Industrial Crops & Products 219 (2024) 118975

Contents lists available at ScienceDirect

Industrial Crops & Products

journal homepage: www.elsevier.com/locate/indcrop

Cell type specific regulation of phenolic acid and flavonoid metabolism in *Taxus mairei* leaves

Xiaori Zhan, Xueshuang Liang, Wanting Lin, Ruoyun Ma, Yue Zang, Huizhong Wang, Lilin Wang, Yanjun Yang *, Chenjia Shen *

College of Life and Environmental Sciences, Hangzhou Normal University, Hangzhou 311121, China

ARTICLE INFO

Keywords:
MALDI-IMS
Multi-omics
Open chromatin region
ScATAC-seq
ScRNA-seq
Transcriptional regulation

ABSTRACT

The *Taxus* tree is an important industrial resource for the extraction of Taxol, a natural anticancer drug. Recently, numerous active compounds, such as phenolic acids and flavonoids, were identified in the leaves of *Taxus mairei*. However, the distribution and biosynthesis of phenolic acids and flavonoids in *Taxus* leaves are largely unknown. Integrated MALDI-IMS, ScRNA-seq and ScATAC-seq analysis was used to reveal the cell type-specific regulation of phenolic acid and flavonoid biosynthesis. Based on the cluster-specifically expressed genes, ScRNA-seq grouped the cells into 12 clusters, while ScATAC-seq grouped the cells into 9 clusters, indicating a high degree of leaf cell heterogeneity. Several marker genes were selected for cell type annotation, some of which were effectively verified by *in situ* hybridization. ScRNA-seq identified four phenolic acid and flavonoid biosynthesis-related genes (each one > 5 % cell coverage), which significantly expressed in the leaf epidermal cells. The cell-specific open chromatin regions surrounding the four phenolic acid and flavonoid biosynthesis-related genes were enriched in the epidermal cells. We further identified several cell type-specifically expressed transcription factors that might be involved in the biosynthesis of phenolic acids and flavonoids. Our study provided valuable information for studying the cell type specific regulation of phenolic acid and flavonoid biosynthesis in *Taxus* leaves.

1. Introduction

The genus *Taxus* is widely distributed throughout the world, with more than 20 species and 50 varieties (Moller et al., 2020). To date, most attention has been focused on taxoids, a class of well-identified anti--cancer active ingredients (Perez-Matas et al., 2023). Taxol (generically named paclitaxel), one of the most typical taxoids, is currently administrated for various cancers as well as AIDS-related diseases (Perez-Matas et al., 2022). In addition to taxoids, numerous active compounds with pharmacological activities have been isolated from *Taxus* trees, such as phenolic acids, flavonoids, and polysaccharides (Elansary et al., 2019; Zhang et al., 2020). The high commercial value results in serious illegal logging and pushes wild *Taxus* trees to the edge of extinction (Majeed et al., 2020; Xu et al., 2022). Thus, it is crucial for us to enhance exhaustive resource utilization of *Taxus* species.

Phenolic acids, such as caffeic acid and ferulic acid, are derived from the complex phenylpropanoid pathway (Valinas et al., 2015). Flavonoids and their derivatives are also derived from the phenylpropanoid pathway through a series of acylation, methylation, and glycosylation reactions (Hurtova et al., 2022). The highly conserved phenolic acid biosynthesis pathway has been well identified in plants (Fang et al., 2016; Waki et al., 2021; Zhang et al., 2019). To date, the regulation of phenolic acid and flavonoid biosynthesis in *Taxus* trees is largely unknown.

Although phenolic acids and flavonoids are byproducts of taxoids in industrial production, their important functions in human health remains highly valued (Ruan et al., 2013; Zhan et al., 2023). Flavonoids in *Taxus* trees are also recognized as natural bioactive ingredients that have multiple active effects, such as antibacterial, antioxidant and antiaging, anti-Alzheimer, and anticancer activities (Cai et al., 2022; Gauchan et al., 2020; Tabaszewska et al., 2021). So far, a large number of flavonoids have been isolated from the needles, fruits, twigs, and barks of various *Taxus* trees (Wei et al., 2023). Twenty years ago, several important flavonoids, such as 3-*O*-rutinosides 7-*O*-glucosides, kaempferol, quercetin, were isolated and identified from the needles of *T. baccata* by HPLC separation (Krauze-Baranowska, 2004). Seven

* Corresponding authors.
E-mail addresses: yjyang@hznu.edu.cn (Y. Yang), shencj@hznu.edu.cn (C. Shen).

https://doi.org/10.1016/j.indcrop.2024.118975
Received 13 April 2024; Received in revised form 25 May 2024; Accepted 10 June 2024
Available online 24 June 2024

classic flavonoids (isoquercitrin, quercitrin, flavonoids, quercetin, bilobetin, ginkgetin, sciadopitysin, and isoginkgetin) were identified in the twigs of three different *Taxus* species by an untargeted HPLC-MS/MS analysis (Gai et al., 2020). In the leaves of *T. media* and *T. mairei*, 197 flavonoids were detected by a untargeted UPLC-ESI-MS/MS method (Wang et al., 2019). Interestingly, seasonal variations of flavone in the leaves of *T. mairei* suggest a complex regulatory network in *Taxus* species (Yang et al., 2016). Flavonoid compounds with different skeletons displayed significant antibacterial, antiaging, anti-Alzheimer's, antidiabetes, anticancer, antidepressant, antileishmaniasis and anti-inflammatory (Wei et al., 2023). Further pharmacological and clinical experiments for phenolic acid and flavonoid could be accomplished to promote the utilization of the *Taxus* species.

The complex and diverse phenylpropanoid biosynthesis pathways ensure that plants can produce specific phenolic acids and flavonoids (Pu et al., 2022). Different transcription factor (TF) families have been reported to be involved in the transcriptional regulation of phenolic acid and flavonoid biosynthesis (Dixon et al., 2013; Zhao et al., 2022). In *Lycoris radiata*, WRKY3 and WRKY27 play roles in the biosynthesis of pelargonidin-3-O-glucoside-5-O-arabinoside, suggesting an essential function of the WRKY family in the regulation of flavonoid biosynthesis (Wang et al., 2023a). In grapevine, VqWRKY56 and VqZIPC22 are required for proanthocyanidin biosynthesis and resistance to powdery mildew (Wang et al., 2023b). In *Tartary buckwheat*, the nucleus-localized FtERF-EAR3 inhibits flavonoid biosynthesis by targeting the GCC-box element in the promoter of *FtF3H* (Ding et al., 2022). In *Salvia miltiorrhiza*, several TFs, such as SmSPL7, SmMYB36/76, and SmERF6/15, are involved in phenolic acid biosynthesis (Chen et al., 2021). Mining TFs is, therefore, an effective approach to uncover the transcriptional regulatory mechanism underlying the phenolic acid and flavonoid biosynthesis in *Taxus* trees.

Recently, several new technologies have been applied to the studies of plant secondary metabolism. For example, matrix-assisted laser desorption/ionization mass spectrometry imaging (MALDI-MSI) is used for *in situ* imaging of flavonoids in litchi seed sections (Liu et al., 2023). High-throughput single-cell RNA sequencing (scRNA-seq) gives an opportunity to reveal the transcriptional regulation at single-cell level (Liao and Wang, 2023). Assay for Transposase-Accessible Chromatin sequencing (ATAC-seq) is a novel method used to reveal key regulatory sequences, which in turn are associated with TFs (Marand et al., 2021). Although many phenolic acid and flavonoid biosynthesis-related genes were identified in *Taxus* trees, their cell type-specific expression regulation in *Taxus* leaves is largely unknown. To investigate the cell-specific regulation of phenolic acid and flavonoid biosynthesis in *T. mairei* leaves, three newly developed technologies, including MSI, scRNA-seq, and scATAC, were integrated in the present study. Our data provides new insights into the cell-type specific regulation of phenolic acid and flavonoid biosynthesis in medicinal plants.

2. Materials and methods

2.1. MS imaging and laser radiation parameters

Young leaves from five-year-old *T. mairei* trees were harvested for tissue sectioning. After being vacuum-dried with nitrogen gas, the resulting slices were stored at −80℃ until used. A commercial MALDI-2 special matrix solution was purchased from Sigma-Aldrich (Shanghai, China) and uniformly sprayed onto the glass slides.

The matrix coated slides were processed and screened using the TimsTOF flex MALDI-2 system with Bruker Data imaging software. The detection area was divided into a series of 2-D points according to the imaging resolution setting. Ionized molecules released molecules from sample slices were detected and imaged at the target sites to generate MS raw data files. The experimental parameters are set as follows as our previous work (Zhan et al., 2023).

2.2. Analysis of MS imaging data

The raw imaging data was imported into SCiLS™ Lab2021 software for parameter calculation to obtain the average MS data. The list of all ion peak mass was converted into imzML format and normalized using SCiLS software with the Root MeanSquare method. The normalized dataset can provide reliable results for spatial denoising and region segmentation. Spatial positions with the same metabolite accumulation patterns were mapped with the same color-coded to intuitively divide the slices. MS spectrum data was extracted and qualified by the Bruker Metaboscape™ workstation to produce annotation information for each metabolite. The resulting data was searched against the Bruker Library MS Metabase 3.0 database (https://www.bruker.com/en/products-and-solutions/mass-spectrometry/ms-software/metabolomics-spectral-libraries.html).

2.3. Single-cell nucleus extraction and isolation

All cells were resuspended and collected by a 20 μm cell sieve. The gel beads containing different barcodes were treated with the nuclei, and wrapped by oil droplets using a microfluidic system to form Gel Beads in Embodiments (GEMs). GEMs then were collected to release barcode sequences. To label the cell samples, the barcode sequences were reverse transcribed into cDNA fragments. The DEMs were broken to release cDNA fragments for PCR amplification.

2.4. ScRNA-seq library construction and data processing

The cDNA was digested into small fragments (200–300 bp) and used for second-generation sequencing processes, including sequencing connectors and primers. The library was constructed for high-throughput sequencing using the dual-terminal sequencing mode of the Illumina platform. The sequencing data was converted into FASTQ format using BCL2fastq software (v. 2.30). The raw data was filtered and quantified to ultimately produce gene expression matrices for each cell using Cell Ranger software. Detailed analyses, such as cell filtration, standardization, cell subgroup classification, and marker gene screening, were carried out using Seurat software (ver. 6). Low-quality cells were removed from the raw data using DoubletFinder software (v.2.0.2).

2.5. ScRNA-seq cell clustering

After cell filtration, all high-quality cells were used for further analyses, including cluster dimensionality reduction, cluster classification, and cluster visualization. All cell data was normalized and searched for highly variable genes using Seurat software with its built-in parameters. Downstream PCA calculation of the highly variable genes was performed to obtain the corresponding PCA values. Based on the PC values, cells with closer expression patterns were identified by utilizing the 'Find_Neighbors' function in Seurat software. The cells with closer expression patterns were classified into the same cell subgroup, referred to as a cluster.

To analyze the pseudo-time trajectory, all detected cells were reordered using Monocle 2 software (ver. 3.0). The gene counts were converted into the 'CellDataSet' using the 'importCDS' function in Monocle 2. Monocle 2 was used to pre-calculate parameters, which were used to trace differences during the pseudo-time process.

2.6. ScATAC-seq library construction

After cell nucleus lysis, the suspension was adjusted to the working concentration. The gel beads were dissolved to release the DNA fragments from the broken nucleus, and a thermal cycle produced single stranded DNA with a P5 connector and 10 × barcode. The labeled DNA was recovered, and the P7 junction was introduced through PCR to complete the library construction. After quality checking, sequencing

was performed on an Illumina sequencing platform.

