BMC Genomics

RESEARCH

Open Access

Transcriptome and genome-wide analysis of the mango glycosyltransferase family involved in mangiferin biosynthesis

Yibo Bai[1,2†], Xinran Huang[2,3†], Rundong Yao[2,3], Muhammad Mubashar Zafar[2*], Waqas Shafqat Chattha[4], Fei Qiao[1,2,5*] and Hanqing Cong[1,2*]

Abstract

Mangiferin, a C-glucosyl xanthone, is a biologically active glycoside naturally synthesized in mango. Glycosyltransferase can catalyze the biosynthesis of mangiferin. In this study, we identified 221 members of the UGT glycosyltransferase family in mango. The 221 *MiUGT* genes were grouped into 13 subfamilies through phylogenetic tree analysis with *Arabidopsis*, Chinese bayberry, and mango. All UGT family members in mango were unevenly distributed on 17 chromosomes and found that tandem duplication dominated the expansion of UGT family members in mango. Purification selection primarily influenced the evolution of the mango UGT family members. In addition, cis-element analysis of the mango UGT gene family revealed the presence of MYB binding sites, which are involved in flavonoid biosynthesis; which further supports the role of UGT family members in the synthesis of flavonoids. To verify these results, we analyzed the expression of UGT family members in mango leaves, stems, and different developmental stages of fruit peel. The RNA-seq and qRT-PCR results showed significant differences in the expression patterns of *MiUGT* genes in various tissues and developmental stages of mango. We identified *MiUGT* gene-specific expression at different stages of fruit development. These results lay a theoretical foundation for research on the relationship between members of the mango UGT family and the synthesis of flavonoids, mangiferin.

Keywords Flavonoids, Gene-specific expression, *MiUGT*, Tissues

†Yibo Bai and Xinran Huang contributed equally to this work.

*Correspondence:
Muhammad Mubashar Zafar
m.mubasharzafar@gmail.com
Fei Qiao
fei.qiao@catas.cn
Hanqing Cong
hanqing.cong@catas.cn
[1] National Key Laboratory for Tropical Crop Breeding, Tropical Crops Genetic Resources Institute, Chinese Academy of Tropical Agricultural Sciences, Sanya 572024, China
[2] Key Laboratory of Crop Gene Resources and Germplasm Enhancement in Southern China, Ministry of Agriculture and Rural Affairs, Tropical Crops Genetic Resources Institute, Chinese Academy of Tropical Agricultural Sciences, Haikou 571101, China
[3] National Key Laboratory for Germplasm Innovation & Utilization of Horticultural Crops, Huazhong Agricultural University, Wuhan 430070, China
[4] Department of Plant Breeding and Genetics, University of Agriculture, Faisalabad, Punjab 38000, Pakistan
[5] Sanya Research Institute, Chinese Academy of Tropical Agricultural Sciences, Sanya 572024, China

Introduction

Flavonoids are a diverse class of secondary metabolites that play multiple functions in plants [1–3]. Currently, there are over 6000 flavonoids have been identified based on their different structures and functions [3–5]. Flavonoids in plants mainly exist in the form of glycosides and aglycones, with glycosides being the predominant form [6]. Mangiferin, a natural antioxidant C-glucoside flavonoid, was originally isolated from mango [7]. Mangiferin has various pharmacological properties, including antioxidant, anti-inflammatory, anti-diabetic, and antibacterial [8–10]. However, its practical use is limited due to low solubility in water, low efficacy, and low bioavailability [11]. Glycosylation of mangiferin through biotransformation can enhance its bioavailability [12]. As a crucial step in the biosynthesis of natural compounds, glycosylation is recognized as a key biological transformation process that regulates the levels, activities, and positions of cellular metabolites and is involved in the degradation of most biomass [13, 14].

