

# BULKED SEGREGANT RNA SEQUENCING (BSR-SEQ) IDENTIFIES A NOVEL ALLELE ASSOCIATED WITH WEEPING TRAITS IN *PRUNUS MUME*

Xiaokang ZHUO\*<sup>1,2</sup>, Tangchun ZHENG\*<sup>1,2</sup>, Zhiyong ZHANG<sup>1,2</sup>, Suzhen LI<sup>1,2</sup>, Yichi ZHANG<sup>1,2</sup>, Lidan SUN<sup>1,2</sup>, Weiru YANG<sup>1</sup>, Jia WANG<sup>1</sup>, Tangren CHENG<sup>1</sup>, Qixiang ZHANG (✉)<sup>1,2</sup>

- 1 Beijing Key Laboratory of Ornamental Plants Germplasm Innovation and Molecular Breeding, National Engineering Research Center for Floriculture, Beijing Laboratory of Urban and Rural Ecological Environment, Engineering Research Center of Landscape Environment of Ministry of Education, Key Laboratory of Genetics and Breeding in Forest Trees and Ornamental Plants of Ministry of Education, School of Landscape Architecture, Beijing Forestry University, Beijing 100083, China.
- 2 Beijing Advanced Innovation Center for Tree Breeding by Molecular Design, Beijing Forestry University, Beijing 100083, China.

\*These authors contribute equally to the work

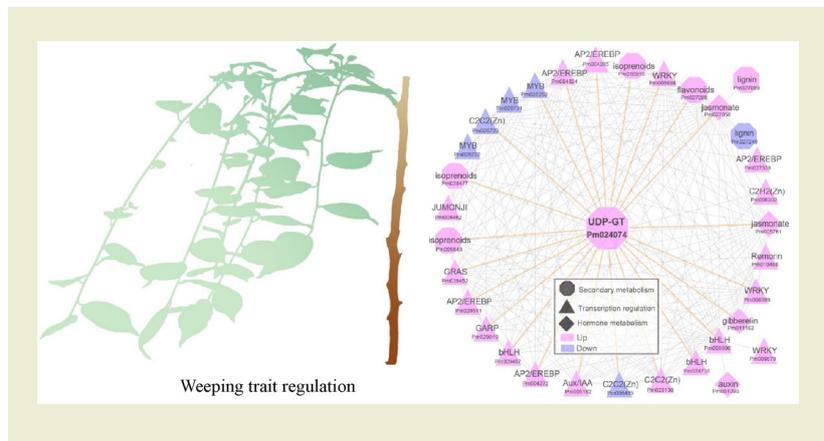
## KEYWORDS

BSR-seq, *PmUGT72B3*, *Prunus mume*, UDP-glycosyltransferase, weeping shoots, WGCNA

## HIGHLIGHTS

- Five QTLs associated with weeping traits on chromosome 7 were identified by BSR-seq.
- The novel allele *PmUGT72B3* has a synonymous transition of T<sup>66</sup> (upright) to C (weeping) in the coding sequence and a 470-bp deletion in the promoter region.
- *PmUGT72B3* was associated with hormone and lignin regulation by WGCNA.

## GRAPHICAL ABSTRACT



## ABSTRACT

Weeping species are used both as ornamental plants and for breeding dwarf plant types. However, exploration of casual genes controlling weeping traits is rather limited. Here, we identified individuals with contrasting phenotypes from an F<sub>1</sub> bi-parental mapping population of *Prunus mume* which was developed from a cross between the upright cultivar ‘Liuban’ and the weeping cultivar ‘Fentai Chuizhi’. Bulked segregant RNA sequencing was used and five QTLs on chromosome 7 were identified. The *Pm024074* (*PmUGT72B3*) allele, belonging to the UDP-glycosyltransferase superfamily containing the coniferyl-alcohol glycosyltransferase domain, was identified in a genomic region overlapping with a previously identified QTL, and had a synonymous transition of T<sup>66</sup> (upright) to C (weeping) in the coding sequence and a 470-bp deletion in the promoter region. *Pm024074* had exceptionally high expression in buds and stems of weeping *P. mume*. Weighted correlation network analysis indicates that genes

Received August 18, 2020;

Accepted November 19, 2020.

Correspondence: zqxbjfu@126.com

neighboring *Pm024074* were significantly associated with plant architecture. In addition, a reliable single nucleotide polymorphism marker was developed based on the variation in the *Pm024074* gene, providing precise marker-assisted breeding for weeping traits. This study provides insights into the genetic mechanism governing the weeping trait in *P. mume*, and indicates potential applications for the manipulation of tree architecture.

© The Author(s) 2020. Published by Higher Education Press. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0>)

## 1 INTRODUCTION

Tree architecture is pivotal to horticultural productivity and management. Trees with appropriate canopy and height more efficiently intercept light, increasing productivity and reducing labor costs. In addition, trees with unique architecture (e.g., weeping growth habit) are generally considered to have value as ornamental trees. Over recent decades, favorable architectural traits have been selected to greatly increase field and horticultural crop yields<sup>[1]</sup>. Plant architecture is not only determined by the organization and activities of meristems, stems, branches, leaves and inflorescence<sup>[2]</sup>, but is also affected by environmental factors<sup>[3,4]</sup>. Previous studies on plant architecture have focused mainly on herbaceous model species such as *Arabidopsis thaliana*<sup>[5]</sup>, rice<sup>[6]</sup> and maize<sup>[7]</sup>. However, genetic information of tree architecture is poorly understood in perennial woody plants, and this is attributed to their complex segregation patterns, heterozygous genome and extended juvenility.

Currently, inexpensive next generation sequencing (NGS) has led to the identification of genetic information associated with the control of tree architecture, and several causative genes or alleles controlling specific aspects of shoot architecture have been identified, mainly dwarf<sup>[8–10]</sup>, columnar<sup>[11–13]</sup>, pillar<sup>[14]</sup> and weeping traits<sup>[15–17]</sup>. Bulk segregant analysis (BSA) is a frequently used method in genetic mapping and candidate identification by analyzing two pools of genomes of contrasting phenotype<sup>[18]</sup>, allowing for inexpensive mapping and causal identification of mutation from forward genetic screens.

RNA sequencing (RNA-seq) is a widely adopted NGS technology which reflects relative transcript concentrations and can also be used to mine for single nucleotide polymorphisms (SNPs) derived from coding sequences. A modification of BSA strategy uses bulk segregant RNA-seq (BSR-seq) reads to rapidly and efficiently map genes controlling mutant phenotypes in segregating populations<sup>[19]</sup>. One of the advantages of BSR-seq is that it can be used for both the identification of QTLs and tests for global patterns of gene expression. In addition, the BSR-seq

strategy has also been extended to mapping genes defined by major QTL loci or dominant mutants<sup>[19]</sup>. Some latest studies using a BSR-seq strategy identified key genes or major QTLs in plant architecture, including the gibberellic acid receptor *GID1c* for dwarf phenotype<sup>[10]</sup> and a conserved sterile alpha motif domain gene for weeping growth habit in peach<sup>[17]</sup>, the major QTL mapping for weeping trait in apple<sup>[16]</sup>, and the genes associated with leafy head formation in Chinese cabbage<sup>[20]</sup>. In addition, similar studies using BSR-seq data have also been adapted for many other traits in plants such as plant growth and development<sup>[21,22]</sup>, disease resistance<sup>[23,24]</sup> and stress responses<sup>[25]</sup>. Although BSR-seq has been used extensively to map important genes in plants, it also has the potential to generate false-positive SNPs or allele-specific expression caused by low-quality reference genomes or uncorrected decisions of tissues<sup>[19,26]</sup>. It would therefore be desirable to combine BSR-seq with causal genes or QTLs using more than two approaches and replicates<sup>[27]</sup>.

The weeping growth habit, a specific phenotype in tree architecture, is common in woody species, representing a particular form of tree architecture and a unique element in landscape architectural aesthetics. Compared with upright trees that grow upwards at certain angles, the shoots of weeping trees initially grow upwards and then their tips gradually grow downward for unknown reasons during shoot growth and development<sup>[28]</sup>. The inheritance and allelism of the weeping trait is highly complex and variable. Previous studies indicate that weeping phenotypes in different species have different segregations including a single dominant locus in apple<sup>[16,29]</sup>, recessive loci in eastern redbud and peach<sup>[30,31]</sup>, and both recessive and dominant loci in chestnut<sup>[32]</sup>. In addition, observations of the weeping phenotype in peach suggest that it is a consequence of incomplete dominance and genetic interactions<sup>[33]</sup>. In *P. mume*, the weeping trait is controlled by a major recessive locus on linkage group 7 coupled with some unknown minor loci<sup>[15]</sup>. In our recent studies, we also revealed genetic interaction involved in the control of weeping traits in *P. mume* and the epistatic loci and minor loci have been

identified. However, the expression patterns and regulated relationships of genes located in these epistatic loci and QTLs remain poorly defined. Therefore, a better understanding of how these genes coordinate the regulation of the weeping trait in *P. mume* would provide important insights into shoot orientation regulation in woody plants with potential horticultural benefits.

Here, we report the use of BSR-seq and comparative transcriptome analysis to reveal molecular pathways involved in weeping trait formation in *P. mume*. Based on a detailed evaluation of phenotypic variants in weeping and upright trees derived from F<sub>1</sub> hybrids of cultivars ‘Liuban’ (upright) × ‘Fentai Chuizhi’ (weeping), we profiled transcript abundance in the buds and stems of plants with contrasting phenotypes. By comparing weeping pools with upright pools, we identified the *Pm024074* allele encoding a glucosyltransferase in an overlapping major QTL that may specifically relate to the control of the weeping trait. In addition, weighted correlation network analysis (WGCNA) shows that many genes in the first and second neighborhood of *Pm024074* were relevant to plant architecture. Our study provides new insights into understanding the genetic basis of weeping trait formation and molecular breeding in *P. mume*.

## 2 MATERIALS AND METHODS

### 2.1 Plant materials and experimental procedures

The F<sub>1</sub> progenies of artificial crossing *P. mume* cultivars ‘Liuban’ (upright phenotype as the female) with ‘Fentai Chuizhi’ (weeping phenotype as the male) were used in this study. The full-sib family of 432 hybrids were grown in 2012 in Deqing County, Zhejiang Province, China (30.53° N, 119.97° E) in a randomized complete block design. After the evaluation of phenotypes in the field, we created bulks of weeping and upright phenotypes originating from the hybrid population. Each bulk contained 20 genotypes that were observed to be contrasting phenotypes for either weeping or upright branching. To avoid other contrasting traits, traits associated with leaf, basal stem, plant height and crown width were further measured and had effectively normal distributions (Fig. S1). The young buds and stems of seedlings of these 20 genotypes were collected for RNA-seq. Finally, we obtained bulks of contrasting traits of bud and stem with three biological replicates per bulk.

The buds and stems of weeping and upright trees derived from two different growth conditions were collected for the validation of gene expression using qRT-PCR. Samples for the tissue-specific expression analysis were collected from grafting

individuals grown in a greenhouse. Different parts of shoots were collected when they sprouted to 5–10 cm in length in spring (March). All samples were stored at –80°C prior to use.

### 2.2 Evaluation of shoot weeping growth habits

Real-time imaging of shoot weeping growth habit in *P. mume* was conducted to monitor the growth dynamics in the studio using time-lapse photography (Fig. S2). The weeping and upright shoots were grafted on the same rootstock to observe the initial status of lateral bud sprouting (Fig. S2). In addition, we also fixed the base of new weeping shoots (<2 cm) to maintain their growth upright, and then observed the growth characteristics before and after removing the shoot apex (Fig. S3). Twenty weeping grafting plants and 20 upright grafting plants were used to observe the growth habit of lateral branches. The weeping phenotype was measured using the nested phenotyping method<sup>[34]</sup>. One-year-old dormant shoots were collected and divided into five parts (T1 to T5) and the T5 angle was defined as the weeping angle and measured using ImageJ software<sup>[35]</sup>. The branch angle at dormancy stage (defined as A2) was also measured as the weeping related traits (Fig. 1(e)). Contrasting individuals were selected through joints of T5 and A2 angles.