2.7. ScATAC peak screening and cell calling

The 3' adaptor sequences were removed from the raw data using cutadapt tool. The clean sequences were aligning using the BWA-MEM program with default parameters. After aligned the clean reads onto genome, shifted the 5' end to the right by 4 bp and the 3' end to the left by 5 bp. The sequence of adjusted positions was defined as a fragment and stored in data files. By counting the 401 bp moving window nearby each site, a smooth contour of the transposition event is generated. Set a signal threshold based on a 1/5 advantage ratio to determine whether an area is a peak signal or noise. Merge peaks within a distance of 500 bp into one and store them in a BED file sorted by position.

The chromatin accessibility analysis showed the features of the targeted site were entered and defined by ENCODE software. The proportion of fragments located in the peak of each barcode was compared with those located in the 2000 bp flanks of each peak. A fixed value was subtracted from all barcode counts to establish a whitelist model. Two negative binomial distribution mixed models are fitted to obtain signal and noise.

2.8. ScRNA-seq and ScATAC-seq cell visualization and annotation

'CellRanger Aggr' tool was used to normalize the number of reads that can be mapped per cell. The well-identified leaf tissue-related marker genes in *Arabidopsis* were obtained from the PlantCellMarker website (Jin et al., 2022). The protein encoding sequences of *Arabidopsis* markers were treated as queries to search the *T. mairei* genome. The hits with the highest score were selected as marker genes to annotate the *T. mairei* leaf tissue.

For the scATAC data, the peak-cell matrix was imported into Cell Ranger software (ver. 6.1.0) with harmony algorithm corrected LSI embeddings were employed to generate UMAP dimensionality reduction. The features nearest the known markers were used to annotate the cell types. All the marker features were imported as an 'Active Feature List' into Loupe Browser to analyze the gene/feature expression pattern.

2.9. Whole-mount in situ hybridization

Similar leaf samples were collected from the *T. mairei* trees and fixed in FAA solution at 4℃. The samples were decolorized by tert-butyl alcohol, which was prepared with 95 % alcohol. The leaf samples were placed in a wax dissolved embedding box. After being cooled to 4℃, the wax block was sliced into 7 μm using a slicer and laid out on an exhibition table. The slide was dried in a 37℃ oven overnight and was put into xylene to remove all the wax. The glass slide was put in a series of different concentration of alcohol solutions. Then, the slide was put into 1 ug/mL of Proteinase K for digestion at 37 ℃ for 30 minutes, and washed it with 1 × PBS for 3 times.

The *in situ* hybridization was carried out using a DIG RNA labeling Kit (SP6/T7). In brief, the slides were soaked in preheated DIG-EASY-HYB buffer solution for 30 min, and soaked in another DIG-EASY-HYB buffer solution with a gene probe (10 pg/mL). Kept the slides at 42℃ for 18 h. The glass slides were washed twice at room temperature with 2 × SSC buffer and then twice at 65℃ with 0.5 × SSC buffer for 30 min. Rinse the slides with washing buffer for 5 min, incubate them in a closed solution for 2 h, and then place them at 37℃ in an antibody solution for 2 h. The resulting slides were washed twice with washing buffer for 15 min, and were transferred into a detection buffer for 5 min. The slides were put in the substrate chromogenic solution for incubation until the gene was detected.

2.10. Functional enrichment analysis of peak targets and TF motifs

Treating all peak-related genes as background, enrichment analysis of genes associated with peaks in each cluster was performed using adjust P value < 0.05 with the Benjamini–Hochberg method. Based on the detection of the TF motif, consider all peaks that match the specified TF. Then, the peak fragments that can be mapped to the peak barcode matrix were integrated, and the number of fragments matching the same TF motif in the barcode was calculated. Detect the enrichment of TFs by analyzing the distribution of a given TF ratio on barcodes for Z-score. In order to make it more stable and reduce the impact of outlier, we use the modified Z-Score calculated by the absolute deviation of median and standardized median (MAD) instead of the mean and the standard deviation. The 'add_Peak_Annotations' function was used to annotate the TF genes.

2.11. Promoter cis-element, subcellular localization and yeast two-hybrid analysis

The promoter sequences of the downstream genes were analyzed using the PlantCARE online tool. Several classic binding elements, including ERF binding sites (GCC-box), ARF binding sites (AuxRE), WRKY binding site (W-box), and MYB binding sites (MBE) were screened and visualized by TBtools.

The full-length sequences of *MYB48* (ctg782_gene.1), *bHLH1* (ctg183_gene.1), and *WRKY1* (ctg6511_gene.5) were cloned into the pH7FWG2.0 vector to produce GFP-constructs. The Clontech Two Hybrid System (Dalian, China) was used to analyze the interaction relationship between MYB48 and bHLH5 proteins. The full-length CDSs of *MYB48* and *bHLH5* were cloned into pGBKT7 and pGADT7 vectors, respectively. The completed constructs were con-transformed into AH109 yeast cells. The empty vectors were treated as the negative control and the P53 and T proteins were treated as the positive control. 3-AT is used to inhibit the self-activation effect. All primer sequences used for subcellular localization and Y2H experiments are shown in Table S6.

2.12. Electrophoretic mobility shift assay (EMSA), and dual-luciferase reporter assays

The full-length sequences of the above three TF genes were inserted into the pGEX4-T expression vector. The completed constructs were transformed into *Escherichia coli* cells (DE3) to yield His-tagged TF proteins. The expression of recombinant proteins was induced by 1 mM IPTG and purified using a Clontech His60 Ni Superflow Resin according to its instructions.

EMSA was performed using the Light Shift Chemiluminescent EMSA Kit (GS009, Beyotime, China) according to our previously described (Yu et al., 2020). Briefly, probes containing various *cis*-elements were derived from the open chromatin regions (OCRs) nearby *C4H* (ctg2905_gene.4), *COMT* (ctg15240_gene.2), *CHI* (ctg5664_gene.1), and *FLS* (ctg2147_gene.1) promoters and were labeled with fluorescent dye. Sequences without labels were applied as competitive probes and sequences with point mutations were applied as block probes. The full-length CDS sequences of *MYB48* and *bHLH5* genes were inserted into the pGreenII62-SK vector as effectors. The promoters of *C4H*, *COMT*, and *CHI* were inserted into the pGreenII0800-LUC vector as reporters. All primer and probe sequences used for prokaryotic expression, EMSA, and Dual-LUC are showed in Table S6.

3. Results

3.1. Identification and visualization of metabolites in the leaves of T. mairei

Our previous study published a MALDI-2 dataset to investigate the locations of various active ingredients in *T. mairei* leaves (Zhan et al., 2023). To visualize phenolic acids and flavonoids in the *T. mairei* leaves, the same dataset was deeply analyzed in the present study.

Segmentation analysis showed a colored *T. mairei* leaf map with a large number of feature points and separated the feature points into different patterns. Principal component analysis (PCA) grouped all the feature points into five PCs referring to five classic leaf tissues, including vein and bundle sheath cells (VB, PC1), epidermal cells (EP, PC2), palisade mesophyll cells (PM, PC3), and spongy mesophyll cells (SM, PC4/5, and Fig. 1**a**). A heatmap displayed the ion loading intensities referring to each feature point in five PCs (Fig. S1a and Table S1). All ion features were clustered into eight clusters. Detailly, Cluster I contains 361 PC1 highly accumulated metabolites, Cluster II contains 519 PC2 specifically accumulated metabolites, Cluster III and IV contain 535 (329 + 206) PC3 specifically accumulated metabolites, Cluster V contains 444 PC4 greatly accumulated metabolites, Cluster VI contains 723 PC5 specifically accumulated metabolites, Cluster VII contains 407 PC4/5 highly accumulated metabolites, and Cluster VIII contains 498 PC1/4 highly accumulated metabolites (Fig. S1b).

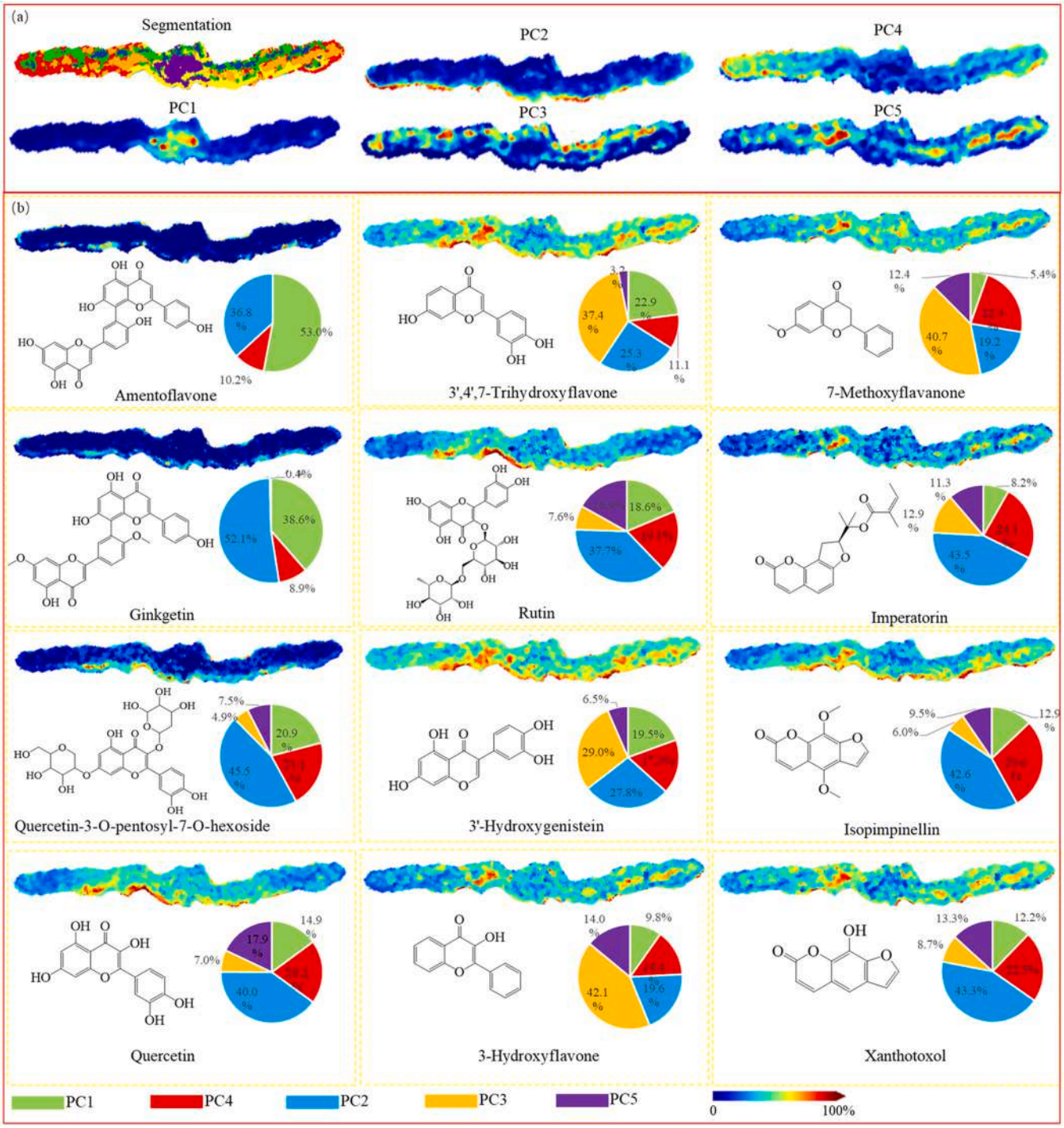

Fig. 1. Visualization of phenolic acid and flavonoid accumulation in *T. mairei* leaves. (a) The segmentation analysis provides a colored map with various data points. The picture showed five PCAs of all detected ion features in *T. mairei* leaves. Each PC referred to a specific metabolite accumulation pattern in *T. mairei* leaves. (b) The leaf tissue specific accumulation of 12 phenolic acids and flavonoids were determined by the MALDI-IMS analysis. The color scale ranges from 0 % to 100 %. Blue indicated low accumulation and red indicated high accumulation. The proportion of various phenolic acids and flavonoids in different leaf tissues was shown by pie charts. Green referring to PC1, blue referring to PC2, yellow referring to PC3, red referring to PC4, and purple referring to PC5.