The process of glycosylation is catalyzed by glycosyltransferases [15]. UDP-glycosyltransferases (UGTs) are a specific type of glycosyltransferase involved in the biosynthesis of various glycosides [16]. In *Arabidopsis,* 117 *UGT* genes have been identified, with five *AtUGT* genes confirmed to be involved in the synthesis of plant flavonoids [17]. In Panax ginseng, 253 *UGT* genes have been identified, highlighting the crucial role *UGTs* play in the biosynthesis of ginsenosides [18, 19]. The quantitative trait locus *GSA1* in rice regulates metabolic flux from lignin biosynthesis to flavonoid glycoside accumulation by encoding UDP glucosyltransferase [20]. UDP-glycosyltransferase can improve mangiferin's water solubility and antioxidant properties by catalyzing the transfer of sugar moieties from donors to receptors [21]. As a member of UDP glycyltransfer, the *MiCGT* genes have been proven to participate in the biosynthesis of mangiferin by producing C-xyloside through UDP xylose [22].

UDP glycyltransfer was widely distributed in natural organisms, including animals, plants, yeast, and bacteria. Therefore, to distinguish *UGT* genes in different organisms, the UGT naming committee assigned 200 UGT subfamilies to different biological categories. Among them, UGT 71–100 belongs to the UGT subfamily of plants [23]. As a superfamily of enzymes that catalyze glucosylation, the structural characteristics of UGT family members are typical UDPGT domains, which contain binding sites for UDP glycosyl donors [24]. Glycosyltransferases are responsible for transferring the glycosyl portion from activated donor molecules to receptor molecules such as sugars, nucleic acids, and proteins, thereby forming various glycoside compounds [25].

The mango (*Mangifera indica* L.) is an evergreen large tree plant in the family *Rhus*, genus mangifera. Due to its rich taste and potential as a traditional Chinese medicine, it has high economic cultivation value [26]. Mangiferin, a bioactive compound that is commonly found in mango leaves, stems, and rind, is renowned for its extensive pharmacological properties, which include antioxidant, anti-inflammatory, antidiabetic, and anticancer effects [27]. The health benefits of mango leaves, which are high in mangiferin, have been traditionally employed to manage diabetes and reduce inflammation, as evidenced by scientific studies [28]. Even though the stems contain reduced concentrations, they still possess valuable antioxidant and antimicrobial properties [29]. Another substantial source of mangiferin is the mango peel, which is frequently discarded. It is used in the production of dietary supplements, cosmetic products, and food preservation due to its antioxidant and antimicrobial properties [27]. With the development of science and technology, the genome of mango has been deciphered, laying the foundation for studying the biological functions of mango genes at the molecular level [30]. In previous reports, it has been demonstrated that mangiferin can be biosynthetically synthesized in mango by glycosyltransferases [22]. However, a comprehensive genome-wide analysis of the glycosyltransferase family members and their roles in mangiferin synthesis has not been conducted. Therefore, this study aims to identify the members of the glycosyltransferase family in mango, along with their evolutionary relationships and expression patterns. This will provide a theoretical foundation for further analysis of their roles in mangiferin biosynthesis.

Materials and methods

Genome-wide identification of the UGT gene family in three plant species

The protein sequences of UGT gene family members in *Arabidopsis* were identified from the TAIR website (https://www.arabidopsis.org), and the protein sequences of UGT gene family members in Chinese bayberry were downloaded from previous studies [31]. The whole genome sequence of mango was downloaded from the NCBI (https://www.ncbi.nlm.nih.gov) [30]. UDP-glucoronosyl and UDP-glucosyl transferase domain (PF00201) and Myb-like DNA-binding domain (PF00249) retrieve hidden Markov Model (HMM) profiles were downloaded from the Pfam website (http://pfam.xfam.org), HMMER3.0 was used for identifying members of the UGT gene family in mango. Using Pfam and SMART (http://smart.embl.de) for *UGT* gene validation to ensure the presence of the UDPGT domain [32].

The protein sequence of UGT gene family members of *Arabidopsis*, Chinese bayberry, and mango were aligned

using the AGLIN function in MEGA11, the UGT phylogenetic tree was constructed by IQ-tree using the neighbor-joining (NJ) method with default parameters, and bootstrap values from 1000 replicates [33].

Physicochemical characterization of UGT family protein sequences

The isoelectric point, molecular weight, and GRAVY information of the *UGT* genes in mango were predicted using the ExPASY website (https://www.expasy.org) [34]. Using WoLF PSORT (https://wolfpsort.hgc.jp) website for subcellular localization prediction of *UGT* gene in mango [35].