### 2.3 RNA extraction, quantification, library construction and sequencing

Equal amounts of total RNA (2 µg each) from buds and stems of 20 strongly weeping and upright individuals were pooled to generate four segregant bulks. Tissue collection for RNA-seq is shown in Fig. 2(a). The RNA was extracted using an EasySpin RNA Reagent (Aidlab Biotechnologies, Beijing, China) according to the manufacturer’s instruction. The RNA concentration, degradation, and integrity were assessed using a Nanodrop spectrophotometer, 1% agarose gel and an Agilent 2100 system, respectively. A total of 3 µg of RNA per bulk was used to construct the cDNA library. This library was generated using a NEBNext Ultra RNA Library Prep Kit for Illumina (NEB, Ipswich, MA) following the manufacturer’s instructions. The library quality was checked using the Agilent Bioanalyzer 2100 system. The amplified fragments were sequenced using an Illumina HiSeq X-10 platform and 150 bp paired-end reads were generated. The insertion size was 270 bp. All raw data were stored in the Genome Sequence Archive under project ID CRA001273.

### 2.4 Bulked segregant RNA-sequence data processing and analysis

The raw reads of four bulked samples were filtered by removing

adapter, poly-N sequences and low-quality sequences using Trimmomatic v.0.39 software<sup>[36]</sup>. Then high-quality clean reads of four bulked samples and whole genome resequencing parents were mapped to the reference genome v.1.0 of *P. mume* using BWA under the default parameters<sup>[37]</sup>. SNPs and InDels were identified using GATK 3.8 software<sup>[38]</sup>. VCFtools v.0.1.15<sup>[39]</sup> was used to filter low-quality variations with minor allele frequency > 0.2, read depth of > 6, and Phred-quality score > 20. The difference in allele frequency between weeping and upright traits was calculated using G' method via QTLseqr package<sup>[40,41]</sup>. Manhattan plots were used to show the results using the 'qqman' R package<sup>[42]</sup>. Raw counts of each read were normalized to fragments per kilobase of transcript per million mapped read (FPKM) values to evaluate the expression levels based on the alignments of reads.

## 2.5 Identification of differentially expressed genes and annotation analysis

Differentially expressed genes (DEGs) of the four bulked samples were analyzed using the DESeq R package<sup>[43]</sup>. The Benjamini-Hochberg approach was then used to correct *P*-values to control a false discovery rate < 0.001. The DEGs were identified with an adjusted *P*-value < 0.05 and an absolute fold-change  $\geq 1.5$  ( $|\log_2\text{Ratio}| \geq 0.585$ ). Functional annotation of all DEGs was conducted using the gene ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) databases<sup>[44,45]</sup>. MapMan software and the Mecnator web tool<sup>[46]</sup> were used to identify DEG categories using TAIR version 10 as reference, with *P* < 0.05 for ontology. Diagrams of PCA analysis, Venn and hierarchical clustering heat maps were generated using R v3.5.3 software.

## 2.6 Co-expression network analysis using differentially expressed genes

To identify genes significantly correlated with the traits (weeping and upright) or tissues (bud and stem) of *P. mume*, the normalized read counts from 1959 DEGs with 2-fold cutoff were used to conduct co-expression gene networks with the WGCNA package in R v3.5.3 software<sup>[47]</sup>. The DEGs between upright and weeping samples in bud and stem tissues were initially grouped together. Redundant genes and genes with FPKM < 1 were filtered to generate a set of confident DEGs to carry out co-expression analysis. The unsigned weighted network was initially created with the parameters following threshold power of 9, minimum module size of 30, and a branch merge cut height of 0.05. The co-expression gene modules were identified by dynamic hybrid tree cut algorithm and visualized using Cytoscape v3.6.1<sup>[48]</sup>. Genes with similar expression patterns

across phenotypes, tissues or their combinations were classified into a module.

## 2.7 Comparative analysis of nucleotide polymorphisms between weeping and upright landraces

The resequencing data of 217 landraces (27 weeping accessions and 190 upright accessions) were used to calculate the nucleotide polymorphisms between weeping and upright landraces. The accession number and Bio-Project number are RP093801 and PRJNA352648, respectively. The methods of data preprocessing, SNP calling and quality control were described by Zhang et al.<sup>[49]</sup>. Using VCFtools v.0.1.15<sup>[39]</sup> we calculated the genome-wide nucleotide diversity, allele frequency, and genetic differentiation. SNPs were validated using the MassARRAY system<sup>[50]</sup>. The genotyping primers were: forward primer ACGTTGGATGGGTGTGTTTCTTTCTAACGAG, reverse primer: ACGTTGGATGATTGACCACAATCCCAGCAG and single nucleotide extension primer: CGGATGTGCAAGAAGTA. The samples and phenotypes are shown in [Data S1](#). Transcription factor binding sites and miRNA targeted sites for variation sequences were predicted using the JASPAR database and miRBase, respectively.

## 2.8 Candidate gene expression validated by qRT-PCR

The buds and stems from weeping and upright samples were collected to validate the gene expression of putative candidates. Each sample was analyzed in triplicate for all three biological repeats. Total RNA was extracted using the same method described for RNA-seq. First-strand cDNA templates were synthesized from ~ 2 µg of total RNA using a PrimeScript™ RT Reagent Kit (TaKaRa, Shiga, Japan) according to the manufacturer's instructions. A set of primers were designed for qRT-PCR using Integrated DNA Technologies online tool and checked their specificity using the Primer-BLAST tool on NCBI. The *protein phosphatase 2A* (*Pm006362*) gene was used as an internal reference, in line with previous reports<sup>[51]</sup>. One-step qPCR programs were conducted using a CFX96 Real-Time PCR Detection System (Bio-Rad, Hercules, CA) with SYBR Premix ExTaqII (TaKaRa) and the qRT-PCR reaction system as described by Xu et al.<sup>[51]</sup>. Expression levels were calculated using the delta-delta CT method<sup>[52]</sup>.

## 2.9 Cloning of the *Pm024074* gene

Specific primers of the *Pm024074* gene were designed based on

the CDS sequences from the *P. mume* genome. Total RNA was extracted from buds of three heterozygous weeping accessions (cultivars ‘Moshan Chuizhi’, ‘Hanxue Chuizhi’ and ‘Fenpi Chuizhi’) and upright accessions (cultivars ‘Caishan Lve’, ‘Danban Lve’ and ‘Changrui Lve’). After mRNA was isolated and transcribed to cDNA, the gene of *Pm024074* was amplified with the specific primers using methods in our laboratory<sup>[53]</sup>. The specific primers were: forward primer ATGGACCCGGTT-CAAGACCG and reverse primer TTAGATCTTGAGGCCCT-TCCATACC. All target sequences were confirmed by Sanger sequencing.

### 3 RESULTS

#### 3.1 Assessment of the weeping growth habit

When dormant lateral buds sprouted the shoots of weeping trees immediately exhibited a horizontal growth tendency but the upright trees were vertical (Fig. 1(a)). These observations indicate that the weeping orientation is generated at an early stage of bud development. We also found that the shoots of weeping trees showed upright growth when the shoot base was fixed. After removal of the shoot apex the new lateral branches regenerated with the weeping growth habit and the original branch angle increased in size (Fig. 1(b)). We therefore suggest that mechanical stress from the increased weight of the shoots might further accelerated the branch and weeping angles. We then further investigated the effects of gravitropism using time-lapse photography to capture the grafted bud movement over a 12-d period until the shoots showed significant downward growth (Fig. 1(d)). We found that the shoots of weeping trees initially grew upwards but the branch angle remained larger than in upright trees. As the shoots grew the branch angles of weeping trees gradually increased and the shoot tips began growing downwards. However, the branch angles were minor in upright trees (Fig. 1(c,d)). This phenomenon supports the notion that gravistimulation or a lack of mechanical rigidity of elongating stems might be an important factor accelerating weeping by accelerating branches downward growth.

In addition, we further evaluated the relationship between branch angles and weeping angles through  $F_1$  segregating hybrids of ‘Liuban’ (upright phenotype)  $\times$  ‘Fentai Chuizhi’ (weeping phenotype). Linear regression analysis shows that the branch angle (A2) was strongly correlated with the weeping angle (T5) ( $R^2 = 0.628$ ) (Fig. 1(e)). This indicates that A2 and T5 angles can be used to discriminate phenotypes between weeping and upright traits. Accordingly we identified 67 strongly weeping individuals with  $A2 \geq 70^\circ$  and  $T5 \geq 70^\circ$ , and 72 strongly upright individuals with  $A2 \leq 50^\circ$  and  $T5 \leq 50^\circ$  (Table S1).

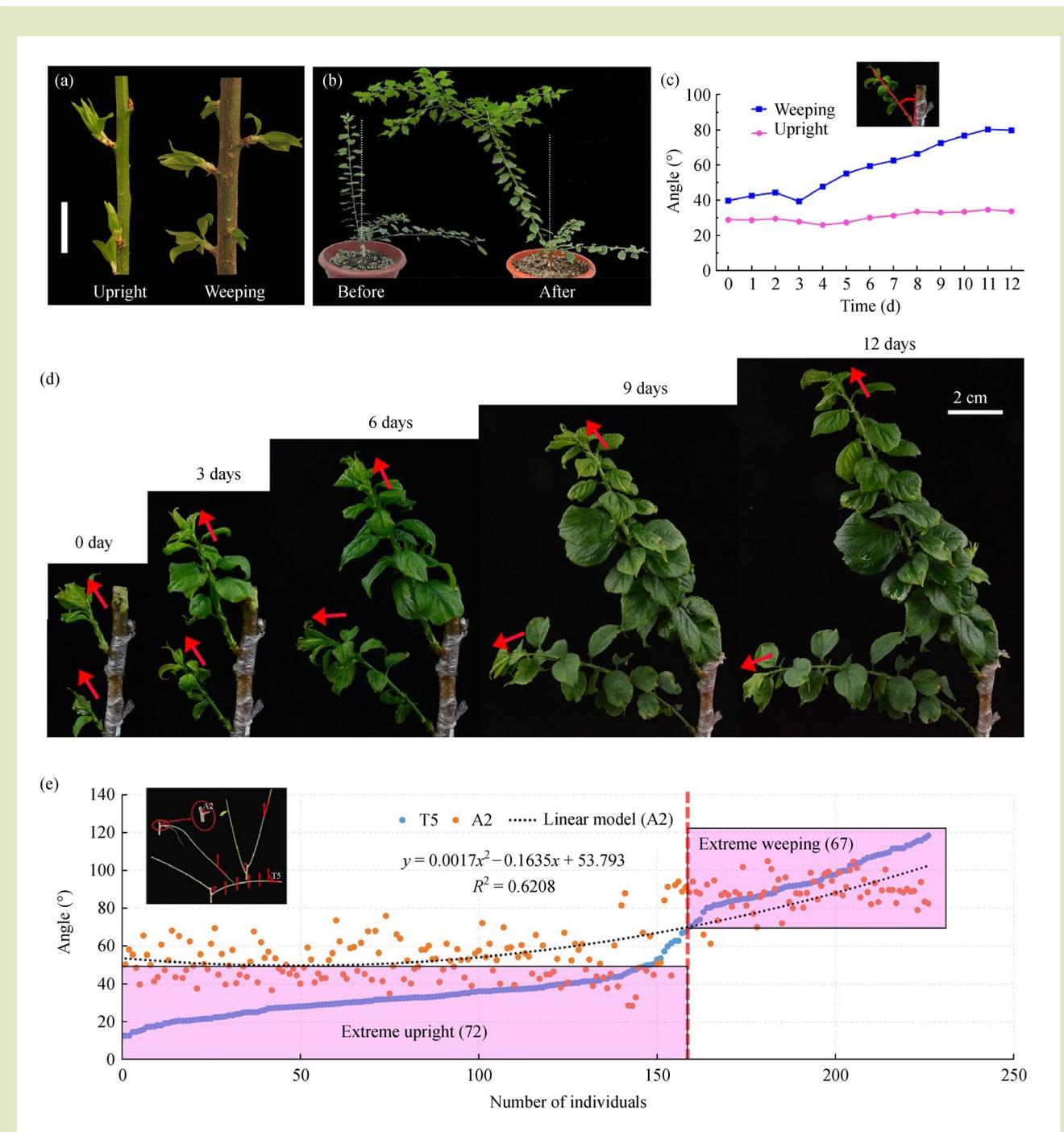
#### 3.2 RNA sequence data processing and global analysis

An RNA-seq method using a BSA strategy was used to identify functionally causative alleles and corresponding gene expression dynamics responsible for weeping traits<sup>[19]</sup>. RNA-seq data were generated from contrasting trait bulks of young buds and stems (Fig. 2(a)). A total of 174 Gb clean data were generated with high quality (Q30 > 91.6%) from four resulting bulks and slightly higher percentages (75%–79%) of the reads were mapped to the reference genome (Tables S2–S3). Normalized read counts (fragments per kb of transcript per million mapped reads, FPKM) for each gene were calculated, and 17,456 genes were identified to be expressed (FPKM > 1.0) in at least one bulk (Data S2). In total, 24,201 genes were successfully annotated using at least one of six databases (COG, GO, KEGG, Swiss-Prot, eggNOG and Nr) and the annotated results of each database are summarized in Fig. 2(d) and Table S4. The detailed annotation of all genes is listed in Data S3.