3.2. Tissue-specific accumulation of phenolic acids and flavonoids in T. mairei leaves

In total, 12 phenolic acid and flavonoid-related metabolites were identified by MS imaging analysis (Table S2). Although there were different addcuts referring to one compound, the addcut that have clear boundary and a relative high content was selected for further visual analysis. Visual analysis showed that various phenolic acid and flavonoid-related metabolites preferentially accumulated in the EP cells. For examples, Amentoflavone, Ginkgetin, Quercetin-3-O-pentosyl-7-O-hexoside, Quercetin, Rutin, Imperatorin, and Xanthotoxol highly accumulated in the EP cells (PC2, > 30 %). Besides, 3′,4′,7-Trihydroxyflavone, 3-Hydroxyflavone, and 7-Methoxyflavanone accumulated significantly in the PM cells (PC3, > 30 %). Interestingly, the proportion of 3′-Hydroxygenistein in any PCs is less than 30 %, suggesting its relative even distribution in different cell types (Fig. 1**b**).

3.3. ScRNA-seq and ScATAC-seq of the T. mairei leaves

To integrate the previously published datasets, re-clustering was performed using Seurat Loupe software. For the ScRNA-seq dataset, all 8846 cells were re-grouped into 12 clusters (Fig. 2**a**). Cluster 1 contains

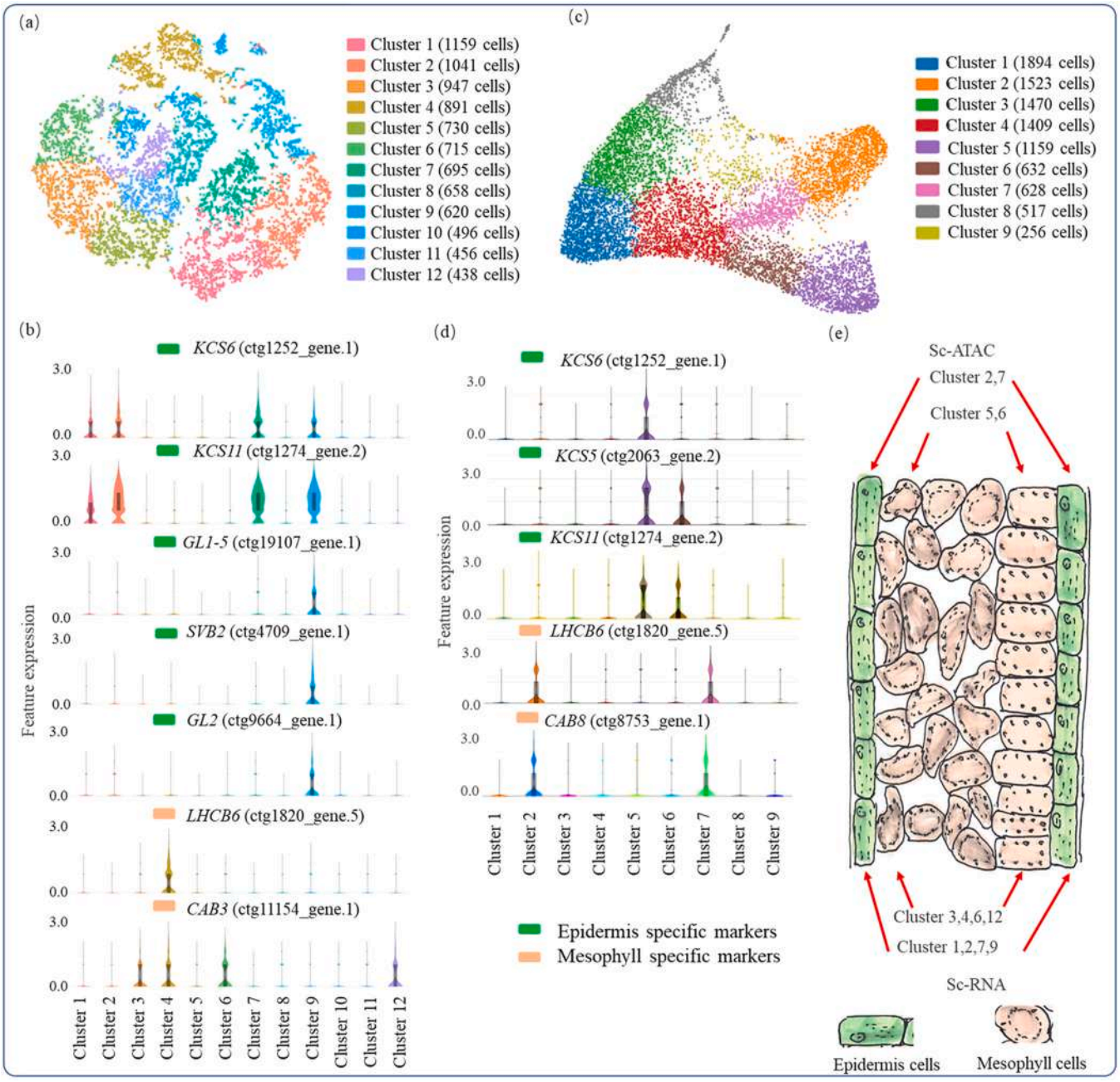

Fig. 2. ScRNA- and ScATAC-seq of the *T. mairei* leaves. (a) Visualization of 12 cell clusters in scRNA-seq dataset using *t*-SNE method. Dots indicated individual cells and colour indicated different cell clusters. (b) Violin plot analysis of expression patterns of selected seven cell type-specific markers for scRNA-seq data, including *KCS6* (ctg1252_gene.1), *KCS11* (ctg1274_gene.2), *GL1–5* (ctg19107_gene.1), *SVB2* (ctg4709_gene.1), *GL2* (ctg9664_gene.1), *LHCB6* (ctg1820_gene.5), and *CAB3* (ctg11154_gene.1). (c) Visualization of 9 cell clusters in scATAC-seq dataset using *t*-SNE method. Dots indicated individual cells and colour indicated different cell clusters. (d) The violin charts of cell type specific marker genes for scATAC-seq data, including *KCS6* (ctg1252_gene.1), *KCS5* (ctg2063_gene.2), *KCS11* (ctg1274_gene.2), *LHCB6* (ctg1820_gene.5), and *CAB8* (ctg8753_gene.1). (e) The sketch map of different clusters of ScRNA- and ScATAC-seq data. Red cells indicated mesophyll tissue and green cells indicated epidermis tissue.

the largest number of cells (1159 cells) and the Cluster 12 contains the smallest number of cells (438 cells). Five leaf epidermal cell-specifically expressed marker genes, such as *3-KETOACYL-COA SYNTHASE 6* (ctg1252_gene.1), *KCS11* (ctg1274_gene.2), *GLABRA1–5* (*GL1–5*, ctg19107_gene.1), *SMALLER with Variable Branches 2* (*SVB2*, ctg4709_gene.1), and *GL2* (ctg9664_gene.1), and two leaf mesophyll cell-specifically expressed marker genes, such as *LIGHT HARVESTING COMPLEX PHOTOSYSTEM II SUBUNIT 6* (*LHCB6*, ctg18492_gene.1), and *CHLOROPHYLL A/B BINDING PROTEIN 3* (*CAB3*, ctg11154_gene.1), were selected for cell type annotation (Jin et al., 2022). Violin plots showed that *KCS6/11* specifically expressed in Clusters 1/2/7/9, *GL1–5/SVB2/GL2* specifically expressed in Cluster 9, *LHCB6* specifically expressed in Cluster 4, and *CAB3* specifically expressed in Clusters 3/4/6/12 (Fig. 2**b**).

For the ScATAC-seq dataset, all 9488 cells were re-grouped into nine clusters (Fig. 2**c**). Cluster 1 contains the largest number of cells (1894 cells) and the Cluster 9 contains the smallest number of cells (256 cells). Cell type was annotated according to the expression patterns of the genes near the chromosome opening peaks. Here, three leaf epidermal cell specifically expressed genes (*KCS5*, *KCS6*, and *KCS11*) and two mesophyll cell-specifically expressed genes (*LHCB6* and *CAB8*) were applied to define the cell types of different clusters (Table S3). Violin plots showed that the peaks nearby the epidermis cell specifically expressed genes were enriched in Clusters 5/6, and the peaks near mesophyll specifically expressed genes were enriched in Clusters 2/7 (Fig. 2**e**).

For the ScRNA-seq data, most of the epidermal cells belonged to Clusters 1/2/7/9 and most of the mesophyll cells belonged to Clusters 3/4/6/9. For the ScATAC-seq data, the leaf epidermal cells were enriched in Clusters 5/6 and the leaf mesophyll cells were enriched in Clusters 2/7 (Fig. 2**e**).

3.4. Verification of the marker genes by in situ hybridization

To validate the two types of leaf cell specific expressed marker genes, seven marker genes, including five leaf epidermal cell markers and two leaf mesophyll cell markers, were selected for *in situ* hybridization. Antisense probes for *KCS6* (ctg1252_gene.1), *KCS11* (ctg1274_gene.2), *GL1–5* (ctg19107_gene.1), *SVB2* (ctg4709_gene.1), *GL2* (ctg9664_gene.1), *LHCB6* (ctg1820_gene.5), and *CAB3* (ctg11154_gene.1) were used for *in situ* hybridization experiment. Our data showed that the positive signals of *KCS11*, *KCS6*, *GL1–5*, *SVB2*, and *GL2* were greatly enriched in the leaf epidermal cells and the positive signals of *LHCB6* and *CAB3* were largely enriched in the leaf mesophyll cells (Fig. 3). The *in situ* hybridization experiment effectively verified the reliability of the selected marker genes.

3.5. Cell type specificity for the phenolic acid and flavonoid biosynthesis

For the biosynthesis of phenolic acid and flavonoid, scRNA-seq identified multiple core enzyme encoding genes (Fig. 4**a** and Table S4). The cell coverage analysis identified four major genes (> 5 % cell coverage) involved in phenolic acid and flavonoid biosynthesis, including ctg2905_gene.4 (C4H encoding gene, 14.84 % cell coverage), ctg5664_gene.1 (CHI encoding gene, 16.32 % cell coverage), ctg15240_gene.2 (COMT encoding gene, 8.7 % cell coverage), and ctg2147_gene.1 (FLS encoding gene, 6.01 % cell coverage, and Fig. 4**b**).