Gene structure and conserved motif analysis of the UGT genes

GSDS2.0 was used for *MiUGT* gene structure analysis (https://gsds.gao-lab.org) [36], and the MEME website for conserved motif analysis of the *MiUGT* genes with default parameters (https://meme-suite.org/meme/) [37].

Chromosomal locations and Ka/Ks range of MiUGT genes

Visualize the distribution of *UGT* genes on chromosomes in mango by Mapchart software [38]. The duplication events were calculated by the DupGen_finder tool (https://github.com/qiao-xin/DupGen_finder) [39]. Perform collinearity analysis of mango genome and *Arabidopsis* genome by MCScanX. The value of Ka, Ks, and Ka/Ks ratios were calculated using MEGA X [40].

Cis-elements analysis in MiUGT genes promoter region

The 2 kb promoter sequence of the *MiUGT* genes was extracted by Perl and predicted the *cis*-elements of *MiUGT* genes by the PlantCARE website (http://bioinformatics.psb.ugent.be/webtools/plantcare/html/) [41].

Sample collection and transcriptome sequencing

The plant of mango was planted in the dry and hot valley of the Youjiang River in Guangxi (Tropical Crops Genetic Resources Institute, Chinese Academy of Tropical Agricultural Sciences), select the leaves, stems, and fruits of "Hong xiangya" were about 15 days after flowering, and collect the fruit peel every week as experimental materials (Fig. S1). A total of 12 developmental stages of experimental materials were collected, with three biological replicates set for each stage. All collected samples were rapidly frozen in liquid nitrogen and stored in a -80 ℃ freezer for RNA extraction.

The RNA-prep Pure Plant Plus Kit (Polysaccharides & Polyphenolics rich) kit (Tiangen, China) was used to extract RNA from mango leaves, stems, and fruit peel. The purity and integrity of RNA were detected by spectrophotometer (NanoDrop, Thermo Scientific), agarose gel electrophoresis (1%), and Agilent 2100 bioanalyzer. At least 1 μg of RNA in each sample for transcriptome sequencing by the Illumina Novaseq 6000 platform. After quality control, clean reads were mapped onto the reference genome of Alphonso by HISAT2 (https://www.ncbi.nlm.nih.gov/sra), and gene expression levels were calculated by TPM (Transcripts Per Million) [42].

Co-expression network construction between MiUGT and MiMYB genes

The *MiUGT* genes with MYB binding site involvement in flavonoid biosynthetic gene regulation were screened by *cis*-elements analysis. The Pearson correlation coefficient between *MiUGT* and *MiMYB* genes was performed from transcriptome sequencing of samples from different tissue parts of mango by the 'cor' function in R. The gene pairs with correlation coefficients $\geq|0.9|$ were considered highly correlated. The gene pairs with high correlation were visualized by Cytoscape software [43].

RNA isolation and qRT-PCR analyses

Reverse transcription of RNA from all samples into cDNA using the HiScript IV RT SuperMix for qPCR kit (Vazyme, Nanjing, China). The qRT-PCR assay was performed by ChamQ Universal SYBR qPCR Master Mix (Vazyme, Nanjing, China) on the QuantStudio 6 Flex real-time PCR system (ThermoFisher, MA, USA). All primers were listed in Table S5. The gene relative expression level was calculated using the $2^{-\Delta\Delta Ct}$ method [44], with the *MiTUBB* gene as the reference gene in mango.

Results

Genome-wide identification and characterization of UGT genes in mango

221 *UGT* genes were identified in mango based on Pfam and HMMER searches. Physicochemical properties analysis showed that the range of theoretical relative molecular weight for *MiUGT* genes was 11.31 ~ 95.81 KDa, the range of isoelectric point was 4.67 ~ 8.94, and the range of the grand average of hydropathicity was -0.457 ~ 0.137. The prediction of subcellular localization showed that 115 MiUGT proteins were located in chloroplast, 55 MiUGT proteins were located in cytoplasm, 8 MiUGT proteins were located in endoplasmic reticulum, 10 MiUGT proteins were located in extracellular, 2 MiUGT proteins were located in mitochondrion, 20 MiUGT proteins were located in nuclear, 3 MiUGT proteins were located in peroxisomes, and 5 MiUGT proteins were located in vacuole (Table S1).