The gene number for four levels of FPKM values in each bulk was calculated. We found the four bulks had similar numbers of expressed genes (around 16,000) and the largest number of expressed genes was at  $10 < \text{FPKM} < 50$  level (Fig. 2(c)). DEGs with 2-fold cutoff were used for each pairwise comparison. The repeatability and relatedness of RNA-seq data of different samples was evaluated by a cluster dendrogram using 2600 DEGs. The three biological replicates of all bulks showed good correlation (Fig. 2(b)). Based on these analyses a total of 495 DEGs in UB-WB pairwise, 328 DEGs in US-WS pairwise, 1533 DEGs in UB-US pairwise and 1507 in WB-WS pairwise were identified (Fig. 2(e) and Data S4). Most DEGs were identified from UB versus US and WB versus WS compared groups, indicating a large number of DEGs caused by tissue specificity. Most DEGs were upregulated in weeping pools for comparison of upright pools versus weeping pools in the same tissues (Fig. 2(e)).

#### 3.3 Identification of specific DEGs and enrichment analysis

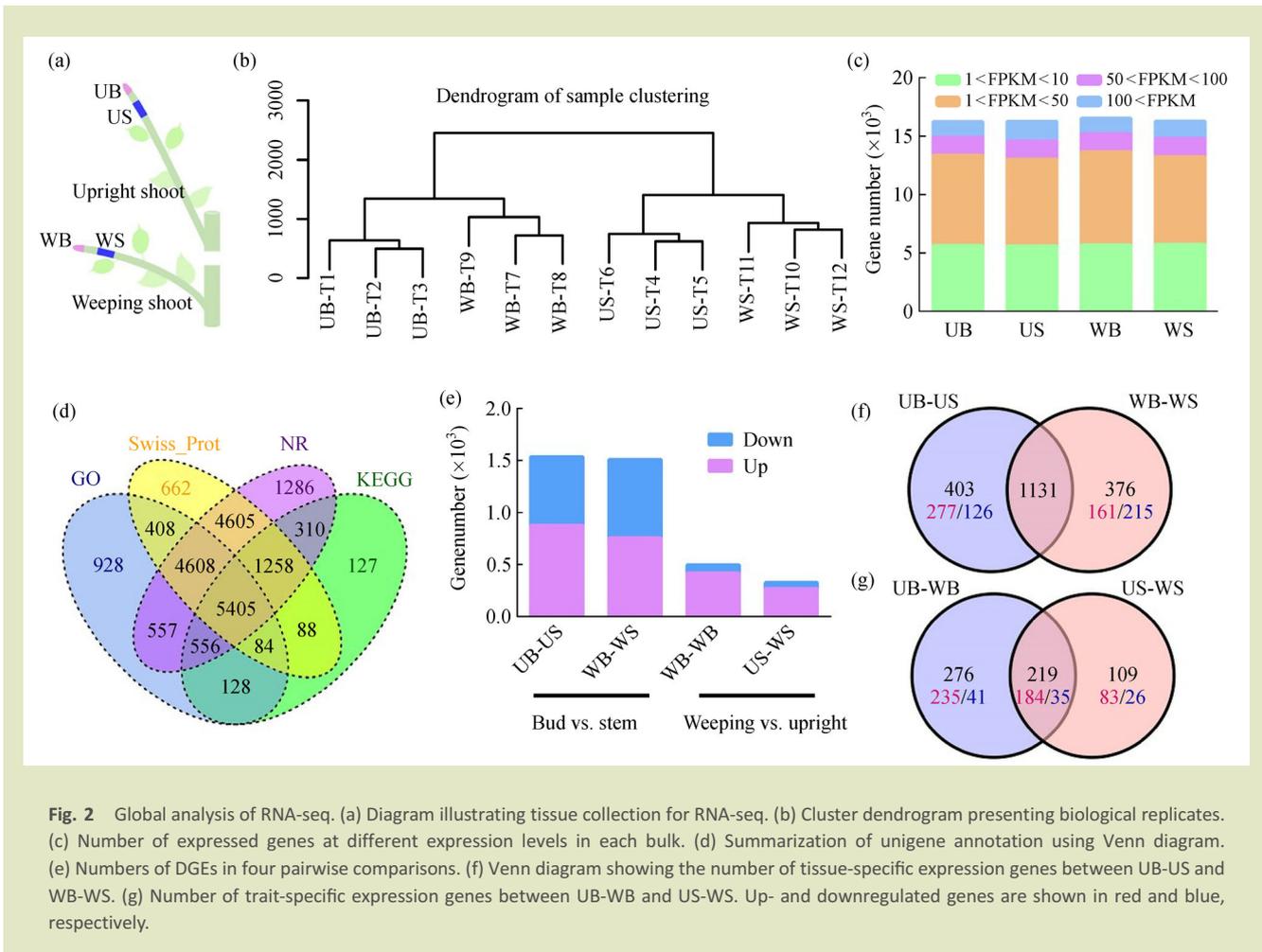
Venn analysis shows that 403 (upregulation/downregulation, 277/126) and 376 (161/215) specific DEGs from bud vs. stem of the same phenotype (UB-US vs. WB-WS) were identified in upright bulks and weeping bulks, respectively, whereas 1131 DEGs overlapped between the two pairwise comparisons (Fig. 2(f)). Comparing DEGs from weeping vs. upright of the same tissues (UB-WB vs. US-WS), 276 (235/41) and 109 (83/26), DEGs were specifically expressed in bud and stem pools, respectively, and 219 (184/35) overlapping DEGs were identified in both buds and stems (Fig. 2(g)). These results indicate that



**Fig. 1** Observation and assessment of weeping phenotype. (a) Growth orientation of the lateral bud between weeping and upright *P. mume* (scale bar: 2 cm). (b) Growth habit of weeping shoot before and after removing apex. (c) Changes in the angle between shoot growth direction and negative gravity direction during plant growth. (d) Phenotype of weeping and upright grafted shoots on different growing days, red arrow indicates toward the direction of the apex. (e) Relationship between branch angle (A2) and weeping angle (T5).

gene expression levels were overall higher in weeping than in upright trees, but more genes downregulated in the stems of weeping than of upright trees.

MapMan functional categories show that specific DEGs of the four comparison groups (UB-US, WB-WS, UB-WB and US-WS) were mainly enriched for secondary metabolism, hormone



**Fig. 2** Global analysis of RNA-seq. (a) Diagram illustrating tissue collection for RNA-seq. (b) Cluster dendrogram presenting biological replicates. (c) Number of expressed genes at different expression levels in each bulk. (d) Summarization of unigene annotation using Venn diagram. (e) Numbers of DGEs in four pairwise comparisons. (f) Venn diagram showing the number of tissue-specific expression genes between UB-US and WB-WS. (g) Number of trait-specific expression genes between UB-WB and US-WS. Up- and downregulated genes are shown in red and blue, respectively.

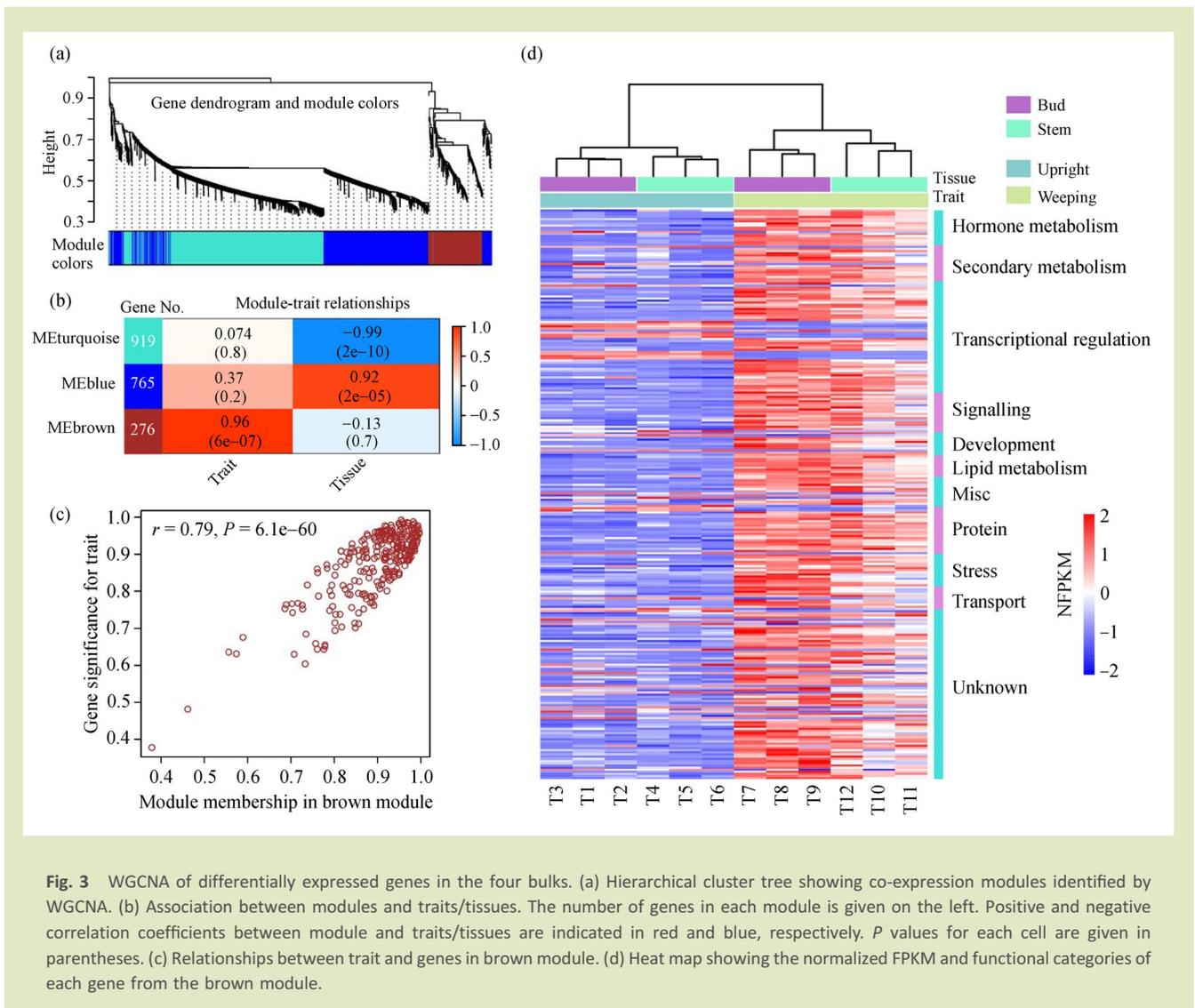
metabolism, stress, RNA, protein, signaling and transport (Fig. S4(a) and Table S5). Most downregulated specific DEGs in the stems of weeping trees were mainly associated with secondary metabolism, hormone metabolism, RNA and signaling (Fig. S4(b) and Data S4), including *ELI3-1* related to lignin biosynthesis, MYB-related transcription factor involved in circadian rhythm regulation, *SAUR-like* auxin-responsive protein, and NPH3 family protein involved in phototropic-response downregulated more than 2-fold in the stems of weeping *P. mume*.

GO enrichment analysis shows that most specifically expressed genes in biological processes were downregulated in the stems of weeping trees (Fig. S5(a,b)). In addition, we found that four processes were specifically enriched in weeping trees, namely immune system processes, growth, and rhythmic processes and supramolecular fibers (Fig. S5(c) and Table S6). Genes among these processes might be related to the weeping trait, such as *bZIP* (*Pm004596*) involved in light-regulated transcriptional activation and *mannitol dehydrogenase* (*Pm002264*) involved in

lignin biosynthesis.