The four major genes involved in the phenolic acid and flavonoid biosynthesis pathway showed significantly cell type specific expression. Violin plot analysis showed that *C4H* (ctg2905_gene.4) specifically expressed in Clusters 1/2/9, *CHI* (ctg5664_gene.1) specifically expressed in Clusters 2/9, and *FLS* (ctg2147_gene.1) and *COMT* (ctg15240_gene.2) specifically expressed in Cluster 9 (Fig. 4**c**). Uniform Manifold Approximation and Projection (UMAP) visualization analysis confirmed that the four selected genes were enriched in Cluster1/2/9 (Fig. 4**d**). According to the annotation information, the four major genes significantly expressed in the EP cells.

The successive differentiation trajectory analysis identified two branch points, which separated all Clusters 1/2/9 cells into five stages. Stage 1 was considered to be the early stage of epidermis development and stage 5 was considered to be the late stage of epidermis development, respectively (Fig. 4**e and f**). Most of the phenolic acid and flavonoid metabolism-related genes highly expressed in the early and middle stages of epidermis development (Fig. 4**g**).

3.6. Chromatin accessibility of the phenolic acid and flavonoid biosynthesis-related genes in the T. mairei leaves

To reveal the chromatin accessibility, the phenolic acid and flavonoid biosynthesis-related promoter sequences and coding sequences were obtained (Xiong et al., 2021). According to the ScATAC-seq data, several OCRs nearby the target genes were detected (Table S5). Within the range of 400 kb around the target genes, several target genes processed more than three OCRs, such *C3H* (ctg11289_gene.1, three OCRs), *COMT* (ctg2470_gene.7, four OCRs), *4CL* (ctg13514_gene.2, three OCRs), *CHS* (ctg13079_gene.1, five OCRs), *CHS* (ctg13079_gene.8, three OCRs), *CHS* (ctg3517_gene.14, three OCRs), *FLS* (ctg2147_gene.1, three OCRs), *FLS* (ctg2413_gene.16, three OCRs). Interestingly, most of the OCRs were localized within the core region (-50 kb ~ 50 kb) of target genes (Fig. 5**b)**. Distribution of various important *cis*-elements is shown in Fig. 5**c**.

Although most of the OCRs could be detected in both of the epidermal and mesophyll cells, some OCRs were significantly enriched in a specific cell type. Our study, evaluated the differences in enrichment of OCRs between leaf mesophyll and epidermal cells (Fig. 5**d**). Importantly, the OCRs surrounding the four phenolic acid and flavonoid biosynthesis-related genes were enriched in Cluster 5/6, implying active transcription regulation occurring in the EP cells.

3.7. Predication of TFs involved in the regulation of phenolic acid and flavonoid biosynthesis

Although a number of phenolic acid and flavonoid biosynthesis-related TFs were identified in various plants, knowledge of their transcriptional regulation in *Taxus* is limited. According to the ScRNA-seq data, several cell-type specifically expressed TFs were identified, including six mesophyll cell specifically expressed TFs, nine epidermal cell specifically expressed TFs, and four mesophyll/epidermal cell commonly expressed TFs (Fig. 6**a-c**). The epidermal cell specifically expressed TFs included four *MYB* genes (ctg73_gene.1, ctg17994_gene.1, ctg5392_gene.1, and ctg10439_gene.1), one *bHLH* gene (ctg18566_gene.1), and one *WRKY* gene (ctg4847_gene.1). The mesophyll cell specifically expressed TFs included one *WRKY* gene (ctg7850_gene.3), two *MYB* genes (ctg340_gene.4 and ctg6731_gene.1), two *ERF* genes (ctg5587_gene.8 and ctg505_gene.1), and three *GT-2* genes (ctg2253_gene.1, ctg2396_gene.1, and ctg21853_gene.2). The mesophyll/epidermal cell commonly expressed TFs included two *bHLH* genes (ctg46_gene.21 and ctg183_gene.1) and two *MYB* genes (ctg4124_gene.1 and ctg872_gene.1).

According to two screening criteria, similar cell type specific expression pattern and matching binding *cis*-elements, a potential regulatory network of phenolic acid and flavonoids biosynthesis in *T. mairei* leaves was predicted (Fig. 6**d**). In detail, ten TFs, including six MYBs, one WKRY, three GT-2 s, and three bHLHs, were predicted to be the regulators of *C4H* (ctg2905_gene.4) gene; nine TFs, including six MYBs, and three bHLHs, were predicted to be the regulators of *COMT* (ctg15240_gene.2) gene; ten TFs, including three bHLHs, one WKRY, and six MYBs, were predicted to be the regulators of *CHI* (ctg5664_gene.1) gene; seven TFs, including six MYBs, and one WRKY, were predicted to be the regulators of *FLS* (ctg2147_gene.1) gene.

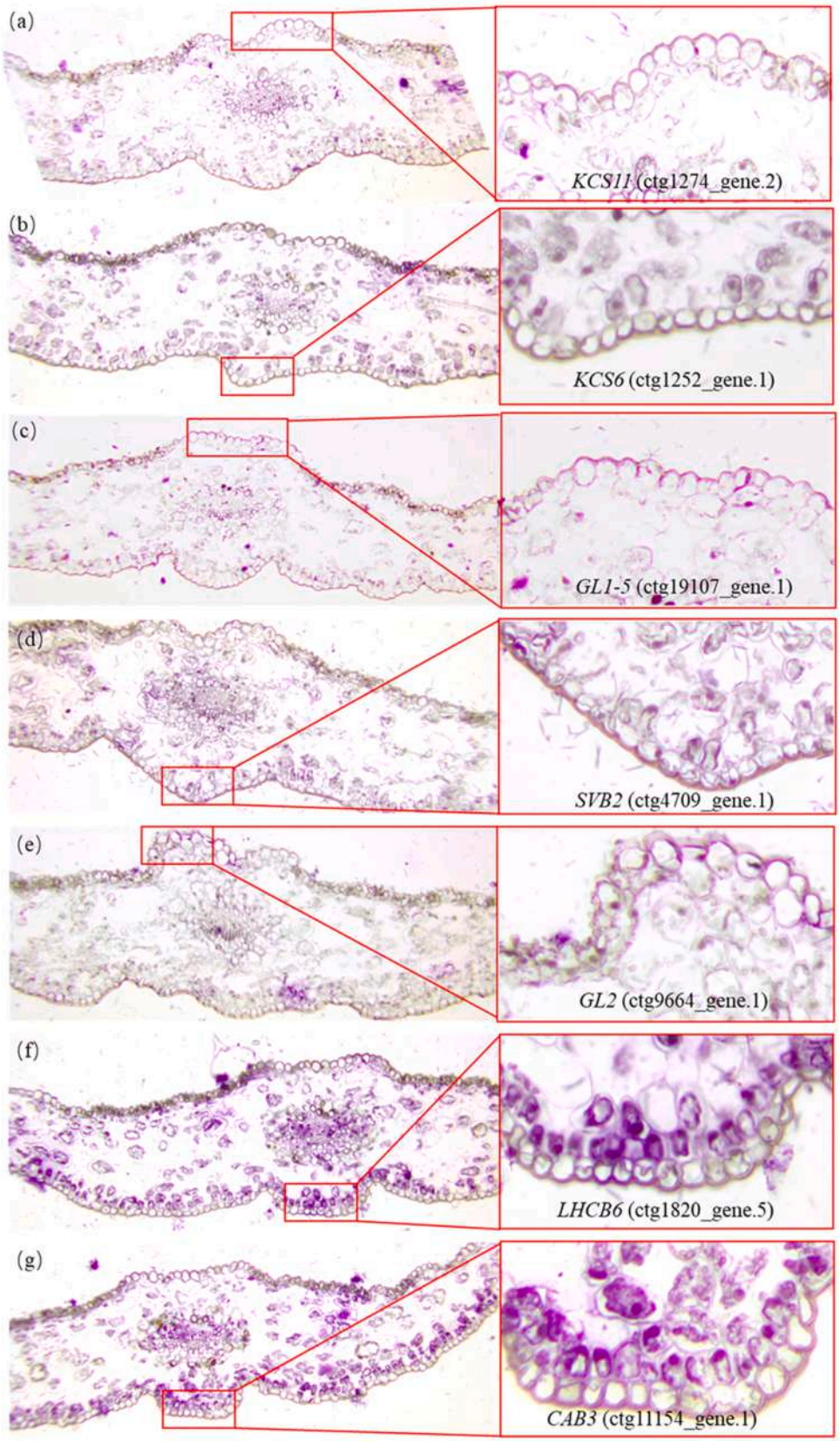

Fig. 3. Verification of the marker genes by *in situ* hybridization. *In situ* hybridizations of *T. mairei* leaves with distinct gene-specific probes to illustrate their tissue specific expression patterns. Localization of (a) *KCS11* (ctg1274_gene.2), (b) *KCS6* (ctg1252_gene.1), (c) *GL1–5* (ctg19107_gene.1), (d) *SVB2* (ctg4709_gene.1), (e) *GL2* (ctg9664_gene.1), (f) *LHCB6* (ctg1820_gene.5), and (g) *CAB3* (ctg11154_gene.1) in leaf cross-sections.

