Dafang were filtered out. The SNPs were first assigned to groups based on their physical locations in the reference genome, and linkage analysis was carried out for each pair of markers in each group. The linkage phases of markers were established based on recombination frequencies and mLOD values, which were further used to establish the linkage phases of parents. The genotypes of markers were corrected using the hidden Markov (HMM) model and Viterbi algorithm. Markers without recombination were partitioned into bins, which were used for linkage map construction using the HighMap software. Only markers that were fully informative with phasing information available for both alleles for each parent were selected to construct the consensus map.

QTL mapping for fruit weight was carried out using the interval mapping model in MapQTL 6.0 software\(^8^8\). To determine the threshold of LOD scores, 1000 permutations were used. The LOD threshold was set at 6.2 corresponding to the 99.5% confidence interval. The collinearity between constructed linkage maps with the physical locations was evaluated as
previously described\textsuperscript{61}.

**Comparative genomic analysis of small-fruited and big-fruited loquat cultivars**

To determine the genetic relationship of 20 commonly known small- or big-fruited loquat cultivars, a phylogenetic tree was constructed using the filtered SNPs and PHYLIP with 100 bootstraps\textsuperscript{89}. Principal component analysis (PCA) was carried out using Plink\textsuperscript{90}. Genetic differentiation ($F_{st}$) values (small-fruited vs big-fruited) were calculated using VCFtools\textsuperscript{91} with the parameters “-fst-window-size 10000 --fst-window-step 5000”.

**RNA-seq analysis**

Library construction and RNA-seq were performed by Wuhan Metware Biotechnology Co., Ltd. (Wuhan, China). HISAT2\textsuperscript{92} was used to map the clean reads generated from RNA-seq to the reference genome of ‘Jiefangzhong’\textsuperscript{52}, which became available to the public when this RNA-seq experiment was initiated in 2021. Novel transcript sequences were assembled using StringTie\textsuperscript{93}. Annotation of novel transcripts was carried out by comparing against databases including KEGG, GO, NR, Swiss-Prot, trEMBL, and KOG. Principal component analysis was performed to determine the correlations among replicates. The differentially expressed genes (DEGs) were identified using read counts required by DESeq2\textsuperscript{94}, and under the threshold of $|\log_2 \text{Fold Change}| \geq 1$ and FDR < 0.05. To identify and compare gene expression patterns across different time points, the genes were clustered using the K-means method.

**Auxin metabolite profiling and analysis**

Fruit tissues collected from ZP44 and ZP65 at above five time points with three biological replicates were used for auxin metabolite profiling. Auxin metabolite extraction and quantification were performed using the AB Sciex QTRAP® 6500 ultra-performance liquid chromatography-tandem mass spectrometry (LC-MS/MS) system following the standard procedures. The same procedure was applied as previously described\textsuperscript{95}, except that 1 mL methanol/water/formic acid (15:4:1, V/V/V) was used. Auxin content was expressed as ng/g fresh weight based on the external standard method, with a series of standard solutions
ranging from 0.01 ng/mL to 500 ng/mL, except for TRP and SAG (0.2-10000 ng/mL).

Candidate gene searching

To make use of the two available reference genomes for candidate gene search, the coordinates of flanking markers/borders for identified QTL regions and top 1% \(F_{st}\) regions were extended from both directions and sequences (500 bp in length) were extracted from the ‘Seventh Star’ genome. The extracted sequences were mapped to the ‘Jiefangzhong’ genome using Blastn to identify the corresponding genomic regions. To identify loquat homologs to known genes potentially associated with fruit weight, keywords “cell division/elongation/enlargement/expansion/size”, “organ growth/size”, and “auxin signaling” were searched in the Swiss-Prot database. Related Arabidopsis genes were used to identify loquat homologs in both reference genomes using an All-Against-All Blast and OrthoMCL approach\(^96\) (inflation value 1.5). The loquat homologs within the identified QTL, top 1% \(F_{st}\) regions, or those that are DEGs were prioritized for candidate gene searching. These candidate genomic regions and genes were plotted on the two reference genomes using TBtools\(^97\).

qRT-PCR assays

Expression levels of the candidate gene were validated using quantitative real-time PCR. Primers were designed using Primer3 (https://primer3.ut.ee/). \(Ej\beta\)-actin was used as the reference gene\(^98\). Primer sequences were provided in Table S14. Each biological replicate had three technical replications. A LightCycler 480 (Roche) and the iTaq™ universal SYBR Green Supermix (Bio-Rad, USA) were used for the gene expression assays.