### 3.4 Construction of the gene co-expression networks using weighted correlation network analysis

DEGs (1959) with 2-fold cutoff and median absolute deviation > 0.01 were used for WGCNA to investigate the network relationships between DEGs. Three distinct co-expression modules were identified as shown by the dendrogram (Fig. 3(a)). Genes within the same module had high correlation coefficients and similar expression trends across samples, and significantly correlated relationships between the traits (weeping and upright) or tissues (buds and stems) of *P. mume* and the modules were identified (Fig. 3(b)). The brown module was significantly related to trait with highly positive coefficients (0.96). Genes in the turquoise and blue models showed significant tissue-specificity with high coefficients of -0.99 and 0.92, respectively. We further investigated the relationship between genes in modules and traits or tissues, indicating that genes in the module were strongly associated



**Fig. 3** WGCNA of differentially expressed genes in the four bulks. (a) Hierarchical cluster tree showing co-expression modules identified by WGCNA. (b) Association between modules and traits/tissues. The number of genes in each module is given on the left. Positive and negative correlation coefficients between module and traits/tissues are indicated in red and blue, respectively.  $P$  values for each cell are given in parentheses. (c) Relationships between trait and genes in brown module. (d) Heat map showing the normalized FPKM and functional categories of each gene from the brown module.

with corresponding traits or tissues (Fig. 3(c), Fig. S6). Genes in each significant module can therefore be considered representative of distinct trait or tissue types. Expression patterns of genes in each module are shown in Fig. S7.

In the brown module, most genes were highly and specifically expressed in both buds and stems of weeping tree but slightly expressed in upright trees (Fig. 3(d), Fig. S8). The functional classification shows that most genes were associated with hormone metabolism, transcription regulation, signaling, and proteins (Fig. 3(d) and Data S5). Genes associated with transcriptional regulation, hormone metabolism, and secondary metabolism have important involvement in the regulation of shoot architecture, especially the genes regulating gibberellin (GA), auxin, brassinosteroid (BR), cytokinin (CK) and lignin biosynthesis pathways<sup>[1,54,55]</sup>.

Networks of genes associated with transcriptional regulation, hormone metabolism, and secondary metabolism in brown modules with edge weight > 0.45 were constructed to investigate the relationships of hub genes (Fig. S9(a) and Data S6). Ten genes associated with IAA, BR, GA, CK and lignin were of particular interest in the network (Fig. S9(b) and Table S7). Strikingly, seven of ten genes are involved in the regulation of auxin and lignin, including four auxin related genes (*ILR-like5*, *Prunusmume\_newGene\_2032*; *Cytochrome b561*, *Pm012502*; *GH3*, *Pm021243*; and *Aux/IAA*, *Pm005182*) and three lignin related genes (*AMP-dependent synthetase*, *Pm012986*; *Cytochrome P450*, *Pm027089*; and *Caffeoyl-CoA O-methyltransferase (CCoAMT)*, *Pm027248*). The other three genes are related to CK (*Zeatin O-xylosyltransferase*, *Pm023083*), GA (*Gibberellin 2-oxidase*, *Pm011163*), and BR (*Squalene monooxygenase*, *Pm021002*). All these genes are upregulated in weeping

*P. mume* except lignin related CCoAMT (Table S7). AUX/IAA and CCoAMT are two key genes that are involved in IAA regulation<sup>[56,57]</sup> and biosynthesis of lignin<sup>[58]</sup>, respectively. Some genes directly connected with AUX/IAA and CCoAMT in the networks have been shown to be associated with shoot regulation. For instance, GRAS (*SCARECROW-like 21*, *Pm029452*) transcription factor involved in the BR signal transduction pathway, highly expressed *OsGRAS19* can reduce shoot strength<sup>[59,60]</sup>.

### 3.5 Mapping an overlapping QTL on chromosome 7 via bulked segregant RNA-sequencing

We merged RNA-seq clean reads of WB/WS bulks and UB/US bulks into a weeping bulk and an upright bulk, respectively to identify the gene variation in weeping *P. mume*. A total of 3,874,124 SNPs between weeping and upright *P. mume* (including two bulks and parents) were identified by mapping the clean reads to the reference and quality control of SNPs. All SNPs were divided into continuous 50-kb windows and the G' value of each window was calculated and is presented (Fig. 4(a)). Five putative QTLs on chromosome 7 (Pm7) were identified with threshold Q value < 0.001 (Fig. 4(a) and Table 1). Pm7-qt1-4 located at 9,693,685–10,654,831 bp (~ 0.96 Mb) region partially overlapped with our previous hybrid population and landrace population studies<sup>[15,49]</sup>. This QTL is of particular interest. Nucleotide diversity analysis of this QTL region further demonstrates a distinct selective sweep located in the 10.1–10.7 Mb region (Fig. 4(b)). These results suggest that Pm7-qt1-4 is a higher confidence QTL associated with the weeping trait in *P. mume*.

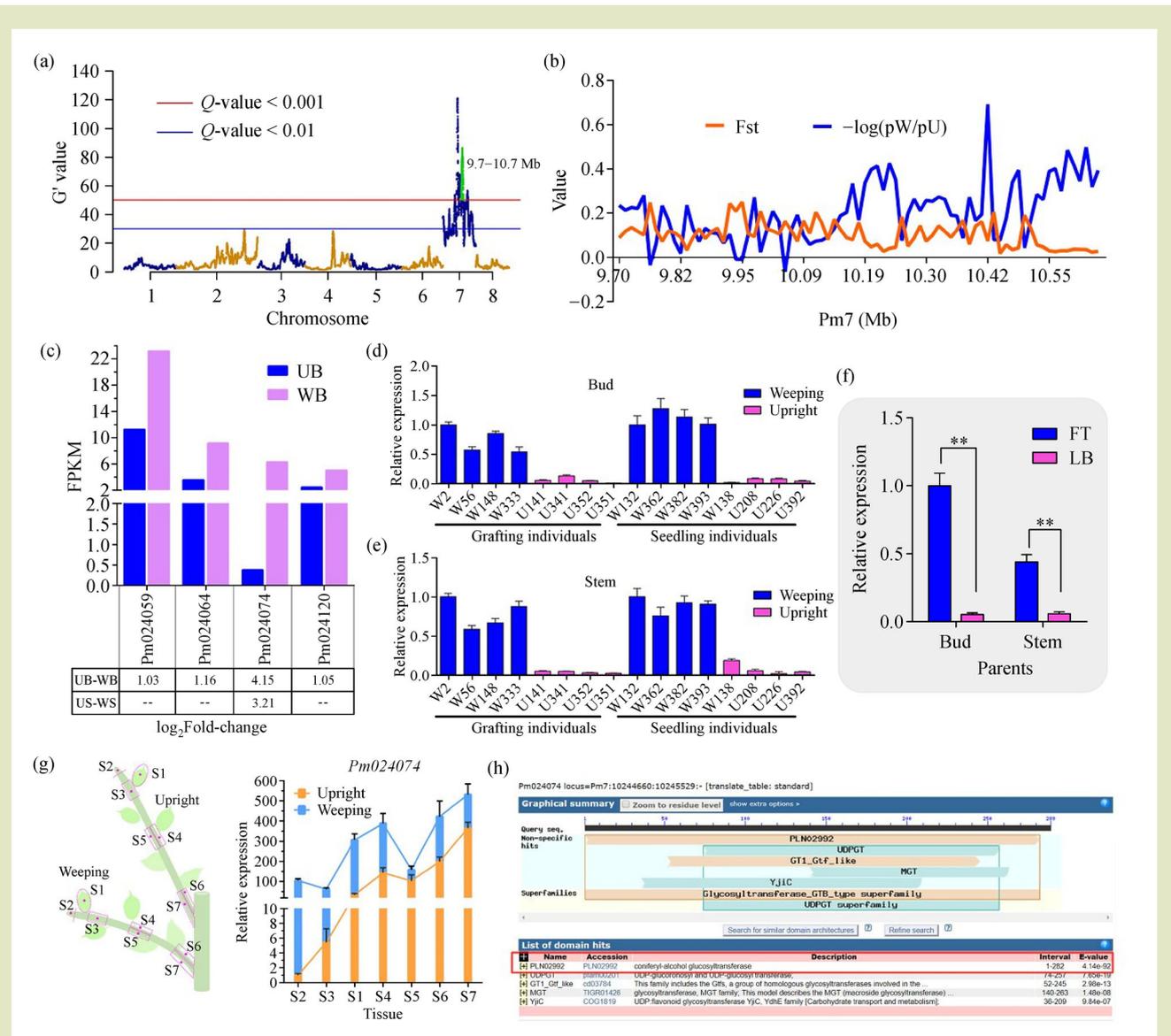
As plant phenotype can be affected by gene expression levels, we further investigated the expression level of genes located in Pm7-qt1-4 QTL interval using RNA-seq data. There were 52 expressed genes in the region but only four DEGs ( $\log_2FC \geq 1$ ) were identified (Fig. 4(c) and Data S7). Functional annotation and fold-change of gene expression indicate that *Pm024074*, belonging to the UDP-glycosyltransferase superfamily, was the most likely associated with weeping trait in *P. mume*. This gene showed greatly different expression between weeping and upright *P. mume* with very large fold-change ( $\log_2FC > 3$ ) (Fig. 4(c)). A search for homologous protein revealed a significant domain: coniferyl-alcohol glucosyltransferase (Fig. 4(h)). Protein contained in this domain was shown to be involved in the glucosylation of lignin precursors and cambial growth<sup>[61,62]</sup>. A recent study indicates that the UDP-glycosyltransferase (UGT) superfamily may also be involved in metabolic redirection and biosynthesis of flavonoids and monolignols<sup>[63]</sup>. The qRT-PCR results further indicate that the

expression level of *Pm024074* was significantly higher in the weeping parents and individuals than that in the upright parents and individuals, and this difference in expression was not impacted by growth environments (Fig. 4(d–f)). Tissue specific expression pattern analysis shows that the overall expression level of *Pm024074* was increasing from shoot tip to basal stem in both weeping and upright *P. mume* and was higher in each part of weeping than in upright *P. mume* (Fig. 4(g)). These results hint that *Pm024074* is potentially associated with weeping traits in *P. mume*.

### 3.6 Discovery of genomic differences between weeping and upright *P. mume*

The weeping growth habit of *P. mume* has important value for landscape architectural aesthetics and most weeping landraces are artificial selections with this aesthetic appeal<sup>[64]</sup>. We wanted to determine if candidate *Pm024074* was a key gene for artificial selection. To this end, SNPs of *Pm024074* and its flanking 2-kb were extracted from 27 weeping and 190 upright resequencing accessions. The nucleotide diversity ( $\pi$ ) and genetic differentiation (Fst) of 100-bp non-overlapping windows were then calculated. The calculation shows that  $\pi$  and Fst between weeping and upright accessions were clearly different at 10,245,401–10,247,301 region (Fig. 5(a)). We further compared the difference in variation sites in *Pm024074* protein coding region between two bulks and the parents. Pm\_10245464 SNP (defined *Pm<sup>W</sup>* allele) was exclusively identified in both weeping bulk and weeping parents (Fig. 5(b)). We successfully genotyped this allele in 235 individuals (Fig. 5(c) and Data S8). Allele frequency shows that the *Pm<sup>W</sup>* allele was strongly associated with weeping traits (Fig. 5(d)). Individuals with CC genotype showed significantly higher branch angle (A2) and weeping angle (T5) compared to individuals with TT genotype (Fig. 5(e)).

The *Pm<sup>W</sup>* allele was a synonymous mutation SNP located at position 66 bp in the cDNA of the *Pm024074* gene, a transition of T<sup>66</sup> in upright *P. mume* to C in weeping *P. mume* (Fig. 6(b)). This variation cannot cause an amino acid change but gene expression levels could also influence plant phenotype<sup>[65]</sup>. We also sought to determine if variation in the promoter region might impact gene expression. We therefore compared the promoter region sequences of the parents, 'Fentai Chuizhi' (weeping) and 'Liuban' (upright) (Fig. 6(a)). A 470-bp deletion was identified in upright *P. mume*, which likely contained four transcription factor binding sites (WRKY18, ARF16, Adof1 and ARR18) (Fig. 6(b)). This indicates that a difference in expression level might be caused by the variation in the promoter region. Notably, we found that *Pm<sup>W</sup>* was a T/C heterozygous locus in most weeping landraces based on allele frequency (Fig. 5(d)).



**Fig. 4** Identification of QTLs and candidate genes. (a) Significant QTL signaling for weeping traits identified by QTLseqr. Plot showing the  $G'$  value with a 50-kb sliding window. QTL overlap with previous study indicated in green. (b) Comparative analysis of SNP polymorphisms between weeping and upright landraces using resequencing data (27 weeping accessions and 190 upright accessions). Values of  $\pi$  and  $F_{st}$  were calculated using SNPs in the overlapped QTL interval, identifying a potential selective sweep region located at 10.1–10.7 Mb on chromosome 7; and (c) Only four differentially expressed genes were identified between weeping and upright *P. mume* in this region. (d–f) Relative expression of *Pm024074* in buds and stems of weeping and upright *P. mume* derived from grafted replicates and seedling individuals and parents. (g) Tissue-specific expression of *Pm024074* in different parts of shoot derived from weeping and upright *P. mume*. (h) Conserved domain of *Pm024074*.