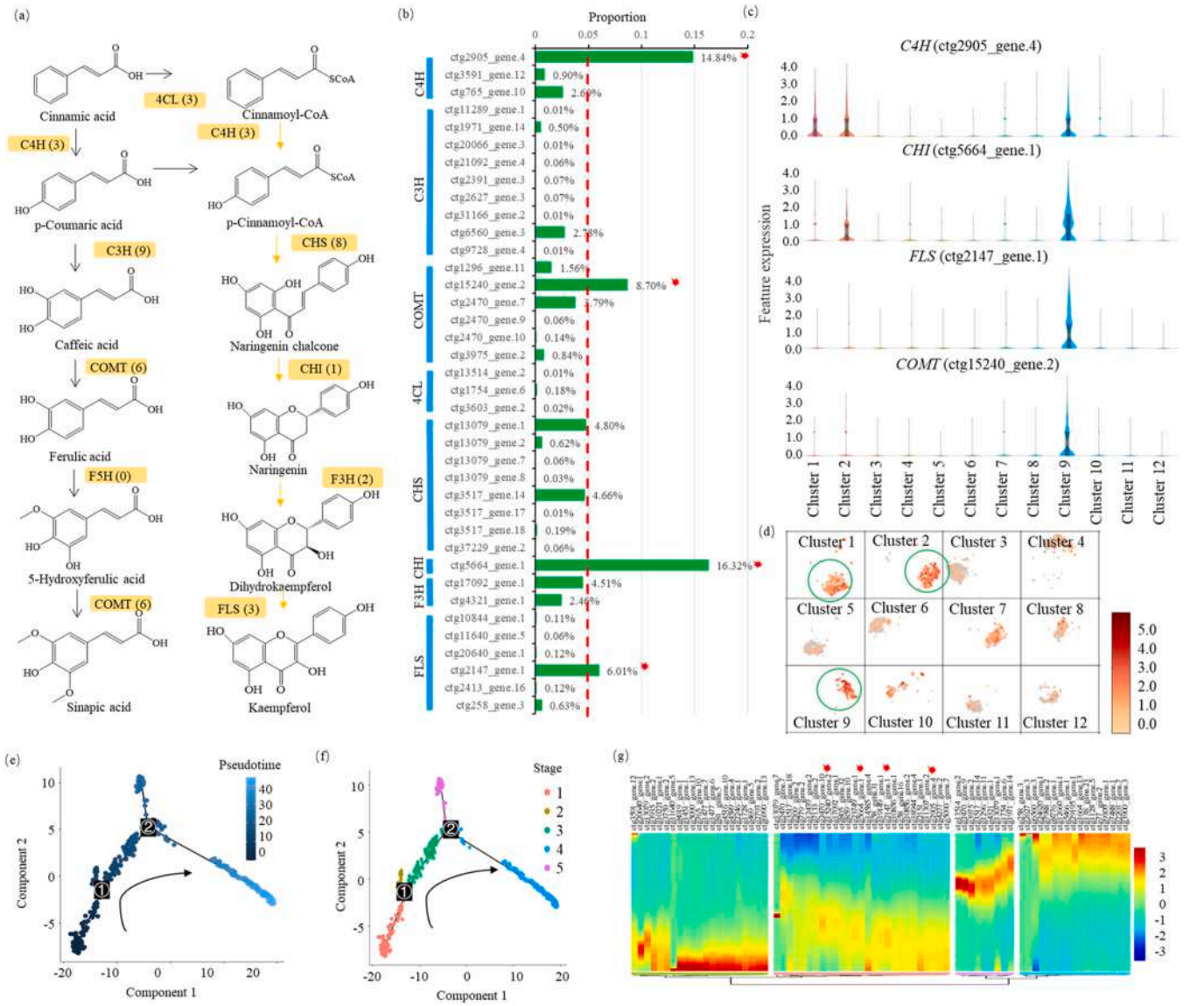

Fig. 4. Cell type specificity of flavonoid metabolisms. (a) Overview of the flavonoid metabolism pathways, including three *C4H* genes, three *4CL* genes, nine *C3H* genes, six *COMT* genes, one *CHI* gene, two *F3H* genes, and three *FLS* genes. The number indicated the encoding genes. (b) The cell coverage of flavonoid metabolism pathway-related genes. Each key enzyme was encoding by at least one gene. Red stars indicated the encoding genes with high cell coverage. Red dashed line indicated 5 % cell coverage. (c) Violin plot analysis of expression patterns of the flavonoid metabolism-related genes, including *C4H* (ctg2905_gene.4), *CHI* (ctg5664_gene1), *FLS* (ctg2147_gene.1), and *COMT* (ctg2147_gene.1). (d) UMAP visualization of expression patterns of flavonoid metabolism pathway-related genes in *T. mairei* leaves. The color scale ranges from 0 to 5.0. (e) Successive differentiation trajectory during the leaf epidermis tissue development process. (f) Successive differentiation trajectory analysis divided the cells into five stages. (g) Heatmap showing expression levels of flavonoid biosynthesis-related genes under leaf epidermis tissue development process. The heatmap scale ranges from +2 to −2 on a log2 scale.

3.8. Identification of two typic cell type specifically expressed TF genes

Two cell type specifically expressed TF genes, MYB48 (ctg872_gene.1), and bHLH5 (ctg183_gene.1), were selected to analyze the cell-specific transcriptional regulation of phenolic acid and flavonoid biosynthesis. The CDSs of *MYB48*, and *bHLH5* were cloned according to the *T. mairei* reference genome. MYB48 is a protein with 379 amino acid residues and bHLH5 is a protein with 521 amino acid residues. MYB48 is typic R2R3-MYB domain and bHLH5 contains a basic helix-loop-helix type domain (Fig. S2 and S3). Subcellular localization analysis showed that enhanced GFP-tagged MYB48 and bHLH5 proteins localized in the nucleus **(**Fig. 7**a)**, Yeast toxicity experiments showed that AD-MYB48 has no toxicity in AH109, while BD-bHLH48 has only weak toxicity in AH109, which does not affect subsequent experiments (Fig. S4). Self-activation analysis indicated that 2.5 mM 3-Amino-1,2,4-triazole can inhibit the self-activation activities of both constructs (Fig. S5). The Y2H results indicated the interaction between MYB48 and bHLH5 (Fig. 7**b**).

3.9. Verifying cell type specific regulation in the T. mairei leaves

To assess the role of MYB48-bHLH5 complex in the transcriptional regulation of phenolic acid and flavonoids biosynthesis, EMSA was carried out to analyze the binding activities to their potential targets. The protomer sequences of *C4H* (ctg2905_gene.4), *COMT* (ctg15240_gene.2) and *CHI* (ctg5664_gene.1) were cloned by PCR amplification. Prokaryotic expression and GST-tagged affinity purification produced the purified MYB48 and bHLH5 proteins (Fig. S6). EMSA results showed that both of MYB48 and bHLH5 directly targeted to the referring elements in the promoters of *C4H* (ctg2905_gene.4), *COMT*

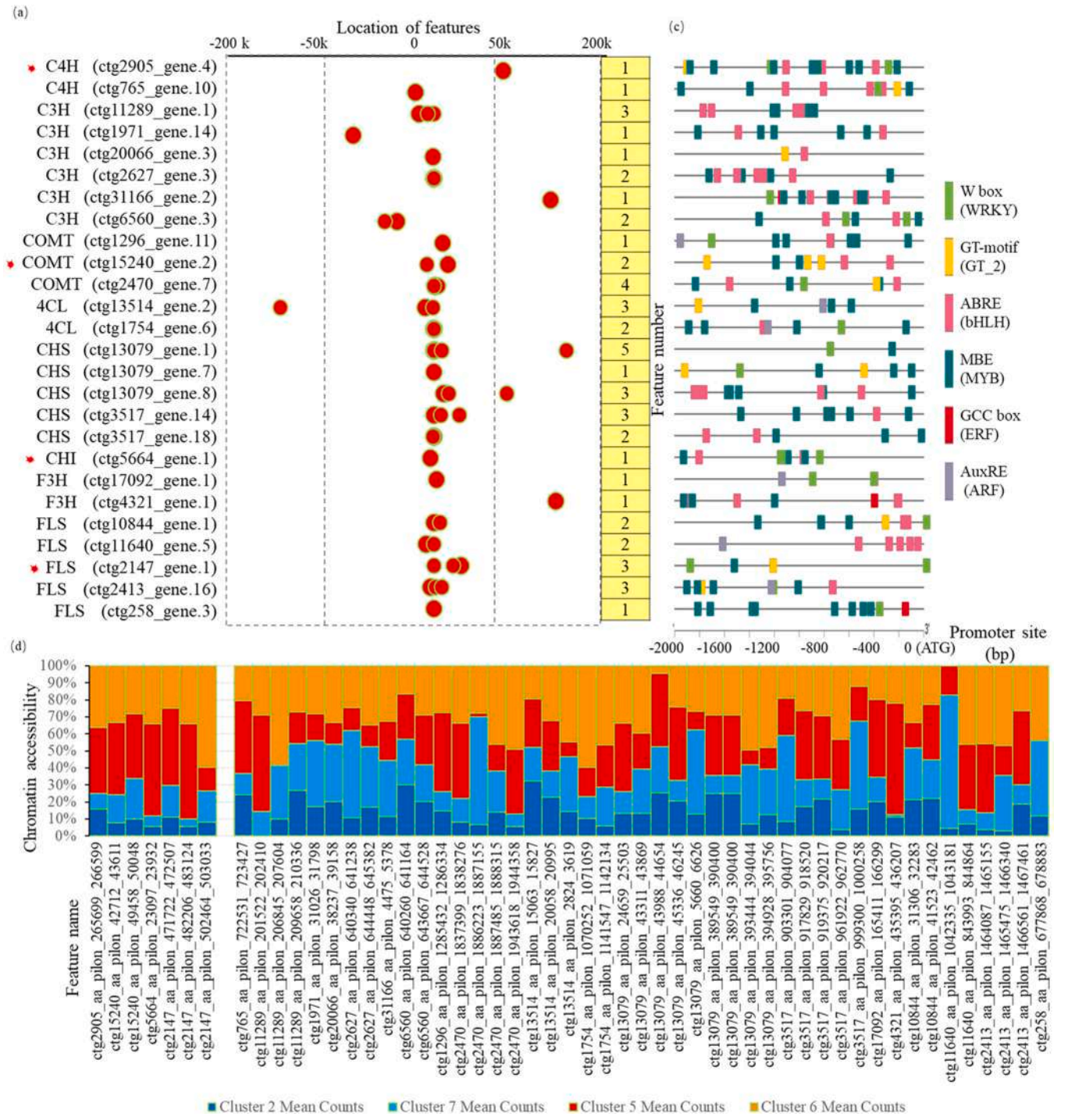

Fig. 5. Chromatin accessibility of the flavonoid metabolism-related genes in *T. mairei* leaves. (a) The flavonoid metabolism-related genes. Red stars indicated the encoding genes with high cell coverage. (b) The red dots indicated the location of OCRs in different flavonoid metabolism-related genes. The number in yellow box indicated the number of OCRs in different flavonoid metabolism-related genes. (c) The distribution of various *cis*-elements in OCRs nearby the flavonoid biosynthesis-related genes. Light green boxes indicated W box, yellow boxes indicated GT-motif, purple boxes indicated ABRE elements, deep green boxes indicated MBE, red boxes indicated GCC-boxes, and grey boxes indicated AuxREs. (a) The biased chromatin accessibility between leaf mesophyll cells and epidermis cells. The read counts of each TF motif in cluster 2&7 (leaf mesophyll cells) and cluster 5&6 (leaf epidermis cells) was used to calculated the percentage of chromatin accessibility.

(ctg15240_gene.2) and *CHI* (ctg5664_gene.1) genes. Addition of excess cold competitive probe effectively weakened the binding of the His-tagged fusion proteins to the DNAs, but addition of excess mutated probe did not affect their binding (Fig. 7**c-h**).

Moreover, *in vivo* dual-LUC reporter assay was used to investigate the transcriptional activities in tobacco leaves (Fig. 7**i**). Dual-LUC results showed that both of MYB48 and bHLH5 effectively enhanced the expression levels of *C4H* (ctg2905_gene.4), *COMT* (ctg15240_gene.2) and *CHI* (ctg5664_gene.1) promoter-driven LUC reporters, indicated that MYB48 and bHLH5 are transcriptional activators of the phenolic acid and flavonoids biosynthesis pathway. Furthermore, co-expression of MYB48 and bHLH5 up-regulated the LUC activity to a higher level