Marker development and genotyping of Eriobotrya species

For the candidate SNP marker potentially associated with fruit weight, primers flanking the SNP were designed and used for PCR amplification with PrimeSTAR® HS DNA Polymerase (TaKaRa, Japan). Sanger sequencing of PCR amplicons of the extremely small- and big-fruited loquat germplasm accessions and F\(_1\) offspring was performed by Sangon Biotech (Sangon Biotech, China). In addition, the primers were also used to genotype all
known and available *Eriobotrya* species currently maintained in our loquat germplasm preservation garden (South China Agricultural University, Guangzhou, China). The Sanger results were visualized using DNA Baser v4 software (http://www.dnabaser.com/). Primers were provided in Table S14.

**Data availability**
All the raw sequencing data have been deposited at China National GeneBank (CNGB) database under Project accession number CNP0002296, Sample accession number CNS0462964-CNS0463140, and Experiment/Run accession number CNR0460669-CNR0460845.

**Acknowledgements**
This research was funded by the National Key R&D Program of China (2019YFD1000200), the Key-Area Research and Development Program of Guangdong Province (2018B020202011), Research Start-up Funding (Z.P.) from South China Agricultural University, the National Natural Science Foundation of China (31901973), and the Collaborative Innovation Project from the People's Government of Fujian Province & Chinese Academy of Agricultural Sciences (XTCXGC2021006).

**Author contributions**
X.Y., W.S., and Z.P. designed and supervised the project. C.Z., Shuqing L., and Y.G. performed phenotyping of fruit weight. J.C. and H.X. maintained the F1 population and helped on phenotyping. Z.L. and G.H. provided funding support. Shunquan L., X.Y., and X.C. secured loquat germplasm accession materials. Z.P. and C.Z. performed data analysis. Z.P. and C.Z. wrote the original manuscript draft. All authors reviewed and approved the manuscript.

**Conflict of interest**
The authors declare that they have no conflict of interest.
References

1. Mauxion, J. P., Chevalier, C. & Gonzalez, N. Complex cellular and molecular events determining fruit size. *Trends Plant Sci*. 26, 1023-1038 (2021).

2. Li, Y. H., Zhang, Z. & Sun, G. M. Changes in cell number and cell size during pineapple (*Ananas comosus* L.) fruit development and their relationship with fruit size. *Aust. J. Bot.* 58, 673-678 (2010).

3. Rawandoozi, Z. J. *et al.* Identification and characterization of QTLs for fruit quality traits in peach through a multi-family approach. *BMC Genomics* 21, 1-18 (2020).

4. Ozga, J. A. & Reinecke, D. M. Hormonal interactions in fruit development. *J. Plant Growth Regul.* 22, 73-81 (2003).

5. Zheng, W. W., Kim, Y. J., Oh, S. M. & Chun, I. J. Anatomical analysis of fruit development of different-sized ‘Hongro’ and ‘Fuji’ apple fruits. *Hortic. Environ. Biote.* 50, 160-165 (2009).

6. Pattison, R. J. & Catalá, C. Evaluating auxin distribution in tomato (*Solanum lycopersicum*) through an analysis of the PIN and AUX/LAX gene families. *Plant J.* 70, 585-598 (2012).

7. Devoghalaeere, F. *et al.* A genomics approach to understanding the role of auxin in apple (*Malus x domestica*) fruit size control. *BMC Plant Biol.* 12, 1-15 (2012).

8. Bu, H. *et al.* Endogenous auxin content contributes to larger size of apple fruit. *Front. Plant Sci.* 11, 1863 (2020).

9. Agustí, M., Almela, V., Aznar, M., El-Otmani, M. & Pons, J. Satsuma mandarin fruit size increased by 2, 4-DP. *HortScience* 29, 279-281 (1994).

10. Liao, X. *et al.* Interlinked regulatory loops of ABA catabolism and biosynthesis coordinate fruit growth and ripening in woodland strawberry. *Proc. Natl. Acad. Sci. U.S.A.* 115, E11542-E11550 (2018).