To investigate the phylogenetic relationship between *UGT* genes in mango and other plants. We selected the reported UGT gene family from *Arabidopsis* and Chinese bayberry, and the identified *UGT* genes in mango

from this study to construct the phylogenetic tree. Based on the distance of genetic relationships, 221 *UGT* genes in mango were divided into 13 subfamilies, namely UGT71, UGT72, UGT91, UGT88, UGT74, UGT84, UGT80, UGT75, UGT89, UGT76, UGT85, UGT78, UGT73 (Fig. 1).

Gene structure and motifs analysis in the MiUGT gene family

To understand the structural diversity of UGT gene family members in mango, we analyzed the gene structures and conserved motifs of 221 *MiUGT* genes (Fig. 2). A total of 15 conserved motifs of the *MiUGT* genes, each gene contains motif 1 (UDPGT domain), and motif 15 (Glycogen Phosphorylase B) was only present in the UGT88 subfamily. In addition, the structure of the *MiUGT* genes was analyzed, 68 *MiUGT* genes (30.77%) have no introns, while 94 *MiUGT* genes (42.53%) contain 1–2 exons.

Genome distribution and the expansion of MiUGTs

As shown in Fig. 3, *MiUGT* genes were unevenly distributed on 20 chromosomes, with a higher number of 45 *MiUGT* genes on Chr 15. There were a small number of