Full-length cDNA of *Pm024074* was cloned and sequenced from three heterozygous weeping and upright landraces to investigate whether allele C was specifically expressed in weeping *P. mume*. The result shows that only allele C was cloned in weeping *P. mume* (Fig. 6(c)). Notably, a putative miR838 was predicted to target this site (Fig. 6(b)). Overall, the above results indicate that variation in the *Pm024074* gene might change its spatial and temporal expression patterns resulting in abnormal develop-

ment of shoots in *P. mume*. Details of the sequence variation information are shown in [Supplementary T1](#).

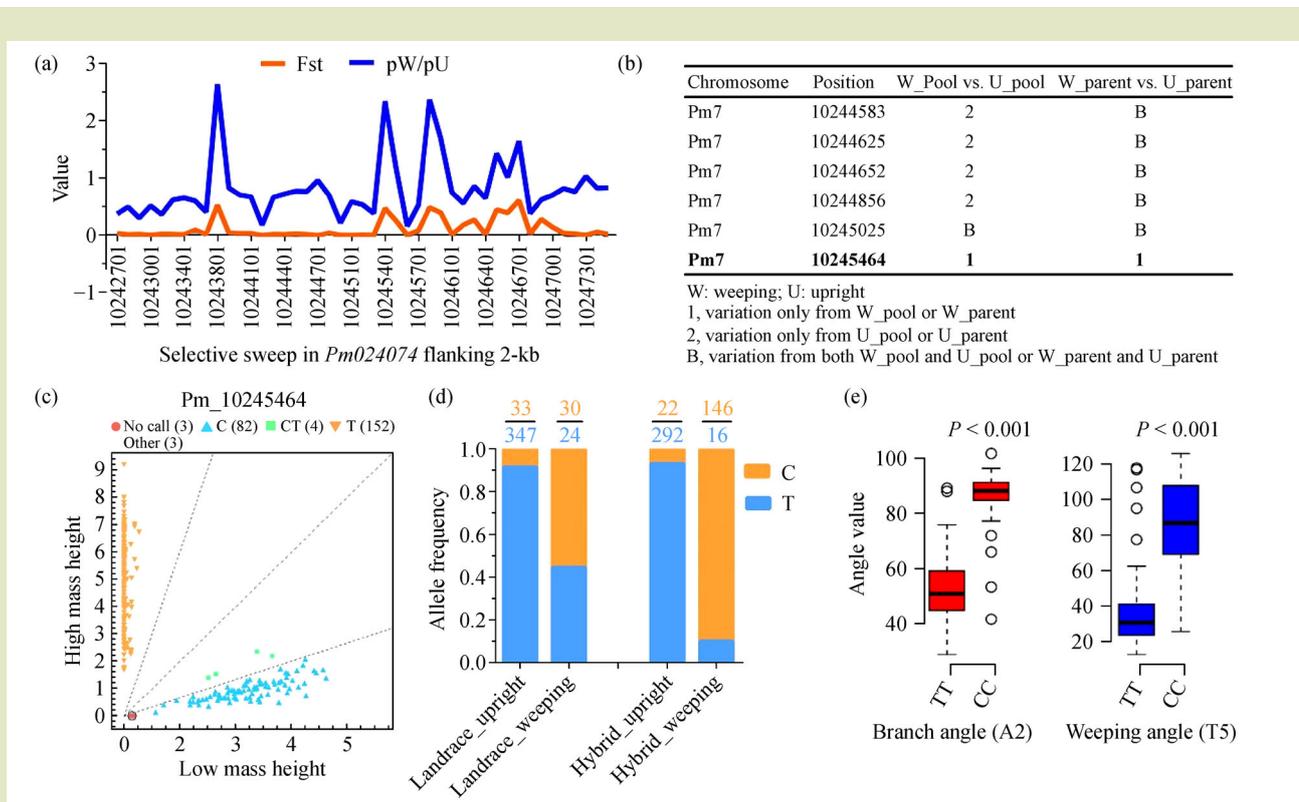
### 3.7 *PmUGT72B3* co-expressed correlation networks

The candidate *Pm02474* gene belongs to the UDP-glycosyltransferase superfamily, best matched *Arabidopsis thaliana*

**Table 1** Five significant QTLs identified by G' method

Chr	QTL	Start	End	Length	nSNPs	Peak Gprime	Peak position	Mean Gprime	P-value	Q-value
Pm7	Pm7-qt1-1	6,303,589	6,546,756	243,167	184	53.888	6,398,896	51.853	9.81E-06	0.000833
Pm7	Pm7-qt1-2	7,292,731	8,162,487	869,756	1050	120.874	7,727,206	89.633	1.78E-06	0.000222
Pm7	Pm7-qt1-3	8,417,790	9,439,749	1,021,959	2229	67.805	8,791,752	62.208	3.13E-06	0.000402
Pm7	Pm7-qt1-4	9,693,685	10,654,831	961,146	2242	86.282	10,122,525	71.930	1.98E-06	0.000279
Pm7	Pm7-qt1-5	12,647,413	13,068,593	421,180	1438	56.611	12,785,876	53.866	7.52E-06	0.000724

Note: Chr, chromosome; nSNP, number of SNPs.

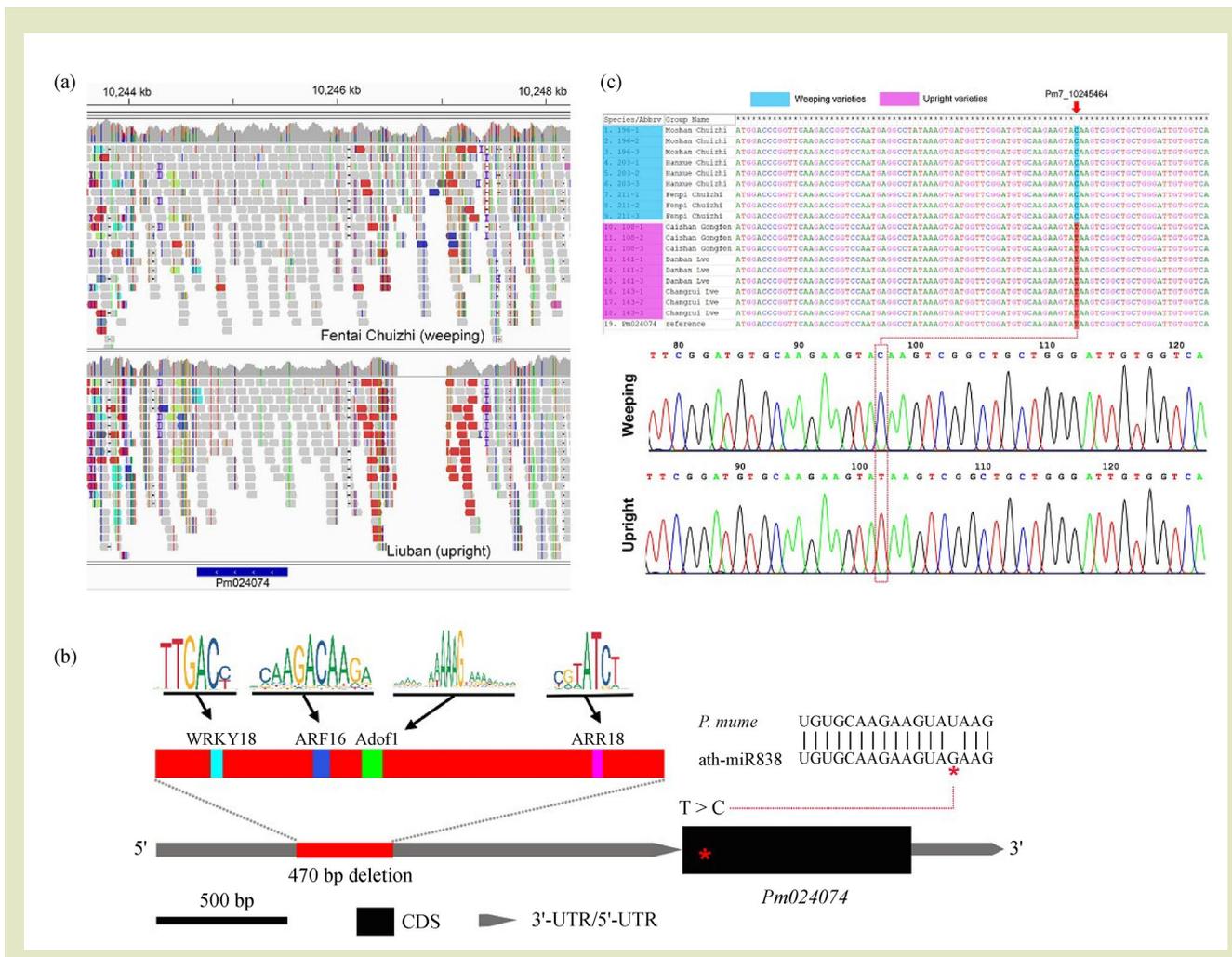


**Fig. 5** Variation analysis of *Pm024074* gene. (a) Analysis of genome selection signaling underlying *Pm024074*. Values of  $\pi$  and  $F_{st}$  were calculated using SNPs in *Pm024074* and its flanking 2-kb with a 100-bp sliding window. (b) Identification of exclusive allele variations in weeping *P. mume*. A weeping specific allele (*Pm\_10245464*) was identified in *Pm024074* coding sequences. (c) Validation of *Pm\_10245464* allele using MassARRAY compact system. (d) Comparative analysis of allele frequency of *Pm\_10245464* between weeping and upright *P. mume* in two different populations. (e) The allele C genotype of *Pm\_10245464* SNP strongly associated with the weeping trait. CC genotype showed larger branch angle and weeping angle than TT genotype.

*UGT72B3* (*AT1G01420.1*) and is defined as *PmUGT72B3*. UGT can catalyze sugar moiety transfer to the specific acceptor which is important in plant growth and development via conjugation of phytohormones with sugars and amino acids [66,67]. Networks around *PmUGT72B3* were extracted from the brown module with edge weight > 0.45 based on WGCNA analysis to investigate the upstream or downstream regulated genes of *PmUGT72B3*. The gene co-expression network shows that 23 genes associated

with secondary metabolism, hormone metabolism and transcription regulation were directly related to *PmUGT72B3*. The first and second neighboring genes of *PmUGT72B3* are shown in Fig. 7 and more detailed information on these genes is listed in Table S8.

The genes associated with auxin, gibberellin and lignin, including Chitin-inducible gibberellin-responsive protein 1



**Fig. 6** Identification of a deletion in candidate *Pm024074* promoter region. (a) Alignment of resequencing reads from upright and weeping parents. A 470-bp deletion was identified in *Pm024074* promoter region of upright parent. (b) Prediction of transcription factor binding sites in a 470-bp deletion sequences and miRNA targeted sites in the flanking Pm7\_10245464 SNP. (c) Cloning of *Pm024074* using cDNA from heterozygous weeping and upright accessions. Allele C was specifically expressed in weeping *Prunus mume*.

(*Pm029452*), Gibberellin 2-oxidase 1 (*Pm011163*), Phytochrome-associated protein 1 (*Pm005182*) and 12-oxophyto-dienoate reductase 2 (*Pm027056*), which correlates with Trans-caffeoyl-CoA 3-O-methyltransferase (*CCoAMT*, *Pm027248*) are of particular interest among the 23 main hub genes in the *PmUGT72B3* network. All these genes were upregulated except the lignin-related *CCoAMT* gene. In addition, transcription factors putatively binding to a 470-bp deletion region of *PmUGT72B3* promoter are also potentially crucial genes involved in the regulation of weeping traits, including WRKY18 (*Pm005698*), Aux/IAA (*Pm005182*), cycling DOF factor 3 (*CDF3*, *Pm006485*), and early-phytochrome-responsive 1 (*EPR1*, *Pm025253*) potentially binding site. Notably, multiple co-expressed genes are associated with the phytochrome signal such as *CIGR1*, *PAP1*, *EPR1* and *CDF3*<sup>[68–70]</sup>, suggesting that

*PmUGT72B3* might control weeping traits via influencing phytochrome. Therefore, a possible hypothesis is that *PmUGT72B3* might influence the expression of genes associated with hormones, lignin and light signal to collaboratively manipulate weeping phenotype in *P. mume*.