11. Wang, H. *et al.* *PbGA20ox2* regulates fruit set and induces parthenocarpy by enhancing GA4 content. *Front. Plant Sci.* 11, 113 (2020).

12. Matsuo, S., Kikuchi, K., Fukuda, M., Honda, I. & Imanishi, S. Roles and regulation of cytokinins in tomato fruit development. *J. Exp. Bot.* 63, 5569-5579 (2012).

13. Xin, T. *et al.* Genetic regulation of ethylene dosage for cucumber fruit elongation. *Plant Cell* 31, 1063-1076 (2019).

14. Nitsch, L. *et al.* ABA-deficiency results in reduced plant and fruit size in tomato. *J.
Plant Physiol. 169, 878-883 (2012).

15 Fu, F. Q. et al. A role of brassinosteroids in early fruit development in cucumber. J. Exp. Bot. 59, 2299-2308 (2008).

16 Perrot-Rechenmann, C. & Napier, R. M. Auxins. Vitam. Horm. 72, 203-233 (2005).

17 Zhao, Y. et al. A role for flavin monooxygenase-like enzymes in auxin biosynthesis. Science 291, 306-309 (2001).

18 Staswick, P. E., Tiryaki, I. & Rowe, M. L. Jasmonate response locus JAR1 and several related Arabidopsis genes encode enzymes of the firefly luciferase superfamily that show activity on jasmonic, salicylic, and indole-3-acetic acids in an assay for adenylation. Plant Cell 14, 1405-1415 (2002).

19 Vanneste, S. & Friml, J. Auxin: a trigger for change in plant development. Cell 136, 1005-1016 (2009).

20 Tiwari, S. B., Hagen, G. & Guilfoyle, T. The roles of auxin response factor domains in auxin-responsive transcription. Plant Cell 15, 533-543 (2003).

21 Lavy, M. & Estelle, M. Mechanisms of auxin signaling. Development 143, 3226-3229 (2016).

22 Tan, X. et al. Mechanism of auxin perception by the TIR1 ubiquitin ligase. Nature 446, 640-645 (2007).

23 Leyser, O. Auxin signaling. Plant Physiol. 176, 465-479 (2018).

24 Giovannoni, J. J. Genetic regulation of fruit development and ripening. Plant Cell 16, S170-S180 (2004).

25 Su, L. et al. The auxin SlIAA17 transcriptional repressor controls fruit size via the regulation of endoreduplication-related cell expansion. Plant Cell Physiol. 55, 1969-1976 (2014).

26 Mounet, F. et al. Down-regulation of a single auxin efflux transport protein in tomato induces precocious fruit development. J. Exp. Bot. 63, 4901-4917 (2012).

27 Guilfoyle, T. J. & Hagen, G. Auxin response factors. Curr. Opin. Plant Biol. 10, 453-460 (2007).

28 Ren, Z. et al. The auxin receptor homologue in Solanum lycopersicum stimulates tomato fruit set and leaf morphogenesis. J. Exp. Bot. 62, 2815-2826 (2011).

29 De Jong, M., Wolters-Arts, M., Feron, R., Mariani, C. & Vriezen, W. H. The Solanum
lycopersicum auxin response factor 7 (SlARF7) regulates auxin signaling during tomato fruit set and development. *Plant J.* **57**, 160-170 (2009).

30 De Jong, M. *et al.* *Solanum lycopersicum* AUXIN RESPONSE FACTOR 9 regulates cell division activity during early tomato fruit development. *J. Exp. Bot.* **66**, 3405-3416 (2015).

31 Nardozza, S. *et al.* Phytohormone and transcriptomic analysis reveals endogenous cytokinins affect kiwifruit growth under restricted carbon supply. *Metabolites* **10**, 23 (2020).

32 Frary, A. *et al.* fw2.2: a quantitative trait locus key to the evolution of tomato fruit size. *Science* **289**, 85-88 (2000).

33 Cong, B., Liu, J. & Tanksley, S. D. Natural alleles at a tomato fruit size quantitative trait locus differ by heterochronic regulatory mutations. *Proc. Natl. Acad. Sci. U.S.A.* **99**, 13606-13611 (2002).