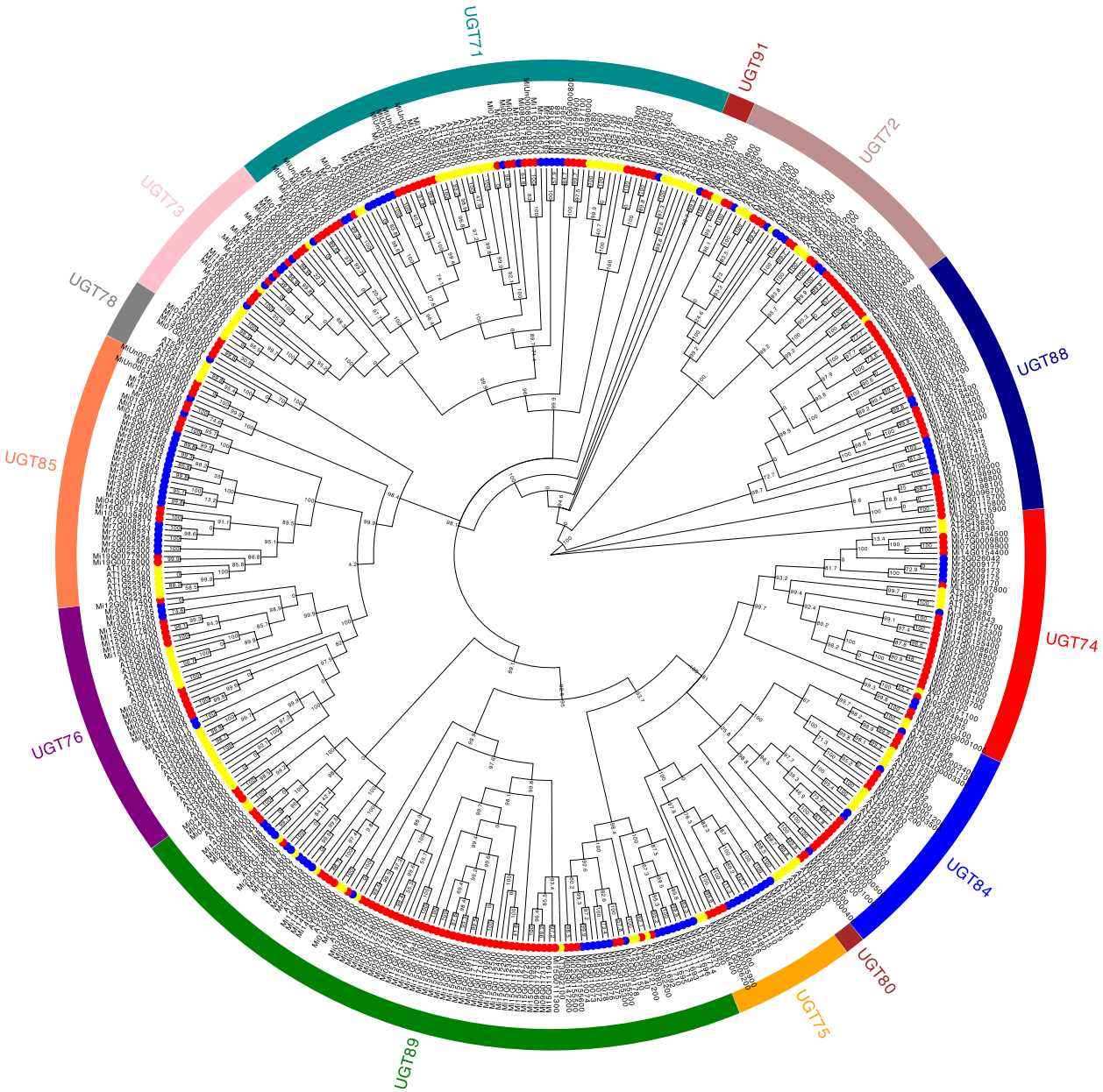

Fig. 1 Phylogenetic tree of the UGT gene family in *Arabidopsis*, Chinese bayberry, and Mango. Different colored circles represent different species, and different colored lines represent different subfamilies