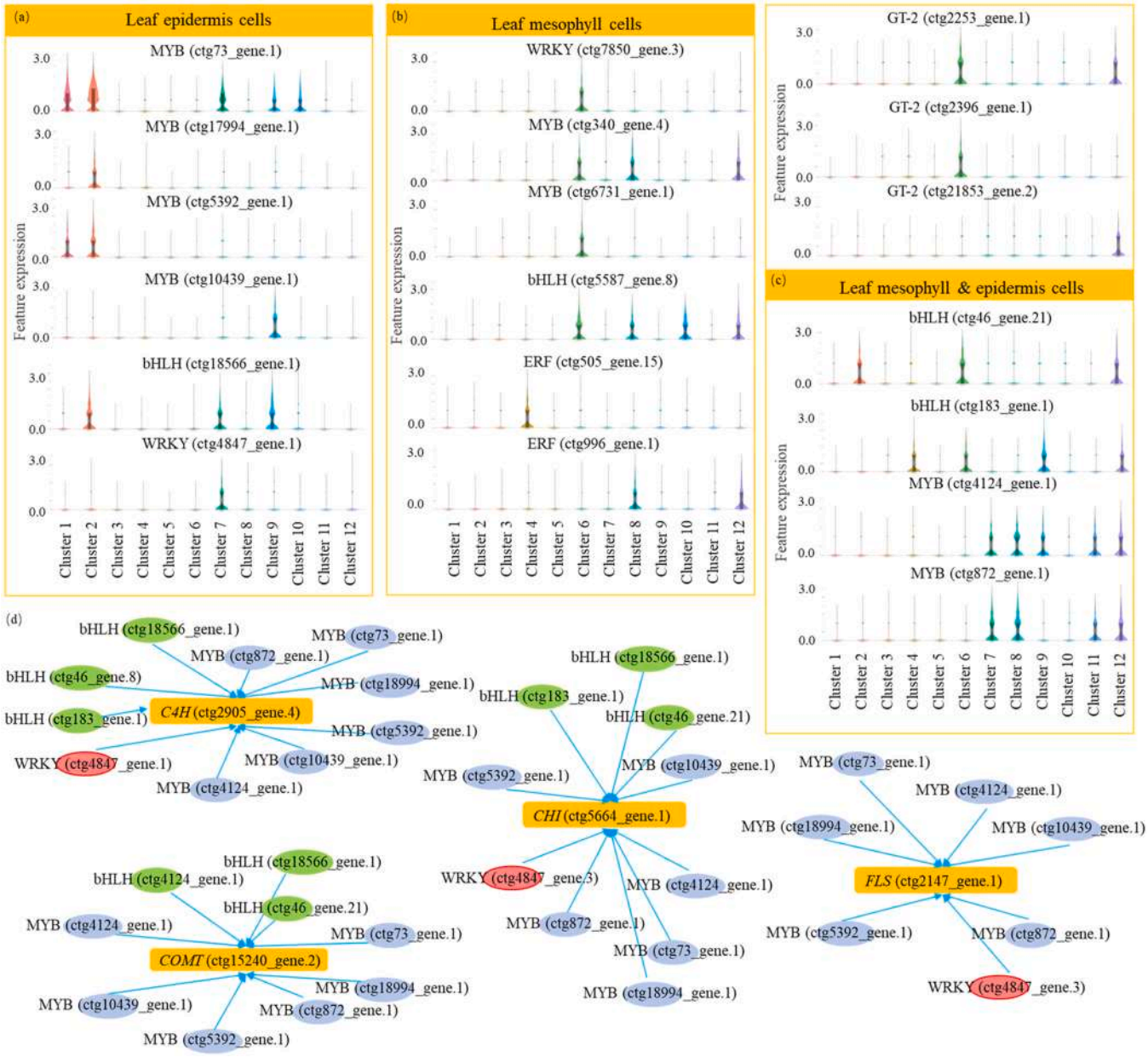

Fig. 6. Predication of TFs involved in the regulation of flavonoid biosynthesis. The Violin charts of cell type specific expression TF encoding genes, including six leaf mesophyll cell-specifically expressed TFs (a), nine leaf epidermis cell-specifically expressed TFs (b), and four leaf mesophyll cell/epidermis cell commonly expressed TFs (c). (d) Prediction of transcription regulation networks according to the cell type specific expression pattern. Red background indicated WKRY family members, green background indicated bHLH family members, and blue background indicated MYB family members.

than any single protein (Fig. 7**j**).

4. Discussion

Chemical localization is controlled by various genetic and epigenetic regulation factors. Nowadays, extremely high attentions have been focused on the effects of phenolic acid and flavonoids on human health care (Szopa et al., 2019; Wen et al., 2021). However, the tissue-specific distribution of phenolic acids and flavonoids in *T. mairei* leaves is largely unknown.

MALDI-MSI is a valuable tool for visualizing phenolic acids and flavonoids in various plant organs (Zhan et al., 2024). In *Cannabis sativa* leaves, the detectable flavonoids localize throughout the leaflet (Lorensen et al., 2023). In *Helianthus annuus* leaves, several flavonoids were detected in the trichomes by MS imaging analysis (Brentan Silva et al., 2017). In *Gliricidia sepium*, MALDI MSI analysis was performed on its leaf surface to detect the spatial distribution of Kaempferol (Hertel Pereira et al., 2022). To date, MSI mainly been used to scan the surface of plant leaves. Our results can reveal the spatial distribution of phenolic acids and flavonoids within the leaf internal structure by slice scanning (Fig. 1**a**). Phenolic acids and flavonoids are important plant protective compounds that help plants adapt to environmental stresses (Casanova et al., 2020). Anthocyanins are reported to be generally accumulated in the leaf epidermal cells (Ahn et al., 2022). Interestingly, most of the detectable flavonoids preferentially accumulated in the leaf epidermal cells of *T. mairei*, which is consistent with previous reports (Ahn et al., 2022). In *Taxus* trees, Amentoflavone and Ginkgetin are important members of phenolic compounds with anti-fungal and anti-cancer activities (Krauze-Baranowska and Wiwart, 2003). In *T. mairei*, MSI data showed strong signals referring to Amentoflavone and Ginkgetin were

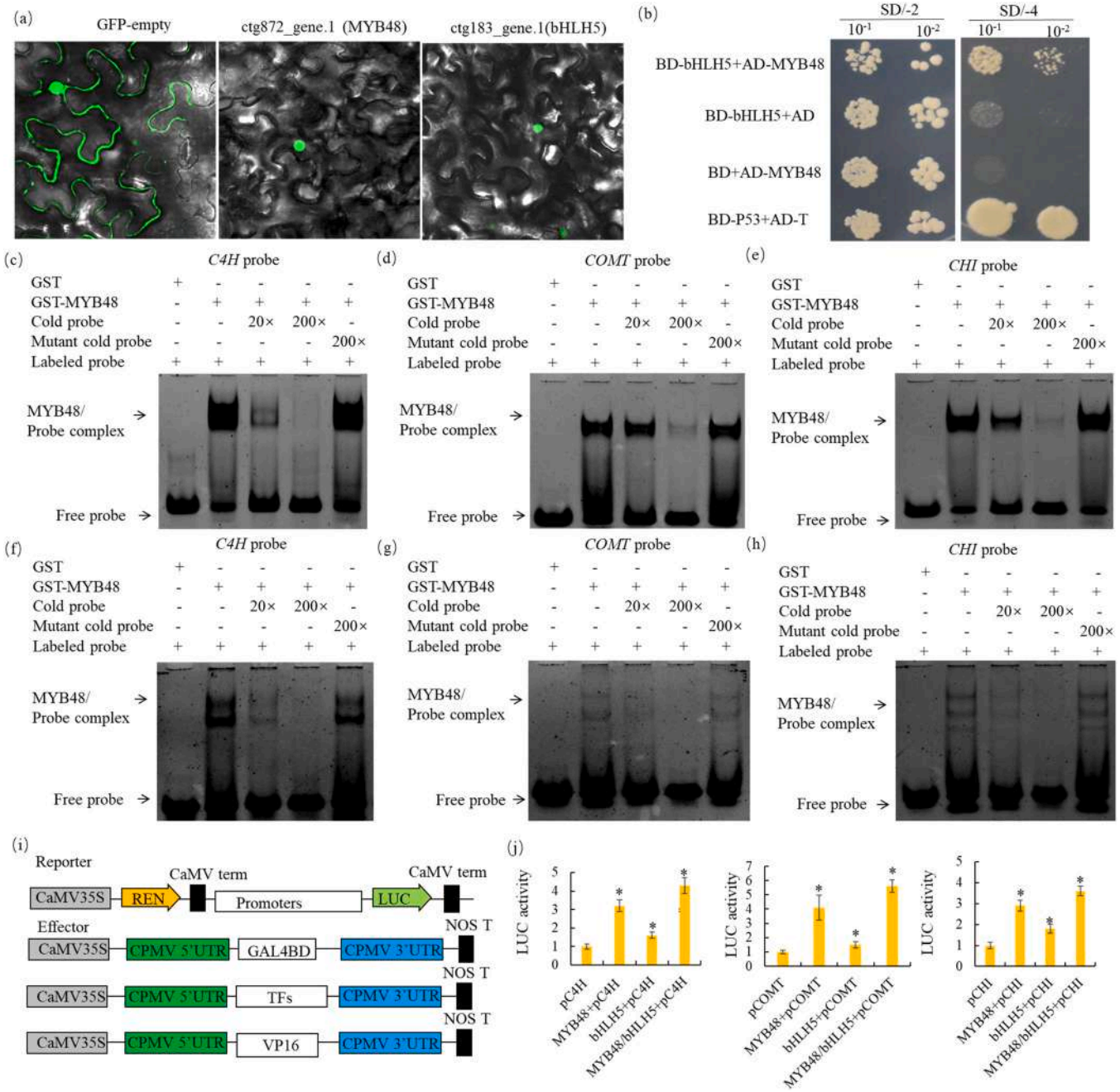

Fig. 7. Verification of the transcriptional activity of two typical TFs to their potential downstream target genes. (a) Subcellular localization of MYB48 and bHLH5. (b) Y2H assay analysis of the interaction between MYB48 and bHLH5. Yeast cells transformed with AD-bHLH5/BD and BD-MYB48 constructs were used as negative controls, while those transformed with the BD-P53+AD-T construct were used as positive control. Design of probes surrounding the elements in several phenolic acid and flavonoid biosynthesis-related promoter regions. GST only or GST-MYB48/GST-bHLH5 fusion protein were incubated with probes containing elements derived from the promoters of *C4H* (c, f), *COMT* (d, g), and *CHI* (e, h), respectively. The symbols '−' and '+' represent absence and presence, respectively. '20×' and '200×' indicate the fold increases in probe amounts. (i) Overview of the constructs for dual-luciferase reporter assays. The promoter fragments were ligated with the pGreenII 0800-LUC vector to produce the reporters. The effectors were produced by inserting the TF genes into the pGreenII-62-SK vector. (j) Effects of MYB48/bHLH5 on the luciferase activities of the C4H, COMT, and CHI reporters. Relative LUC activity represents the activity ratio of the firefly luciferase to Renilla luciferase. Each value is the mean±standard deviation of three biological repeats. '*' indicated $P < 0.05$.

detected in the EP cells (PC2), suggesting their essential roles in the protection of plant leaves against adverse environments at the front line (Fig. 1**b**).

To investigate the regulatory differences in phenolic acid and flavonoid biosynthesis between mesophyll and epidermal cells in *T. mairei* leaves, an integrative analysis of ScRNA-seq and ScATAC-seq was performed (Fig. 2). Cell-type specifically expressed marker genes are crucial for successful annotation of single cell clusters. To date, most marker genes of woody plant were derived from model plants by homology analysis (Liu et al., 2022). Lack of cell type specifically expressed marker genes, which must be verified by experiments, limits the annotation of cell clusters from woody plant ScRNA-seq datasets (Yu et al., 2023; Zhan et al., 2023). In our study, *in situ* hybridization verified several marker genes referring to leaf epidermal and mesophyll cell types, providing reliable marker genes for woody plant research (Fig. 3). Marker genes used for ScRNA-seq annotation are often inaccurate. In

future research, it is necessary to verify the accuracy of the marker genes through in situ hybridization.