34 Liu, J., Cong, B. & Tanksley, S. D. Generation and analysis of an artificial gene dosage series in tomato to study the mechanisms by which the cloned quantitative trait locus fw2.2 controls fruit size. *Plant Physiol.* **132**, 292-299 (2003).

35 Zygier, S. *et al.* QTLs mapping for fruit size and shape in chromosomes 2 and 4 in pepper and a comparison of the pepper QTL map with that of tomato. *Theor. Appl. Genet.* **111**, 437-445 (2005).

36 Illa-Berenguer, E., Van Houten, J., Huang, Z. & van der Knaap, E. Rapid and reliable identification of tomato fruit weight and locule number loci by QTL-seq. *Theor. Appl. Genet.* **128**, 1329-1342 (2015).

37 Chakrabarti, M. *et al.* A cytochrome P450 regulates a domestication trait in cultivated tomato. *Proc. Natl. Acad. Sci. U.S.A.* **110**, 17125-17130 (2013).

38 Li, Q. *et al.* Differential expression of SlKLUH controlling fruit and seed weight is associated with changes in lipid metabolism and photosynthesis-related genes. *J. Exp. Bot.* **72**, 1225-1244 (2021).

39 Mu, Q. *et al.* Fruit weight is controlled by *Cell Size Regulator* encoding a novel protein that is expressed in maturing tomato fruits. *PLoS Genet.* **13**, e1006930 (2017).

40 Potts, S. M., Khan, M. A., Han, Y., Kushad, M. M. & Korban, S. S. Identification of quantitative trait loci (QTLs) for fruit quality traits in apple. *Plant Mol. Biol. Rep.* **32**, 109-116 (2014).

41 Duan, N. *et al.* Genome re-sequencing reveals the history of apple and supports a two-stage model for fruit enlargement. *Nat. Commun.* **8**, 1-11 (2017).
Wu, J. et al. High-density genetic linkage map construction and identification of fruit-related QTLs in pear using SNP and SSR markers. *J. Exp. Bot.* **65**, 5771-5781 (2014).

Li, X. et al. Development of an integrated 200K SNP genotyping array and application for genetic mapping, genome assembly improvement and genome wide association studies in pear (*Pyrus*). *Plant Biotechnol. J.* **17**, 1582-1594 (2019).

Li, Y. et al. Genomic analyses of an extensive collection of wild and cultivated accessions provide new insights into peach breeding history. *Genome Biol.* **20**, 1-18 (2019).

Shi, P. et al. Construction of a high-density SNP-based genetic map and identification of fruit-related QTLs and candidate genes in peach (*Prunus persica* (L.) Batsch). *BMC Plant Biol.* **20**, 1-16 (2020).

Doligez, A. et al. New stable QTLs for berry weight do not colocalize with QTLs for seed traits in cultivated grapevine (*Vitis vinifera* L.). *BMC Plant Biol.* **13**, 1-16 (2013).

Mengist, M. F. et al. High-density linkage map construction and identification of loci regulating fruit quality traits in blueberry. *Hortic. Res.* **8**, 1-13 (2021).

Jiang, S. et al. Comparative analysis of transcriptomes to identify genes associated with fruit size in the early stage of fruit development in *Pyrus pyrifolia*. *Int. J. Mol. Sci.* **19**, 2342 (2018).

Pei, M. S. et al. Comparative transcriptome analyses of fruit development among pears, peaches, and strawberries provide new insights into single sigmoid patterns. *BMC Plant Biol.* **20**, 1-15 (2020).

Muñoz-Espinoza, C. et al. Transcriptome profiling of grapevine seedless segregants during berry development reveals candidate genes associated with berry weight. *BMC Plant Biol.* **16**, 1-17 (2016).

Sharpe, R. H. Loquat: botany and horticulture. *Horticult. Rev.* **23**, 233 (2010).

Su, W. et al. Polyploidy underlies co-option and diversification of biosynthetic triterpene pathways in the apple tribe. *Proc. Natl. Acad. Sci. U.S.A.* **118**, e2101767118 (2021).

Liu, Y., Zhang, W., Xu, C. & Li, X. Biological activities of extracts from loquat (*Eriobotrya japonica* Lindl.): a review. *Int. J. Mol. Sci.* **17**, 1983 (2016).