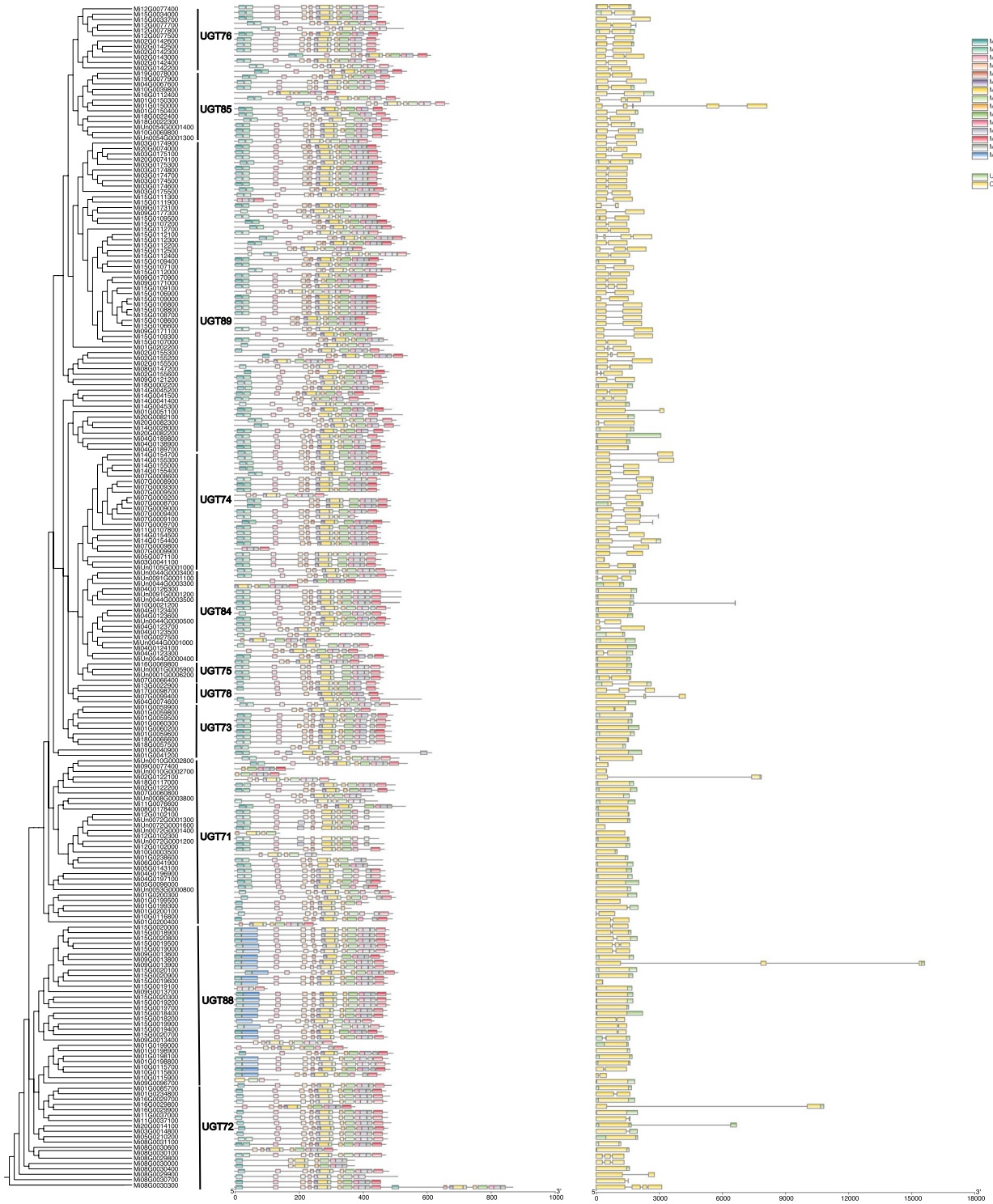

Fig. 2 The motif and structural analysis of the *UGT* genes in mango

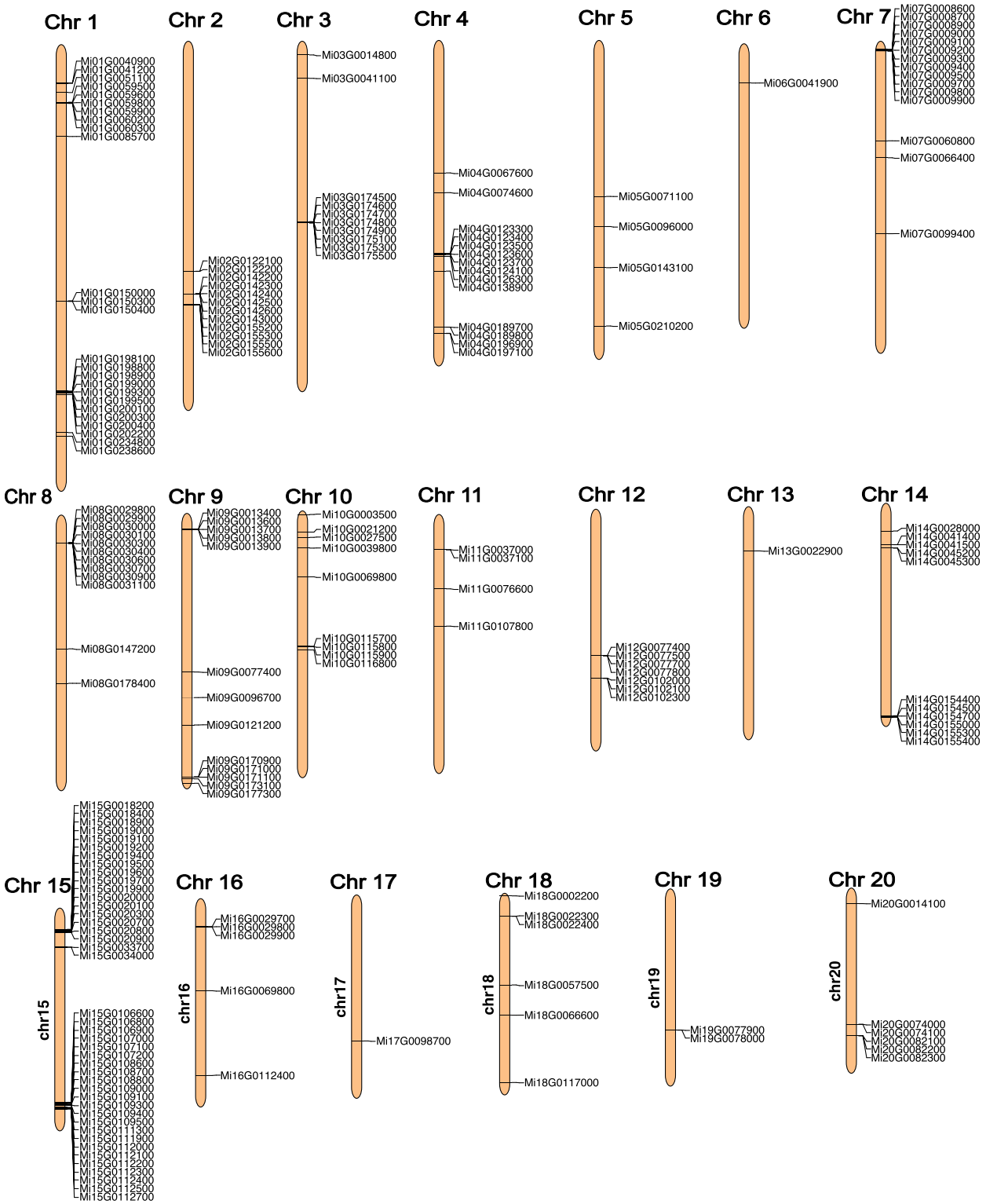

Fig. 3 Chromosome distribution map of MiUGT gene family

MiUGT genes on Chr 6, Chr 13, and Chr 17, which were one respectively. Gene duplication events play a crucial role in the evolution and expansion of gene families. The collinearity analysis showed that there are 43 UGT homologous gene pairs between *Arabidopsis* and mango (Fig. 4 and Table S4). In this study, we analyzed the duplication events of the *MiUGT* genes and found that only tandem duplication events occurred in the MiUGT gene family. This result indicated that tandem duplication was important in *MiUGT* gene expansion. To understand the evolution of *UGT* genes in mango, we calculated the Ka/Ks ratio of duplication events. The results showed that the ratio of Ka/Ks was mainly concentrated between 0–1, indicating that the *MiUGT* genes were mainly affected by purification selection during evolution (Fig. 5).

Analysis of cis-elements in the promoter regions of MiUGT genes

We selected the promoter region of the *MiUGT* genes (2000 bp upstream of the DNA sequence translation start site) for *cis*-element analysis. 1226 *cis*-acting elements have been identified and 14 *cis*-elements were found in the promoter region of the *MiUGT* genes. As expected, the ABRE motifs involved in abscisic acid response and the TGACG motifs involved in MeJA response were enriched in all *MiUGT* genes. The *cis*-element (MBSI) of the MYB binding site involved in flavonoid biosynthetic gene regulation, the *cis*-element (MBS) of the MYB binding site involved in drought-inducibility, and the *cis*-element (MRE) of the MYB binding site involved in light responsiveness were found in the *MiUGT* promoter region. This result suggested that the MYB genes may bind to the *UGT* genes and participate in the growth and development of mango (Fig. 6).

The interaction relationship between MiUGT genes and MiMYB genes

To identify the interaction between the *MiUGT* and *MiMYB* genes, we selected the *MiMYB* genes and *MiUGT* genes with MYB binding site involvement in flavonoid biosynthetic gene regulation in the promoter region for correlation analysis. We selected 20 *MiUGT* genes with MYB binding sites to calculate the Pearson correlation