In plants, secondary metabolism is known to exhibit cell type specificity. In *Arabidopsis* leaves, ScRNA-seq showed that certain genes involved in the phenylpropanoid biosynthesis pathway exhibit an expression pattern enriched in the palisade layer (Procko et al., 2022). Another ScRNA-seq analysis supposed spatial separation of vasculature tissue during glucosinolate biosynthesis, storage, and breakdown (Tenorio Berrio et al., 2022). In *Heliotropium indicum*, the genes encoding of Homospermidine Synthase, which is involved in pyrrolizidine alkaloid biosynthesis, express only in the lower leaf epidermis cells (Sievert et al., 2015). Flavonoids, important phenolic secondary metabolites, consist of a 15-carbon skeleton, including two benzene rings (A and B) and one heterocyclic ring (C), existing in most plants (Vue et al., 2016). Although each important flavonoid biosynthesis-related enzyme has multiple coding genes in the *T. mairei* genome, their expression levels vary greatly (Fig. 4**a**). Our study identified four major genes, *C4H* (ctg2905_gene.4), *COMT* (ctg15240_gene.2), *CHI* (ctg5664_gene.1), and *FLS* (ctg2147_gene. 1), with high cell coverage (> 5 %), speculating that they might play a key role in phenolic acid and flavonoid biosynthesis (Fig. 4**b**). Interestingly, all four selected genes specifically expressed in the epidermal cell-related clusters (1, 2, and 9), suggesting that the epidermal cells are the major site for phenolic acid and flavonoid biosynthesis (Fig. 4**c**). Our ScRNA-seq data further confirmed the reliability of the MALDI-2 data.

Gene expression is programmed and controlled by various epigenetic regulatory factors (Springer and Schmitz, 2017). Recent studies have revealed several epigenetic regulatory events during the taxol biosynthesis process in *Taxus* cells. DNA methylation of the *GGPPS*, *TS*, *DBTNBT*, and *BAPT* promoters affects taxoid biosynthesis in long-term cultured *Taxus* cells (Escrich et al., 2022; Fu et al., 2012; Sanchez-Munoz et al., 2018). Chromatin accessibility in the *T. mairei* leaf was investigated at the singe-cell level, indicating that the chromatin accessibility greatly varied among different leaf tissues. Open chromatin in the promoter, enhancer, and gene ontology can be detected on a genome-wide scale by newly developed ScATAC-seq technology (Buenrostro et al., 2015). OCRs frequently contain various genomic elements for TF recognition and gene expression regulation (Xiong et al., 2019). In *T. mairei* leaves, seven OCRs surrounding the four major genes involved in phenolic acid and flavonoid biosynthesis were enriched in the epidermal cells, giving a reasonable explanation for their cell specific expression (Fig. 5**a**).

Intergrated scRNA-seq and scATAC analysis provided a novel perspective on screening cell-type specifically expressed TFs (Lu et al., 2023). Our data identified six leaf epidermal cell-specific and nine leaf mesophyll cell-specific TF coding genes, predicting a regulation network of phenolic acid and flavonoid biosynthesis at the spatial level (Fig. 6). To date, the function of the MBW complex in flavonoid biosynthesis has been functional identified in plants (Dixon et al., 2013; Zhao et al., 2022). Our results identified four epidermal cell specifically expressed *MYB* genes and one epidermal cell specifically expressed *bHLH* gene. So far, there have been no reports of the MYB-bHLH complex participating in the regulation of phenolic acid biosynthesis (Zhao et al., 2022). To expand the understanding of the biological functions of the MYB-bHLH complex, we will continue to screen for tissue-specifically expressed MYB and bHLH family members in the future. EMSA and dual-LUC experiments confirmed that *C4H*, *COMT*, and *CHI* were downstream targets of MYB48 and bHLH5, suggesting a potential function of MYB-bHLH complex in the regulation of phenolic acid and flavonoid biosynthesis (Fig. 7**c-j**).

In conclusion, MALDI-2 MS imaging revealed the precise spatial locations of several phenolic acid and flavonoid-related metabolites, thereby conforming leaf cell type-specifically accumulated secondary metabolites in *T. mairei* leaves. Integrated ScRNA-seq and ScATAC-seq analysis revealed the genetic regulatory mechanisms underlying phenolic acid and flavonoid biosynthesis in different leaf tissues. Our study identified four phenolic acid and flavonoid biosynthesis-related genes, *C4H* (ctg2905_gene.4), *COMT* (ctg15240_gene.2), *CHI* (ctg5664_gene.1), and *FLS* (ctg2147_gene. 1), which specifically expressed in the epidermal cells. Furthermore, the roles of two cell-specifically expressed TFs (*MYB48* and *bHLH5*) in phenolic acid and flavonoid biosynthesis were predicted. Our study provided a valuable resource outlining the cell type specific regulation of phenolic acid and flavonoid biosynthesis.

CRediT authorship contribution statement

Xiaori Zhan: Resources, Data curation, Conceptualization. **Xueshuang Liang:** Software, Methodology, Data curation. **Lilin Wang:** Funding acquisition, Writing – original draft. **Yanjun Yang:** Conceptualization, Writing – original draft, Writing – review & editing. **Huizhong Wang:** Validation, Conceptualization. **Chenjia Shen:** Writing – review & editing, Writing – original draft, Funding acquisition. **Ruoyun Ma:** Software, Formal analysis. **Yue Zang:** Software, Formal analysis, Data curation. **Wanting Lin:** Resources, Data curation, Conceptualization.

Declaration of Competing Interest

The authors have no relevant financial or non-financial interests to disclose.

Data Availability

scRNA-seq data are available at the NCBI database (BioProject: SRX22599982)

Acknowledgements

This work was funded by National Natural Science Foundation of China (32271905 and 32270382); Zhejiang Provincial Natural Science Foundation of China under Grant (No. LY23C160001, LY18C050005, LY19C150005, and LY19C020003); Postgraduate Research Innovation Project, Hangzhou Normal University (2022HSDYJSKY217); Postgraduate Research Innovation Project, Hangzhou Normal University (4175B20500001); Science and Technology Innovation Activity Plan (New Talent Plan) of Zhejiang Province (2023R445029). We are grateful to LC Sciences company (Hangzhou, China) and Shanghai Applied Protein Technology Co., Ltd (Shanghai, China) for transcriptomic and metabolomic analysis, respectively. We are also grateful to Professor Jianbin Yan (Agricultural Genomics Institute at Shenzhen, Chinese Academy of Agricultural Sciences, Shenzhen, China) for the genome data of *T. mairei*.

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.indcrop.2024.118975.

References

Ahn, J.Y., Kim, J., Yang, J.Y., Lee, H.J., Kim, S., Cho, K.S., Lee, S.H., Kim, J.H., Lee, T.H., Hur, Y., Shim, D., 2022. Comparative transcriptome analysis between two potato cultivars in tuber induction to reveal associated genes with anthocyanin accumulation. Int J. Mol. Sci. 23, 3681.

Brentan Silva, D., Aschenbrenner, A.K., Lopes, N.P., Spring, O., 2017. Direct analyses of secondary metabolites by Mass Spectrometry Imaging (MSI) from sunflower (*Helianthus annuus* L.) trichomes. Molecules 22, 774.

Buenrostro, J.D., Wu, B., Litzenburger, U.M., Ruff, D., Gonzales, M.L., Snyder, M.P., Chang, H.Y., Greenleaf, W.J., 2015. Single-cell chromatin accessibility reveals principles of regulatory variation. Nature 523, 486–490.

Cai, D., Jin, J., Bi, H., Zhong, G., Zhou, M., Guo, J., Cai, Y., Liang, M., Gu, Q., Hu, Z., Lai, Y., Dai, Z., Li, L., Chen, Y., Gao, H., Huang, M., 2022. Paclitaxel-containing extract exerts anti-cancer activity through oral administration in A549-xenografted BALB/C nude mice: synergistic effect between paclitaxel and flavonoids or lignoids. Evid. Based Complement Altern. Med 2022, 3648175.

Casanova, J.M., Dos Santos Nascimento, L.B., Casanova, L.M., Leal-Costa, M.V., Costa, S. S., Tavares, E.S., 2020. Differential distribution of flavonoids and phenolic acids in leaves of Kalanchoe delagoensis Ecklon & Zeyher (*Crassulaceae*). Microsc. Micro 26, 1061–1068.

Chen, R., Cao, Y., Wang, W., Li, Y., Wang, D., Wang, S., Cao, X., 2021. Transcription factor SmSPL7 promotes anthocyanin accumulation and negatively regulates phenolic acid biosynthesis in *Salvia miltiorrhiza*. Plant Sci. 310, 110993.

Ding, M., He, Y., Zhang, K., Li, J., Shi, Y., Zhao, M., Meng, Y., Georgiev, M.I., Zhou, M., 2022. JA-induced FtBPM3 accumulation promotes FtERF-EAR3 degradation and rutin biosynthesis in *Tartary buckwheat*. Plant J. 111, 323–334.

Dixon, R.A., Liu, C., Jun, J.H., 2013. Metabolic engineering of anthocyanins and condensed tannins in plants. Curr. Opin. Biotechnol. 24, 329–335.

Elansary, H.O., Szopa, A., Kubica, P., F, A.A.-M., Mahmoud, E.A., Zin El-Abedin, T.K.A., M, A.M., Ekiert, H., 2019. Phenolic compounds of *Catalpa speciosa*, *Taxus cuspidate*, and *Magnolia acuminata* have antioxidant and anticancer activity. Molecules 24, 412.

Escrich, A., Cusido, R.M., Bonfill, M., Palazon, J., Sanchez-Munoz, R., Moyano, E., 2022. The epigenetic regulation in plant specialized metabolism: DNA methylation limits paclitaxel *in vitro* biotechnological production. Front Plant Sci. 13, 899444.

Fang, S., Qu, Q., Zheng, Y., Zhong, H., Shan, C., Wang, F., Li, C., Peng, G., 2016. Structural characterization and identification of flavonoid aglycones in three *Glycyrrhiza* species by liquid chromatography with photodiode array detection and quadrupole time-of-flight mass spectrometry. J. Sep Sci. 39, 2068–2078.

Fu, C., Li, L., Wu, W., Li, M., Yu, X., Yu, L., 2012. Assessment of genetic and epigenetic variation during long-term Taxus cell culture. Plant Cell Rep. 31, 1321–1331.

Gai, Q.Y., Jiao, J., Wang, X., Liu, J., Fu, Y.J., Lu, Y., Wang, Z.Y., Xu, X.J., 2020. Simultaneous determination of taxoids and flavonoids in twigs and leaves of three *Taxus* species by UHPLC-MS/MS. J. Pharm. Biomed. 189, 113456.

Gauchan, D.P., Kandel, P., Tuladhar, A., Acharya, A., Kadel, U., Baral, A., Shahi, A.B., Garcia-Gil, M.R., 2020. Evaluation of antimicrobial, antioxidant and cytotoxic properties of bioactive compounds produced from endophytic fungi of Himalayan yew (*Taxus wallichiana*) in Nepal. F1000Res 9, 379.

Hertel Pereira, A.C., Auer, A.C., Biedel, L., de Almeida, C.M., Romao, W., Endringer, D. C., 2022. Analysis of *Gliricidia sepium* Leaves by MALDI Mass Spectrometry Imaging. J. Am. Soc. Mass Spectrom. 33, 573–583.

Hurtova, M., Biedermann, D., Osifova, Z., Cvacka, J., Valentova, K., Kren, V., 2022. Preparation of synthetic and natural derivatives of flavonoids using Suzuki-Miyaura Cross-Coupling Reaction. Molecules 27, 967.

Jin, J., Lu, P., Xu, Y., Tao, J., Li, Z., Wang, S., Yu, S., Wang, C., Xie, X., Gao, J., Chen, Q., Wang, L., Pu, W., Cao, P., 2022. PCMDB: a curated and comprehensive resource of plant cell markers. Nucleic Acids Res. 50, D1448–D1455.

Krauze-Baranowska, M., 2004. Flavonoids from the genus *Taxus*. Z. Nat. C. J. Biosci. 59, 43–47.