Agustí, M., Juan, M., Almela, V. & Gariglio, N. Loquat fruit size is increased through the thinning effect of naphthaleneacetic acid. *Plant Growth Regul.* **31**, 167-171 (2000).
Amorós, A. et al. Role of naphthalene acetic acid and phenothiol treatments on increasing fruit size and advancing fruit maturity in loquat. *Sci. Hortic.* **101**, 387-398 (2004).

Reig, C., Mesejo, C., Martínez-Fuentes, A. & Agustí, M. Synthetic auxin 3, 5, 6-TPA increases fruit size of loquat (*Eriobotrya japonica* Lindl.) by reducing cell turgor pressure. *Sci. Hortic.* **210**, 213-219 (2016).

Su, W. et al. The cellular physiology of loquat (*Eriobotrya japonica* Lindl.) fruit with a focus on how cell division and cell expansion processes contribute to pome morphogenesis. *Sci. Hortic.* **224**, 142-149 (2017).

Su, W. et al. *EjBZR1* represses fruit enlargement by binding to the *EjCYP90* promoter in loquat. *Hortic. Res.* **8**, 1-14 (2021).

Zhao, C. et al. Correlation and genetic analysis of fruit traits in Ninghaibai × Dafang loquat F1 hybrid population. *J. Fruit Sci.* **38**, 11 (2021).

Jiang, S., An, H., Xu, F. & Zhang, X. Chromosome-level genome assembly and annotation of the loquat (*Eriobotrya japonica*) genome. *GigaScience* **9**, giaa015 (2020).

Liu, D. et al. Construction and analysis of high-density linkage map using high-throughput sequencing data. *Plos One* **9**, e98855 (2014).

Cnops, G. et al. The *TORNADO1* and *TORNADO2* genes function in several patterning processes during early leaf development in *Arabidopsis thaliana*. *Plant Cell* **18**, 852-866 (2006).

Li, Z. et al. Three SAUR proteins SAUR76, SAUR77 and SAUR78 promote plant growth in Arabidopsis. *Sci. Rep.* **5**, 1-19 (2015).

Liu, S. et al. Tomato *AUXIN RESPONSE FACTOR 5* regulates fruit set and development via the mediation of auxin and gibberellin signaling. *Sci. Rep.* **8**, 1-16 (2018).

Hao, Y. J. et al. Molecular cloning and functional characterization of two apple *S-adenosylmethionine decarboxylase* genes and their different expression in fruit development, cell growth and stress responses. *Gene* **350**, 41-50 (2005).

Cui, X. et al. The *BUD2* mutation affects plant architecture through altering cytokinin and auxin responses in Arabidopsis. *Cell Res.* **20**, 576-586 (2010).

Kim, J. H., Choi, D. & Kende, H. The AtGRF family of putative transcription factors is involved in leaf and cotyledon growth in Arabidopsis. *Plant J.* **36**, 94-104 (2003).
Beltramino, M. *et al.* Robust increase of leaf size by *Arabidopsis thaliana* GRF3-like transcription factors under different growth conditions. *Sci. Rep.* **8**, 1-13 (2018).

Pattison, R. J., Csukasi, F. & Catalá, C. Mechanisms regulating auxin action during fruit development. *Physiol. Plant.* **151**, 62-72 (2014).

Fenn, M. A. & Giovannoni, J. J. Phytohormones in fruit development and maturation. *Plant J.* **105**, 446-458 (2021).

Fukuda, S., Ishimoto, K., Terakami, S., Yamamoto, T. & Hiehata, N. Genetic mapping of the loquat canker resistance gene pse-c in loquat (*Eriobotrya japonica*). *Sci. Hortic.* **200**, 19-24 (2016).

Fukuda, S. *et al.* Construction of a high-density linkage map for bronze loquat using RAD-Seq. *Sci. Hortic.* **251**, 59-64 (2019).

Hu, Z. *et al.* A re-sequencing-based ultra-dense genetic map reveals a gummy stem blight resistance-associated gene in *Cucumis melo*. *DNA Res.* **25**, 1-10 (2018).