## 4 DISCUSSION

BSA-seq is an efficient and economic approach for the identification of genetic variants for important horticultural or agronomic traits<sup>[19,71,72]</sup>. Weeping growth habits are popular in woody rosaceous plants but their genetic mechanisms vary greatly between different species, including recessive, dominant and epistatic genetic effects<sup>[15,16,33]</sup>. Non-allelic weeping alleles

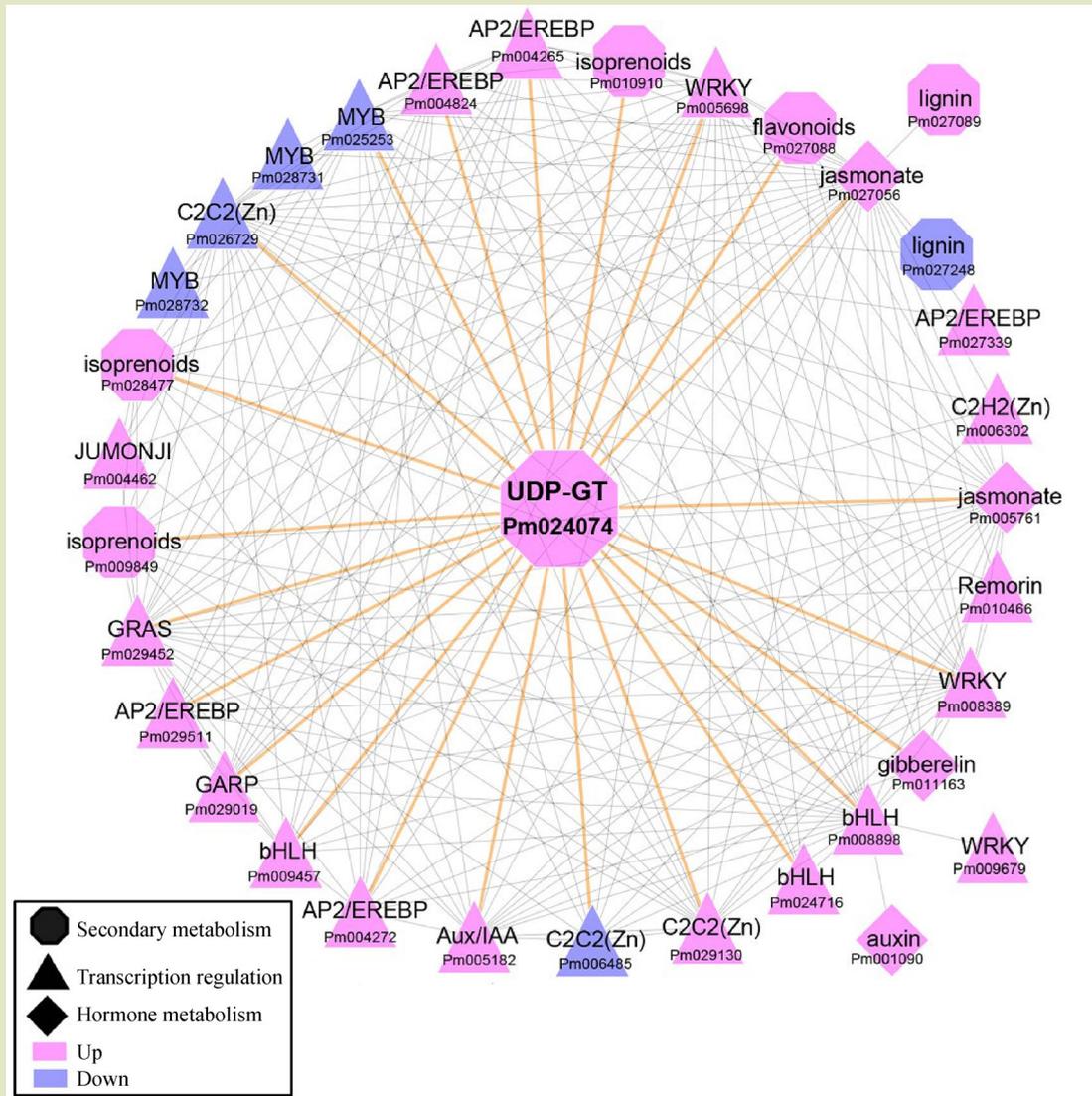


Fig. 7 The co-expression network of *Pm024074* neighboring genes. The first and second neighboring genes associated with hormone, secondary metabolism, and transcription factors with edge weight > 0.45 are included. First neighboring genes of *Pm024074* are indicated by the yellow line.

have been identified in *Cercis canadensis*<sup>[30]</sup>. These findings suggest that the genetic control of weeping traits is very complicated. In *P. mume*, our earlier study indicates that the weeping traits were controlled by a major recessive locus together with some modified loci<sup>[15]</sup>. However, insufficient markers and gene expression information hindered the identification of candidates and application of molecular assisted selection of weeping traits in *P. mume* breeding programs.

Here, dynamic investigation of the shoot growth habits of weeping and upright *P. mume* show that the meristem and elongation regions might be key parts for the determination of weeping traits. Integrated transcriptome analysis, WGCNA,

bulked segregant analysis and selective sweep analysis provide new insights into genetic control and regulation mechanisms of weeping *P. mume*. We identified five QTLs located on Pm7 potentially associated with weeping traits of *P. mume* and a key candidate gene involved in glucosylation of lignin precursors and cambial growth. A causative allele located at the candidate was first identified and validated in *P. mume*. These findings provide valuable information for understanding the genetic regulation of weeping traits and for molecular breeding in *P. mume*.

Gravity and light sensing are two of the most important environmental processes that control shoot growth direction in

plants<sup>[73–76]</sup>. In addition, stem mechanical rigidity via the formation of tension wood on the upper side of stems is also a crucial factor in the upward support of the stem through increasing a contractile force<sup>[77]</sup>. Here, we have found that the shoots of weeping trees displayed gravitropic responses at an early stage of bud sprouting and the stem elongation significantly increased branch weeping, causing a larger branch angle in weeping than in upright shoots (Fig. 1). When we settled the basal branch to reduce the stress of gravity the shoot tips maintained upward growth. In addition, the classic starch statolith hypothesis suggests that starch makes an important contribution to gravity sensing which occurs in the endodermis of the stem<sup>[73]</sup>. In *Arabidopsis*, studies have shown that the endodermis plays an important role in shoot gravity sensing, and the inflorescence stems of mutants show no response to gravity in darkness<sup>[78,79]</sup>. According to the phenotype observation results we suspect that the endodermis in young shoot tissues (meristem or the elongation region) might be a crucial part in the formation of weeping traits. Tree architecture is highly plastic and largely affected by environmental stimuli. Certainly, more investigations into phototropism, gravitropism, tissue anatomy, hormonal composition and wood strength are needed to comprehensively investigate their contributions to the weeping phenotype in *P. mume*.

Bulked segregant analysis strategy RNA-seq was conducted to reveal a potential molecular mechanism underlying the weeping trait in *P. mume*. Overall, the number of expressed genes was higher in weeping than in upright trees, and most genes were upregulated in young buds and stems of weeping trees (Fig. 2(e–f)). Similar expression regulation was also found in weeping shoot tips in peach<sup>[17]</sup>. When comparing buds, numerous specific DEGs associated with secondary, stress, hormone metabolism, RNA, protein, transport, development and signaling were significantly downregulated in the stems of weeping *P. mume* (Fig. S4(b)). Functional categories show that these genes have important roles in the regulation of phototropism, gravitropism and wood strength. Some genes involved in cell walls, lignin biosynthesis, auxin response and phototropic response are significantly downregulated by more than double (Data set S1). For instance, 1-aminocyclopropane-1-carboxylic acid synthase regulates cell wall function via interaction with FEI proteins<sup>[80]</sup>; changes in the expression of auxin related SAUR-like and IAA carboxyl methyltransferase may alter plant stem elongation, gravitropism, phototropism and gene transcription<sup>[81–84]</sup>. In other weeping species, numerous DEGs associated with hormone and light signaling transduction, wood development and tropism are also significantly enriched, for example in crape myrtle<sup>[85]</sup>, *Salix matsudana*<sup>[86]</sup> and peach<sup>[17]</sup>.

Next generation BSA has been a particularly fruitful approach for rapidly identifying QTLs associated with a trait of interest<sup>[19,41]</sup>. Here, five QTL loci on chromosome 7 were successfully mapped based on BSR and one overlapping QTL was confirmed in cultivars and hybrids. This indicates that the targeted QTL region represents a unique source of gene variants that can be exploited in bulked RNA-seq analysis. The present study has uncovered four QTLs significantly linked to weeping traits inherited from ‘Fentai Chuizhi’. This appears to be consistent with the segregation of phenotypes that some offspring represented diverse phenotypes, such as large branch angle or flat branches (Fig. S10). Our previous study also shows 95% confident QTL intervals almost covered the whole chromosome 7<sup>[15]</sup>. These five QTLs might therefore be an important region of chromosome 7 in controlling weeping traits.

Recently, our team found that multiple epistatic loci of weeping and upright *P. mume*, including overlapping Pm7-qt1-4 QTL, strongly interacted with another significant QTL located at 11.0–11.3 Mb on the same chromosome and explained 13.9% of the phenotypic variation. Similar genetic interactions of weeping traits have also been found in other rosaceous species such as weeping peach and weeping apple<sup>[16,33]</sup>. In a future study it would be interesting to determine if these two interacting loci coregulate the weeping traits in *P. mume*. Overlapping Pm7-qt1-4 is considered to be an important QTL for controlling weeping traits, and QTLs and linkage markers around this region were also detected in our previous studies<sup>[15,49,87,88]</sup>. However, the core allele variations and candidate genes remain unclear. Here, we have identified a weeping allele called *Pm<sup>W</sup>*, located at the *PmUGT72B3* (*Pm024074*) gene, potentially coding for a UGT protein with a conserved coniferyl-alcohol glucosyltransferase domain. Proteins with this domain make important contributions to the glucosylation of lignin precursors and cambial growth<sup>[61,89,90]</sup>. The *Pm<sup>W</sup>* allele located in this candidate was a synonymous mutation SNP, which was predicted to be targeted by miR838. The promoter region of the candidate gene has a 470-bp deletion in upright *P. mume* potentially containing multiple TF binding sites. However, whether these variations impact weeping traits remains to be demonstrated by molecular biological evidence, but both our gene expression evidence and SNP validated results support the hypothesis that *PmUGT72B3* controls weeping traits in *P. mume* (Fig. 5, Fig. 6). Several genetic and molecular biology studies have indicated that UDP-glucosyltransferase family genes are involved in the regulation of plant architecture<sup>[91–93]</sup>.

Auxin conjugates are known to be critical in maintaining auxin homeostasis in higher plants, mainly ester-linked IAA-sugar conjugates, amide-linked IAA-amino acid conjugates and

methyl esters of IAA<sup>[82,92,94]</sup>. Glycosylation of auxin is known to be an important mechanism to plant architecture by maintaining auxin homeostasis. In rice, a mutant of the UDP-glucosyltransferase family member,  $\alpha$ ,1,3-fucosyltransferase, can reduce basipetal auxin transport and gravitropic response and increase tiller angle<sup>[93]</sup>. Overexpression of *OsIAAGLU* and *OsIAGT1* also resulted in reduced sensitivity resulting in reduced sensitivity of auxin and altered gravitropic response of the roots in the transgenic rice<sup>[91,95]</sup>. These results support the hypothesis that exceptionally high expression of *PmUGT72B3* in buds and young stems might reduce the sensitivity of the IAA and gravitropic responses, and result in a weeping phenotype in weeping *P. mume*. In addition, several hormone, light and lignin hub genes directly correlated to *PmUGT72B3* are also identified via WGCNA, including (*ILR-like5* and *GH3* related IAA-amino acid conjugate, Aux/IAA related auxin active repressors, *CCoAMT* related lignin biosynthesis, *LHY* related circadian rhythm, *SCARECROW-like 21* related far-red light response, *Zeatin O-xylosyltransferase* related cytokinin, and *Gibberellin 2-oxidase* GA related biosynthesis (Table S8). These genes are essential in the control of gravitropic growth and development in light, affecting shoot or root architecture<sup>[1,28,54,96]</sup>. Although little is known about the functions of UDP-glucosyltransferase in the regulation of tree architecture, futures studies on the

regulated relationship between *PmUGT72B3* and its neighboring genes may help in understanding the genetic mechanism underlying the weeping phenotype in *P. mume*.