Krauze-Baranowska, M., Wiwart, M., 2003. Antifungal activity of biflavones from *Taxus baccata* and *Ginkgo biloba*. Z. Nat. C. J. Biosci. 58, 65–69.

Liao, R.Y., Wang, J.W., 2023. Analysis of meristems and plant regeneration at single-cell resolution. Curr. Opin. Plant Biol. 74, 102378.

Liu, Y., Nie, X., Wang, J., Zhao, Z., Wang, Z., Ju, F., 2023. Visualizing the distribution of flavonoids in litchi (*Litchi chinenis*) seeds through matrix-assisted laser desorption/ionization mass spectrometry imaging. Front Plant Sci. 14, 1144449.

Liu, Z., Yu, X., Qin, A., Zhao, Z., Liu, Y., Sun, S., Liu, H., Guo, C., Wu, R., Yang, J., Hu, M., Bawa, G., Sun, X., 2022. Research strategies for single-cell transcriptome analysis in plant leaves. Plant J. 112, 27–37.

Lorensen, M., Hayat, S.Y., Wellner, N., Bjarnholt, N., Janfelt, C., 2023. Leaves of Cannabis sativa and their trichomes studied by DESI and MALDI mass spectrometry imaging for their contents of cannabinoids and flavonoids. Phytochem Anal. 34, 269–279.

Lu, Z.J., Ye, J.G., Wang, D.L., Li, M.K., Zhang, Q.K., Liu, Z., Huang, Y.J., Pan, C.N., Lin, Y. H., Shi, Z.X., Zheng, Y.F., 2023. Integrative single-cell RNA-seq and ATAC-seq analysis of mouse corneal epithelial cells. Invest Ophthalmol. Vis. Sci. 64, 30.

Majeed, A., Singh, A., Sharma, R.K., Jaitak, V., Bhardwaj, P., 2020. Comprehensive temporal reprogramming ensures dynamicity of transcriptomic profile for adaptive response in *Taxus contorta*. Mol. Genet Genom. 295, 1401–1414.

Marand, A.P., Zhang, X., Nelson, J., Braga Dos Reis, P.A., Schmitz, R.J., 2021. Profiling single-cell chromatin accessibility in plants. STAR Protoc. 2, 100737.

Moller, M., Liu, J., Li, Y., Li, J.H., Ye, L.J., Mill, R., Thomas, P., Li, D.Z., Gao, L.M., 2020. Repeated intercontinental migrations and recurring hybridizations characterise the evolutionary history of yew (*Taxus* L.). Mol. Phylogenet. Evol. 153, 106952.

Perez-Matas, E., Hanano, A., Moyano, E., Bonfill, M., Cusido, R.M., Palazon, J., 2022. Insights into the control of taxane metabolism: molecular, cellular, and metabolic changes induced by elicitation in *Taxus baccata* cell suspensions. Front Plant Sci. 13, 942433.

Perez-Matas, E., Hidalgo-Martinez, D., Escrich, A., Alcalde, M.A., Moyano, E., Bonfill, M., Palazon, J., 2023. Genetic approaches in improving biotechnological production of taxanes: An update. Front Plant Sci. 14, 1100228.

Procko, C., Lee, T., Borsuk, A., Bargmann, B.O.R., Dabi, T., Nery, J.R., Estelle, M., Baird, L., O'Connor, C., Brodersen, C., Ecker, J.R., Chory, J., 2022. Leaf cell-specific and single-cell transcriptional profiling reveals a role for the palisade layer in UV light protection. Plant Cell 34, 3261–3279.

Pu, X., Li, J., Guo, Z., Wang, M., Lei, M., Yang, S., Yang, J., Wang, H., Zhang, L., Huang, Q., 2022. Structure-based identification and pathway elucidation of flavonoids in *Camptotheca acuminate*. Synth. Syst. Biotechnol. 7, 824–836.

Ruan, X., Zhan, L.M., Gao, X.X., Yan, L.Y., Zhang, H., Zhu, Z.Y., Wang, Q., Jiang, D.A., 2013. Separation and purification of flavonoid from *Taxus* remainder extracts free of taxoids using polystyrene and polyamide resin. J. Sep Sci. 36, 1925–1934.

Sanchez-Munoz, R., Bonfill, M., Cusido, R.M., Palazon, J., Moyano, E., 2018. Advances in the regulation of in vitro paclitaxel production: methylation of a Y-patch promoter region alters *BAPT* gene expression in *Taxus* cell cultures. Plant Cell Physiol. 59, 2255–2267.

Sievert, C., Beuerle, T., Hollmann, J., Ober, D., 2015. Single cell subtractive transcriptomics for identification of cell-specifically expressed candidate genes of pyrrolizidine alkaloid biosynthesis. Phytochemistry 117, 17–24.

Springer, N.M., Schmitz, R.J., 2017. Exploiting induced and natural epigenetic variation for crop improvement. Nat. Rev. Genet 18, 563–575.

Szopa, A., Klimek-Szczykutowicz, M., Kokotkiewicz, A., Dziurka, M., Luczkiewicz, M., Ekiert, H., 2019. Phenolic acid and flavonoid production in agar, agitated and bioreactor-grown microshoot cultures of *Schisandra chinensis* cv. Sadova No. 1 - a valuable medicinal plant. J. Biotechnol. 305, 61–70.

Tabaszewska, M., Antoniewska, A., Rutkowska, J., Skoczylas, L., Slupski, J., Skoczen-Slupska, R., 2021. Bioactive components, volatile profile and *in vitro* antioxidative properties of *Taxus baccata* L. Red Arils. Molecules 26, 4474.

Tenorio Berrio, R., Verstaen, K., Vandamme, N., Pevernagie, J., Achon, I., Van Duyse, J., Van Isterdael, G., Saeys, Y., De Veylder, L., Inze, D., Dubois, M., 2022. Single-cell transcriptomics sheds light on the identity and metabolism of developing leaf cells. Plant Physiol. 188, 898–918.

Valinas, M.A., Lanteri, M.L., ten Have, A., Andreu, A.B., 2015. Chlorogenic acid biosynthesis appears linked with suberin production in potato tuber (*Solanum tuberosum*). J. Agr. Food Chem. 63, 4902–4913.

Vue, B., Zhang, S., Chen, Q.H., 2016. Flavonoids with therapeutic potential in prostate cancer. Anticancer Agents Med Chem. 16, 1205–1229.

Waki, T., Takahashi, S., Nakayama, T., 2021. Managing enzyme promiscuity in plant specialized metabolism: A lesson from flavonoid biosynthesis: Mission of a "body double" protein clarified. Bioessays 43, e2000164.

Wang, N., Song, G., Zhang, F., Shu, X., Cheng, G., Zhuang, W., Wang, T., Li, Y., Wang, Z., 2023a. Characterization of the WRKY gene family related to anthocyanin biosynthesis and the regulation mechanism under drought stress and methyl jasmonate treatment in *Lycoris radiata*. Int J. Mol. Sci. 24, 2423.

Wang, Y., Wang, X., Fang, J., Yin, W., Yan, X., Tu, M., Liu, H., Zhang, Z., Li, Z., Gao, M., Lu, H., Wang, Y., Wang, X., 2023b. VqWRKY56 interacts with VqbZIPC22 in grapevine to promote proanthocyanidin biosynthesis and increase resistance to powdery mildew. N. Phytol. 237, 1856–1875.

Wang, T., Zhang, F., Zhuang, W., Shu, X., Wang, Z., 2019. Metabolic variations of flavonoids in leaves of *T. media* and *T. mairei* obtained by UPLC-ESI-MS/MS. Molecules 24, 3323.

Wei, Q., Li, Q.Z., Wang, R.L., 2023. Flavonoid components, distribution, and biological activities in *Taxus*: A review. Molecules 28, 1713.

Wen, K., Fang, X., Yang, J., Yao, Y., Nandakumar, K.S., Salem, M.L., Cheng, K., 2021. Recent research on flavonoids and their biomedical applications. Curr. Med Chem. 28, 1042–1066.

Xiong, X., Gou, J., Liao, Q., Li, Y., Zhou, Q., Bi, G., Li, C., Du, R., Wang, X., Sun, T., Guo, L., Liang, H., Lu, P., Wu, Y., Zhang, Z., Ro, D.K., Shang, Y., Huang, S., Yan, J., 2021. The *Taxus* genome provides insights into paclitaxel biosynthesis. Nat. Plants 7, 1026–1036.

Xiong, L., Xu, K., Tian, K., Shao, Y., Tang, L., Gao, G., Zhang, M., Jiang, T., Zhang, Q.C., 2019. SCALE method for single-cell ATAC-seq analysis via latent feature extraction. Nat. Commun. 10, 4576.

Xu, M., Shen, C., Zhu, Q., Xu, Y., Xue, C., Zhu, B., Hu, J., 2022. Comparative metabolomic and transcriptomic analyses revealed the differential accumulation of secondary metabolites during the ripening process of acerola cherry (Malpighia emarginata) fruit. J. Sci. Food Agric. 102, 1488–1497.

Yang, L., Zheng, Z.S., Cheng, F., Ruan, X., Jiang, D.A., Pan, C.D., Wang, Q., 2016. Seasonal dynamics of metabolites in needles of *Taxus wallichiana* var. *mairei*. Molecules 21, 1403.

Yu, C., Hou, K., Zhang, H., Liang, X., Chen, C., Wang, Z., Wu, Q., Chen, G., He, J., Bai, E., Li, X., Du, T., Wang, Y., Wang, M., Feng, S., Wang, H., Shen, C., 2023. Integrated mass spectrometry imaging and single-cell transcriptome atlas strategies provide novel insights into taxoid biosynthesis and transport in *Taxus mairei* stems. Plant J. 115, 1243–1260.

Yu, C., Luo, X., Zhang, C., Xu, X., Huang, J., Chen, Y., Feng, S., Zhan, X., Zhang, L., Yuan, H., Zheng, B., Wang, H., Shen, C., 2020. Tissue-specific study across the stem of *Taxus media* identifies a phloem-specific TmMYB3 involved in the transcriptional regulation of paclitaxel biosynthesis. Plant J. 103, 95–110.

Zhan, X., Qiu, T., Zhang, H., Kailin, H., Liang, X., Chen, C., Wang, Z., Wu, Q., Wang, X., Li, X.L., Wang, M., Feng, S., Zeng, H., Yu, C., Wang, H., Shen, C., 2023. Mass spectrometry imaging and single-cell transcriptional profiling reveal the tissue-specific regulation of bioactive ingredient biosynthesis in Taxus leaves. Plant Commun., 100630

Zhan, X., Zang, Y., Ma, R., Lin, W., Li, X.L., Pei, Y., Shen, C., Jiang, Y., 2024. Mass spectrometry-imaging analysis of active ingredients in the leaves of *Taxus cuspidata*. ACS Omega 9, 18634–18642.

Zhang, S., Li, L., Hu, J., Ma, P., Zhu, H., 2020. Polysaccharide of *Taxus chinensis* var. *mairei* Cheng et L.K.Fu attenuates neurotoxicity and cognitive dysfunction in mice with Alzheimer's disease. Pharm. Biol. 58, 959–968.

Zhang, L., Sun, X., Wilson, I.W., Shao, F., Qiu, D., 2019. Identification of the genes involved in anthocyanin biosynthesis and accumulation in *Taxus chinensis*. Genes 10, 982.

Zhao, X., Zhang, Y., Long, T., Wang, S., Yang, J., 2022. Regulation mechanism of plant pigments biosynthesis: anthocyanins, carotenoids, and betalains. Metabolites 12, 871.