Shi, T. X. *et al.* Mapping QTLs for 1000-grain weight and genes controlling hull type using SNP marker in Tartary buckwheat (*Fagopyrum tataricum*). *BMC Genomics* **22**, 1-12 (2021).

Gisbert, A. D., Martínez-Calvo, J., Llácer, G., Badenes, M. L. & Romero, C. Development of two loquat (*Eriobotrya japonica* (Thunb.) Lindl.) linkage maps based on AFLPs and SSR markers from different *Rosaceae* species. *Mol. Breed.* **23**, 523-538 (2009).

Fukuda, S. *et al.* Genetic mapping of the loquat canker resistance locus in bronze loquat (*Eriobotrya deflexa*). *Tree Genet. Genomes* **10**, 875-883 (2014).

Mejía, N. *et al.* Identification of QTLs for seedlessness, berry size, and ripening date in a seedless x seedless table grape progeny. *Am. J. Enol. Viticult.* **58**, 499-507 (2007).

Costantini, L., Battilana, J., Lamaj, F., Fanizza, G. & Grando, M. S. Berry and phenology-related traits in grapevine (*Vitis vinifera* L.): from quantitative trait loci to underlying genes. *BMC Plant Biol.* **8**, 1-17 (2008).

Müller, A., Düchting, P. & Weiler, E. W. A multiplex GC-MS/MS technique for the sensitive and quantitative single-run analysis of acidic phytohormones and related compounds, and its application to *Arabidopsis thaliana*. *Planta* **216**, 44-56 (2002).

Agustí, M. *et al.* Effect of the synthetic auxin 2, 4-DP on fruit development of loquat. *Plant Growth Regul.* **41**, 129-132 (2003).

Ulmasov, T., Hagen, G. & Guilfoyle, T. J. Dimerization and DNA binding of auxin
response factors. *Plant J.* 19, 309-319 (1999).

82 Audran-Delalande, C. *et al.* Genome-wide identification, functional analysis and expression profiling of the *Aux/IAA* gene family in tomato. *Plant Cell Physiol.* 53, 659-672 (2012).

83 Guilfoyle, T. J., Ulmasov, T. & Hagen, G. The ARF family of transcription factors and their role in plant hormone-responsive transcription. *Cell. Mol. Life Sci.* 54, 619-627 (1998).

84 Liscum, E. & Reed, J. Genetics of *Aux/IAA* and ARF action in plant growth and development. *Plant Mol. Biol.* 49, 387-400 (2002).

85 Ishida, T., Thitamadee, S. & Hashimoto, T. Twisted growth and organization of cortical microtubules. *J. Plant Res.* 120, 61-70 (2007).

86 Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* 25, 1754-1760 (2009).

87 McKenna, A. *et al.* The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* 20, 1297-1303 (2010).

88 Van Ooijen, J. MapQTL® 6, Software for the mapping of quantitative trait loci in experimental populations of diploid species (Kyazma B.V., Wageningen, 2009).

89 Felsenstein J. PHYLIP-phylogeny inference package (version 3.2). *Cladistics* 5, 164-166 (1989).

90 Purcell, S. *et al.* PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.* 81, 559–575 (2007).

91 Danecek, P. *et al.* The variant call format and VCFtools. *Bioinformatics* 27, 2156-2158 (2011).

92 Kim, D., Paggi, J. M., Park, C., Bennett, C. & Salzberg, S. L. Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. *Nat. Biotechnol.* 37, 907-915 (2019).

93 Pertea, M. *et al.* StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. *Nat. Biotechnol.* 33, 290-295 (2015).

94 Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* 15, 550 (2014).

95 Liu, Q. *et al.* *OsLAGT1* is a glucosyltransferase gene involved in the glucose conjugation of auxins in rice. *Rice* 12, 1-13 (2019).
Fischer, S. et al. Using OrthoMCL to assign proteins to OrthoMCL-DB groups or to cluster proteomes into new ortholog groups. *Curr. Protoc. Bioinformatics* 35, 6.12.11-16.12.19 (2011).

Chen, C. et al. TBtools: an integrative toolkit developed for interactive analyses of big biological data. *Mol. Plant* 13, 1194-1202 (2020).

Jiang, Y. et al. EjTFL1 genes promote growth but inhibit flower bud differentiation in loquat. *Front. Plant Sci.* 11, 576 (2020).