## 5 CONCLUSIONS

The phenotype observations show that the weeping traits might be independently determined by the young meristem and stem. According to the BSR-seq analysis and WGCNA, we have identified a core candidate *PmUGT72B3* coding a protein containing conserved coniferyl-alcohol glucosyltransferase, which is a key protein in regulating lignin and IAA. Compared to upright *P. mume*, *PmUGT72B3* was exceptionally highly expressed in young buds and stems of weeping *P. mume*. Higher expression of *UDP-glucosyltransferase* appears to act by reducing the gravitropic response and inactivating IAA. Variations in *PmUGT72B3* were detected and validated, and were significantly associated with weeping phenotypes. Further molecular experiments are required to verify its molecular functions. Identification of this candidate gene from weeping *P. mume* may greatly contribute to our understanding of the genetics basis of the weeping phenotypes in rosaceous trees and provide benefits for horticultural productivity.

### Supplementary materials

The online version of this article at <https://doi.org/10.15302/J-FASE-2020379> contains supplementary materials.

### Acknowledgements

The research was funded by the National Key Research and Development Program of China (2018YFD1000401), the program for Science and Technology of Beijing (Z18110002418006), the National Natural Science Foundation of China (31870689) and the Special Fund for Beijing Common Construction Project.

### Compliance with ethics guidelines

Xiaokang Zhuo, Tangchun Zheng, Zhiyong Zhang, Suzhen Li, Yichi Zhang, Lidan Sun, Weiru Yang, Jia Wang, Tangren Cheng, and Qixiang Zhang declare that they have no conflicts of interest or financial conflicts to disclose. This article does not contain any studies with human or animal subjects performed by any of the authors.

## REFERENCES

1. Wang B, Smith S M, Li J. Genetic regulation of shoot architecture. *Annual Review of Plant Biology*, 2018, **69**(1): 437–468
2. Barthélémy D, Caraglio Y. Plant architecture: a dynamic, multilevel and comprehensive approach to plant form, structure and ontogeny. *Annals of Botany*, 2007, **99**(3): 375–407
3. Toyota M, Gilroy S. Gravitropism and mechanical signaling in plants. *American Journal of Botany*, 2013, **100**(1): 111–125
4. Osada N, Hiura T. How is light interception efficiency related to shoot structure in tall canopy species? *Oecologia*, 2017, **185**(1): 29–41
5. Reinhardt D, Pesce E R, Stieger P, Mandel T, Baltensperger K, Bennett M, Traas J, Friml J, Kuhlemeier C. Regulation of phyllotaxis by polar auxin transport. *Nature*, 2003, **426**(6964):

- 255–260
6. Yoshihara T, Spalding E P. *LAZY* genes mediate the effects of gravity on auxin gradients and plant architecture. *Plant Physiology*, 2017, **175**(2): 959–969
  7. Satoh-Nagasawa N, Nagasawa N, Malcomber S, Sakai H, Jackson D. A trehalose metabolic enzyme controls inflorescence architecture in maize. *Nature*, 2006, **441**(7090): 227–230
  8. Cantín C M, Arús P, Eduardo I. Identification of a new allele of the *Dw* gene causing brachytic dwarfing in peach. *BMC Research Notes*, 2018, **11**(1): 386
  9. Cheng J, Zhang M, Tan B, Jiang Y, Zheng X, Ye X, Guo Z, Xiong T, Wang W, Li J, Feng J. A single nucleotide mutation in *GID1c* disrupts its interaction with *DELLA1* and causes a GA-insensitive dwarf phenotype in peach. *Plant Biotechnology Journal*, 2019, **17**(9): 1723–1735
  10. Hollender C A, Hadiarto T, Srinivasan C, Scorza R, Dardick C. A brachytic dwarfism trait (*dw*) in peach trees is caused by a nonsense mutation within the gibberellic acid receptor *PpeGID1c*. *New Phytologist*, 2016, **210**(1): 227–239
  11. Wolters P J, Schouten H J, Velasco R, Si-Ammour A, Baldi P. Evidence for regulation of columnar habit in apple by a putative 2OG-Fe(II) oxygenase. *New Phytologist*, 2013, **200**(4): 993–999
  12. Petersen R, Krost C. Tracing a key player in the regulation of plant architecture: the columnar growth habit of apple trees (*Malus × domestica*). *Planta*, 2013, **238**(1): 1–22
  13. Otto D, Petersen R, Brauksiepe B, Braun P, Schmidt E R. The columnar mutation (“*Co* gene”) of apple (*Malus × domestica*) is associated with an integration of a Gypsy-like retrotransposon. *Molecular Breeding*, 2014, **33**(4): 863–880
  14. Dardick C, Callahan A, Horn R, Ruiz K B, Zhebentyayeva T, Hollender C, Whitaker M, Abbott A, Scorza R. *PpeTAC1* promotes the horizontal growth of branches in peach trees and is a member of a functionally conserved gene family found in diverse plants species. *Plant Journal*, 2013, **75**(4): 618–630
  15. Zhang J, Zhang Q, Cheng T, Yang W, Pan H, Zhong J, Huang L, Liu E. High-density genetic map construction and identification of a locus controlling weeping trait in an ornamental woody plant (*Prunus mume* Sieb. et Zucc). *DNA Research*, 2015, **22**(3): 183–191
  16. Dougherty L, Singh R, Brown S, Dardick C, Xu K. Exploring DNA variant segregation types in pooled genome sequencing enables effective mapping of weeping trait in *Malus*. *Journal of Experimental Botany*, 2018, **69**(7): 1499–1516
  17. Hollender C A, Pascal T, Tabb A, Hadiarto T, Srinivasan C, Wang W, Liu Z, Scorza R, Dardick C. Loss of a highly conserved sterile alpha motif domain gene (*WEEP*) results in pendulous branch growth in peach trees. *Proceedings of the National Academy of Sciences of the United States of America*, 2018, **115**(20): E4690–E4699
  18. Shi L, Jiang C, He Q, Habekuß A, Ordon F, Luan H, Shen H, Liu J, Feng Z, Zhang J, Yang P. Bulk segregant RNA-sequencing (BSR-seq) identified a novel rare allele of *eIF4E* effective against multiple isolates of *BaYMV/BaMMV*. *Theoretical and Applied Genetics*, 2019, **132**(6): 1777–1788
  19. Liu S, Yeh C T, Tang H M, Nettleton D, Schnable P S. Gene mapping via bulked segregant RNA-Seq (BSR-Seq). *PLoS One*, 2012, **7**(5): e36406
  20. Li R, Hou Z, Gao L, Xiao D, Hou X, Zhang C, Yan J, Song L. Conjunctive analyses of BSA-Seq and BSR-Seq to reveal the molecular pathway of leafy head formation in Chinese cabbage. *Plants*, 2019, **8**(12): 603
  21. Li M, Li B, Guo G, Chen Y, Xie J, Lu P, Wu Q, Zhang D, Zhang H, Yang J, Zhang P, Zhang Y, Liu Z. Mapping a leaf senescence gene *els1* by BSR-Seq in common wheat. *Crop Journal*, 2018, **6**(3): 236–243
  22. Tan C, Liu Z, Huang S, Feng H. Mapping of the male sterile mutant gene *ftms* in *Brassica rapa* L. ssp. *pekinensis* via BSR-Seq combined with whole-genome resequencing. *Theoretical and Applied Genetics*, 2019, **132**(2): 355–370
  23. Wu P, Xie J, Hu J, Qiu D, Liu Z, Li J, Li M, Zhang H, Yang L, Liu H, Zhou Y, Zhang Z, Li H. Development of molecular markers linked to powdery mildew resistance gene *Pm4b* by combining SNP discovery from transcriptome sequencing data with bulked segregant analysis (BSR-Seq) in wheat. *Frontiers of Plant Science*, 2018, **9**: 95
  24. Hao Z, Geng M, Hao Y, Zhang Y, Zhang L, Wen S, Wang R, Liu G. Screening for differential expression of genes for resistance to *Sitodiplosis mosellana* in bread wheat via BSR-seq analysis. *Theoretical and Applied Genetics*, 2019, **132**(11): 3201–3221
  25. Du H, Zhu J, Su H, Huang M, Wang H, Ding S, Zhang B, Luo A, Wei S, Tian X, Xu Y. Bulk segregant RNA-seq reveals differential expression and SNPs of candidate genes associated with waterlogging tolerance in maize. *Frontiers of Plant Science*, 2017, **8**: 1022
  26. Pastinen T. Genome-wide allele-specific analysis: insights into regulatory variation. *Nature Reviews: Genetics*, 2010, **11**(8): 533–538
  27. Zou C, Wang P, Xu Y. Bulk sample analysis in genetics, genomics and crop improvement. *Plant Biotechnology Journal*, 2016, **14**(10): 1941–1955
  28. Hollender C A, Dardick C. Molecular basis of angiosperm tree architecture. *New Phytologist*, 2015, **206**(2): 541–556
  29. Sampson D, Cameron D. Inheritance of bronze foliage extra petals and pendulous habit in ornamental crabapples. In: *Proceedings of the American Society for Horticultural Science*, 1965, **86**: 717
  30. Roberts D J, Werner D J, Wadl P A, Trigiano R N. Inheritance and allelism of morphological traits in eastern redbud (*Cercis canadensis* L.). *Horticulture Research*, 2015, **2**(1): 15049
  31. Chaparro J X, Werner D J, O’Malley D, Sederoff R R. Targeted mapping and linkage analysis of morphological isozyme, and RAPD markers in peach. *Theoretical and Applied Genetics*, 1994, **87**(7): 805–815
  32. Kotobuki K, Sawamura Y, Saito T, Takada N. The mode of inheritance of weeping habit in Japanese chestnut, *Castanea*

- crenata*. In: *III International Chestnut Congress 693*, 2004, 477–484
33. Werner D J, Chaparro J X. Genetic interactions of pillar and weeping peach genotypes. *HortScience*, 2005, **40**(1): 18–20
  34. Crowell S, Korniliev P, Falcão A, Ismail A, Gregorio G, Mezey J, McCouch S. Genome-wide association and high-resolution phenotyping link *Oryza sativa* panicle traits to numerous trait-specific QTL clusters. *Nature Communications*, 2016, **7**(1): 10527
  35. Ferreira T, Rasband W. ImageJ user guide. Ferreira: *Tiago*, 2012, **1**: 155–161
  36. Bolger A M, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*, 2014, **30**(15): 2114–2120
  37. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics*, 2009, **25**(14): 1754–1760
  38. McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytsky A, Garimella K, Altshuler D, Gabriel S, Daly M, DePristo M A. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Research*, 2010, **20**(9): 1297–1303
  39. Danecek P, Auton A, Abecasis G, Albers C A, Banks E, DePristo M A, Handsaker R E, Lunter G, Marth G T, Sherry S T, McVean G, Durbin R. The variant call format and VCFtools. *Bioinformatics*, 2011, **27**(15): 2156–2158
  40. Mansfeld B N, Grumet R. QTLseqr: an R Package for bulk segregant analysis with next-generation sequencing. *Plant Genome*, 2018, **11**(2): 180006
  41. Magwene P M, Willis J H, Kelly J K. The statistics of bulk segregant analysis using next generation sequencing. *PLoS Computational Biology*, 2011, **7**(11): e1002255
  42. Turner S D. qqman: an R package for visualizing GWAS results using QQ and manhattan plots. 2014: 005165 doi: 10.1101/005165
  43. Anders S. Analysing RNA-Seq data with the DESeq package. *Molecular Biology*, 2010, **43**(4): 1–17
  44. Young M D, Wakefield M J, Smyth G K, Oshlack A. Gene ontology analysis for RNA-seq: accounting for selection bias. *Genome Biology*, 2010, **11**(2): R14
  45. Moriya Y, Itoh M, Okuda S, Yoshizawa A C, Kanehisa M. KAAS: an automatic genome annotation and pathway reconstruction server. *Nucleic Acids Research*, 2007, **35**(suppl 2): W182–W185
  46. Lohse M, Nagel A, Herter T, May P, Schroda M, Zrenner R, Tohge T, Fernie A R, Stitt M, Usadel B. Mercator: a fast and simple web server for genome scale functional annotation of plant sequence data. *Plant, Cell & Environment*, 2014, **37**(5): 1250–1258
  47. Langfelder P, Horvath S. WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics*, 2008, **9**(1): 559
  48. Shannon P, Markiel A, Ozier O, Baliga N S, Wang J T, Ramage D, Amin N, Schwikowski B, Ideker T. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Research*, 2003, **13**(11): 2498–2504
  49. Zhang Q, Zhang H, Sun L, Fan G, Ye M, Jiang L, Liu X, Ma K, Shi C, Bao F, Guan R, Han Y, Fu Y, Pan H, Chen Z, Li L, Wang J, Lv M, Zheng T, Yuan C, Zhou Y, Lee S M Y, Yan X, Xu X, Wu R, Chen W, Cheng T. The genetic architecture of floral traits in the woody plant *Prunus mume*. *Nature Communications*, 2018, **9**(1): 1702
  50. Gabriel S, Ziaugra L, Tabbaa D. SNP genotyping using the Sequenom MassARRAY iPLEX platform. *Current Protocols in Human Genetics*, 2009, **Chapter 2**(1): 12
  51. Xu Z, Sun L, Zhou Y, Yang W, Cheng T, Wang J, Zhang Q. Identification and expression analysis of the SQUAMOSA promoter-binding protein (SBP)-box gene family in *Prunus mume*. *Molecular Genetics and Genomics*, 2015, **290**(5): 1701–1715
  52. Livak K J, Schmittgen T D. Analysis of relative gene expression data using real-time quantitative PCR and the  $2^{-\Delta\Delta CT}$  Method. *Methods*, 2001, **25**(4): 402–408
  53. Zhou Y, Xu Z, Yong X, Ahmad S, Yang W, Cheng T, Wang J, Zhang Q. SEP-class genes in *Prunus mume* and their likely role in floral organ development. *BMC Plant Biology*, 2017, **17**(1): 10
  54. Hill J L Jr, Hollender C A. Branching out: new insights into the genetic regulation of shoot architecture in trees. *Current Opinion in Plant Biology*, 2019, **47**: 73–80
  55. Ferguson B J, Beveridge C A. Roles for auxin, cytokinin, and strigolactone in regulating shoot branching. *Plant Physiology*, 2009, **149**(4): 1929–1944
  56. Sun J, Qi L, Li Y, Zhai Q, Li C. PIF4 and PIF5 transcription factors link blue light and auxin to regulate the phototropic response in *Arabidopsis*. *Plant Cell*, 2013, **25**(6): 2102–2114
  57. Xu F, He S, Zhang J, Mao Z, Wang W, Li T, Hua J, Du S, Xu P, Li L, Lian H, Yang H Q. Photoactivated CRY1 and phyB interact directly with AUX/IAA proteins to inhibit auxin signaling in *Arabidopsis*. *Molecular Plant*, 2018, **11**(4): 523–541
  58. Guo D, Chen F, Inoue K, Blount J W, Dixon R A. Down-regulation of caffeic acid 3-O-methyltransferase and caffeoyl CoA 3-O-methyltransferase in transgenic alfalfa. Impacts on lignin structure and implications for the biosynthesis of G and S lignin. *Plant Cell*, 2001, **13**(1): 73–88
  59. Tong H, Jin Y, Liu W, Li F, Fang J, Yin Y, Qian Q, Zhu L, Chu C. DWARF AND LOW-TILLERING, a new member of the GRAS family, plays positive roles in brassinosteroid signaling in rice. *Plant Journal*, 2009, **58**(5): 803–816
  60. Chen L, Xiong G, Cui X, Yan M, Xu T, Qian Q, Xue Y, Li J, Wang Y. OsGRAS19 may be a novel component involved in the brassinosteroid signaling pathway in rice. *Molecular Plant*, 2013, **6**(3): 988–991
  61. Savidge R A, Förster H. Seasonal activity of uridine 5'-diphosphoglucose: coniferyl alcohol glucosyltransferase in relation to cambial growth and dormancy in conifers. *Canadian Journal of Botany*, 1998, **76**(3): 486–493 doi:10.1139/b98-015
  62. Ibrahim R K. Glucosylation of lignin precursors by uridine diphosphate glucose: coniferyl alcohol glucosyltransferase in

- higher plants. *Zeitschrift für Pflanzenphysiologie*, 1977, **85**(3): 253–262
63. Dong N, Sun Y, Guo T, Shi C, Zhang Y, Kan Y, Xiang Y, Zhang H, Yang Y, Li Y, Zhao H, Yu H, Lu Z, Wang Y, Ye W, Shan J, Lin H. UDP-glucosyltransferase regulates grain size and abiotic stress tolerance associated with metabolic flux redirection in rice. *Nature Communications*, 2020, **11**(1): 2629
64. Chen J. China Mei flower (*Prunus mume*) cultivars in colour. Beijing: *China Forestry Publishing House*, 2017 (in Chinese).
65. Chen Z. Genetic and epigenetic mechanisms for gene expression and phenotypic variation in plant polyploids. *Annual Review of Plant Biology*, 2007, **58**(1): 377–406
66. Maciej O, Anna J. UDP-glycosyltransferases of plant hormones. *Medical Journal of Cell Biology*, 2014, **4**(1): 43–60
67. Piotrowska A, Bajguz A. Conjugates of abscisic acid, brassinosteroids, ethylene, gibberellins, and jasmonates. *Phytochemistry*, 2011, **72**(17): 2097–2112
68. Torres-Galea P, Hirtreiter B, Bolle C. Two GRAS proteins, SCARECROW-LIKE21 and PHYTOCHROME A SIGNAL TRANSDUCTION1, function cooperatively in phytochrome A signal transduction. *Plant Physiology*, 2013, **161**(1): 291–304
69. Kuno N, Møller S G, Shinomura T, Xu X, Chua N H, Furuya M. The novel MYB protein EARLY-PHYTOCHROME-RESPONSIVE1 is a component of a slave circadian oscillator in *Arabidopsis*. *Plant Cell*, 2003, **15**(10): 2476–2488
70. Park D H, Lim P O, Kim J S, Cho D S, Hong S H, Nam H G. The *Arabidopsis* COG1 gene encodes a Dof domain transcription factor and negatively regulates phytochrome signaling. *Plant Journal*, 2003, **34**(2): 161–171
71. Heng S, Huang H, Cui M, Liu M, Lv Q, Hu P, Ren S, Li X, Fu T, Wan Z. Rapid identification of the *BjRCO* gene associated with lobed leaves in *Brassica juncea* via bulked segregant RNA-seq. *Molecular Breeding*, 2020, **40**(4): No. 42
72. Huang Z, Peng G, Gossen B D, Yu F. Fine mapping of a clubroot resistance gene from turnip using SNP markers identified from bulked segregant RNA-Seq. *Molecular Breeding*, 2019, **39**(9): 131
73. Morita M T. Directional gravity sensing in gravitropism. *Annual Review of Plant Biology*, 2010, **61**(1): 705–720
74. Liscum E, Askinosie S K, Leuchtman D L, Morrow J, Willenburg K T, Coats D R. Phototropism: growing towards an understanding of plant movement. *Plant Cell*, 2014, **26**(1): 38–55
75. Gilroy S. Plant tropisms. *Current Biology*, 2008, **18**(7): R275–R277
76. Nakamura M, Nishimura T, Morita M T. Gravity sensing and signal conversion in plant gravitropism. *Journal of Experimental Botany*, 2019, **70**(14): 3495–3506
77. Groover A. Gravitropisms and reaction woods of forest trees—evolution, functions and mechanisms. *New Phytologist*, 2016, **211**(3): 790–802
78. Fukaki H, Fujisawa H, Tasaka M. *SGR1*, *SGR2*, *SGR3*: novel genetic loci involved in shoot gravitropism in *Arabidopsis thaliana*. *Plant Physiology*, 1996, **110**(3): 945–955
79. Fukaki H, Wysocka-Diller J, Kato T, Fujisawa H, Benfey P N, Tasaka M. Genetic evidence that the endodermis is essential for shoot gravitropism in *Arabidopsis thaliana*. *Plant Journal*, 1998, **14**(4): 425–430
80. Xu S L, Rahman A, Baskin T I, Kieber J J. Two leucine-rich repeat receptor kinases mediate signaling, linking cell wall biosynthesis and ACC synthase in *Arabidopsis*. *Plant Cell*, 2008, **20**(11): 3065–3079
81. Leyser H M O, Pickett F B, Dharmasiri S, Estelle M. Mutations in the *AXR3* gene of *Arabidopsis* result in altered auxin response including ectopic expression from the *SAUR-AC1* promoter. *Plant Journal*, 1996, **10**(3): 403–413
82. Qin G, Gu H, Zhao Y, Ma Z, Shi G, Yang Y, Pichersky E, Chen H, Liu M, Chen Z, Qu L J. An indole-3-acetic acid carboxyl methyltransferase regulates *Arabidopsis* leaf development. *Plant Cell*, 2005, **17**(10): 2693–2704
83. Leyser O. Auxin Signaling. *Plant Physiology*, 2018, **176**(1): 465–479
84. Guilfoyle T J, Hagen G. Auxin response factors. *Current Opinion in Plant Biology*, 2007, **10**(5): 453–460
85. Li S, Zheng T, Zhuo X, Li Z, Li L, Li P, Qiu L, Pan H, Wang J, Cheng T, Zhang Q. Transcriptome profiles reveal that gibberellin-related genes regulate weeping traits in crape myrtle. *Horticulture Research*, 2020, **7**(1): 54
86. Liu J, Zeng Y, Yan P, He C, Zhang J. Transcriptional and hormonal regulation of weeping trait in *Salix matsudana*. *Genes*, 2017, **8**(12): 359
87. Li S, Zheng T, Zhuo X, Li L, Qiu L, Wang J, Cheng T, Zhang Q. Identification of SNP markers linked to the weeping trait in *Prunus mume*. *Euphytica*, 2019, **215**(10): 168
88. Zhang J, Zhao K, Hou D, Cai J, Zhang Q, Cheng T, Pan H, Yang W. Genome-wide discovery of DNA polymorphisms in mei (*Prunus mume* Sieb. et Zucc.), an ornamental woody plant, with contrasting tree architecture and their functional relevance for weeping trait. *Plant Molecular Biology Reporter*, 2017, **35**(1): 37–46
89. Ibrahim R K. Glucosylation of lignin precursors by uridine diphosphate glucose: coniferyl alcohol glucosyltransferase in higher plants. *Zeitschrift für Pflanzenphysiologie*, 1977, **85**(3): 253–262
90. Förster H, Steeves V, Pommer U, Savidge R. UDPG: coniferyl alcohol glucosyltransferase and coniferin biosynthesis—a regulatory link to seasonal cambial growth in conifers. Oxford: *Cell and molecular biology of wood formation*, 2000, 189–201
91. Yu X, Wang H, Leung D W, He Z, Zhang J, Peng X, Liu E. Overexpression of *OsIAAGLU* reveals a role for IAA-glucose conjugation in modulating rice plant architecture. *Plant Cell Reports*, 2019, **38**(6): 731–739
92. Tognetti V B, Van Aken O, Morreel K, Vandenbroucke K, van de Cotte B, De Clercq I, Chiwocha S, Fenske R, Prinsen E, Boerjan W, Genty B, Stubbs K A, Inzé D, Van Breusegem F. Perturbation of indole-3-butyric acid homeostasis by the UDP-glucosyltransferase UGT74E2 modulates *Arabidopsis* architecture and water stress tolerance. *Plant Cell*, 2010, **22**(8): 2660–

- 2679
93. Harmoko R, Yoo J Y, Ko K S, Ramasamy N K, Hwang B Y, Lee E J, Kim H S, Lee K J, Oh D B, Kim D Y, Lee S, Li Y, Lee S Y, Lee K O. N-glycan containing a core  $\alpha$ 1,3-fucose residue is required for basipetal auxin transport and gravitropic response in rice (*Oryza sativa*). *New Phytologist*, 2016, **212**(1): 108–122
94. Ludwig-Müller J. Auxin conjugates: their role for plant development and in the evolution of land plants. *Journal of Experimental Botany*, 2011, **62**(6): 1757–1773
95. Liu Q, Chen T, Xiao D, Zhao S, Lin J, Wang T, Li Y, Hou B. *OsIAGT1* is a glucosyltransferase gene involved in the glucose conjugation of auxins in rice. *Rice*, 2019, **12**(1): 92
96. Yu H, Moss B L, Jang S S, Prigge M, Klavins E, Nemhauser J L, Estelle M. Mutations in the TIR1 auxin receptor that increase affinity for auxin/indole-3-acetic acid proteins result in auxin hypersensitivity. *Plant Physiology*, 2013, **162**(1): 295–303