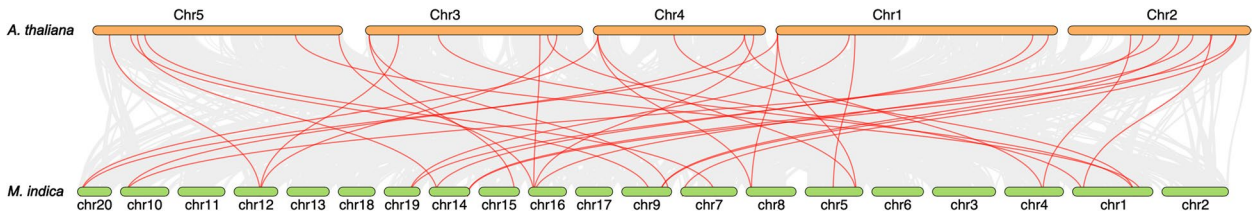

Fig. 4 Collinearity relationships of UGTs among *A. thaliana* and *M. indica*

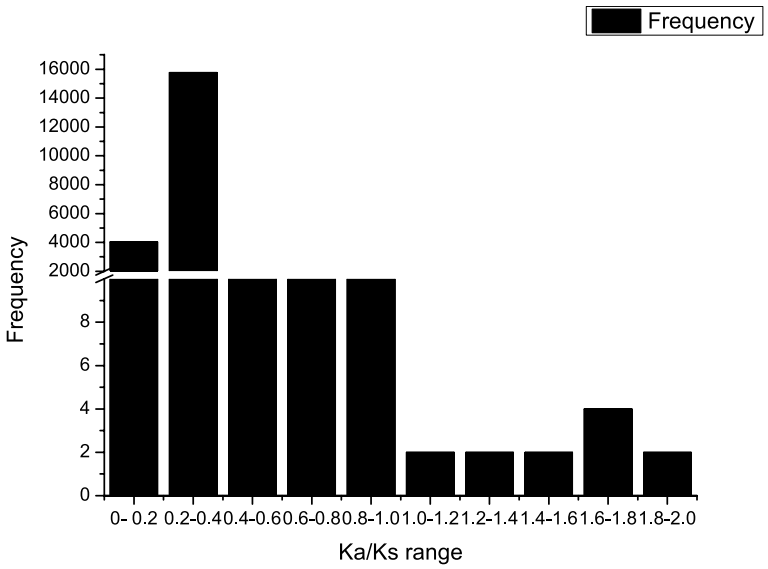

Fig. 5 The Ka/Ks ratio of duplication events in *MiUGT* genes

Fig. 6 2000-bp promoter sequence analysis in members of the MiUGT gene family. Different colors represent different *cis*-elements

coefficient between candidate *MiUGT* genes and *MiMYB* by the expression level in different tissues of mango. Through screening, it was found that there was a high correlation ($P \geq 0.9$) between 60 *MiMYB* genes and 13 *MiUGT* genes. There was a high negative correlation between the *MiMYB* genes (*Mi08G0079000*, *Mi04G0090900*, *Mi03G0073600*) and the *MiUGT* genes (*Mi04G0197100*, *Mi15G0019600*, *Mi04G0197100*). There was a high positive correlation between other gene pairs (Fig. 7).

Expression patterns and co-expression networks of MiUGT in different organs

To further determine the expression patterns of the *MiUGT* genes in different tissues of mango, we performed transcriptome sequencing on the stems, leaves, and fruit peels at different developmental stages of the fruit. The heatmap results have shown that there were similar expression patterns of *MiUGT* genes in stems and leaves, at least half of the *MiUGT* genes were highly expressed in leaves and stems (Fig. 8). There were significant differences in the expression of *MiUGT* genes in different stages of fruit peel development. 40 *MiUGT* genes were mainly expressed in the early stage of mango fruit development, while 50 *MiUGT* genes were highly expressed in the middle stage. Interestingly, 12 *MiUGT* genes were specifically highly expressed in the post-ripening stage of mango. These results indicated that the *MiUGT* genes played a key role in the mango's different tissues and developmental stages.

QRT-PCR analysis of MiUGT genes under different organs in mango

To validate the expression of the *MiUGT* genes in various tissue parts of mango, we selected different subfamily genes for qRT-PCR validation. To verify the expression and potential function of the *MiUGT* genes in different tissues of mango, the *MiUGT* genes that can bind to the *MiMYB* genes were selected for qRT-PCR verification. The results showed that the genes *Mi02G0155500*, *MiUn0001G0005900*, and *Mi14G0155300* were mainly highly expressed in leaves and stems, the gene of *Mi15G0020000* was mainly highly expressed in stems, and the genes of *Mi15G0112400* and *Mi04G0197100* were mainly highly expressed in the peel at the early stage of fruit development. The genes of *Mi02G142500*, *Mi15G0019600*, *Mi01G0200100*, *Mi02G0122200*, *Mi09G0013400*, *Mi01G0085700*, and *Mi05G0071100* were mainly highly expressed in the peel at the middle stage of fruit ripening (Fig. 9).

Discussion

Mango leaves and bark contain various naturally bioactive glycosides [26]. Mangiferin is a glycoside flavonoid compound with medicinal activity, mainly enriched in

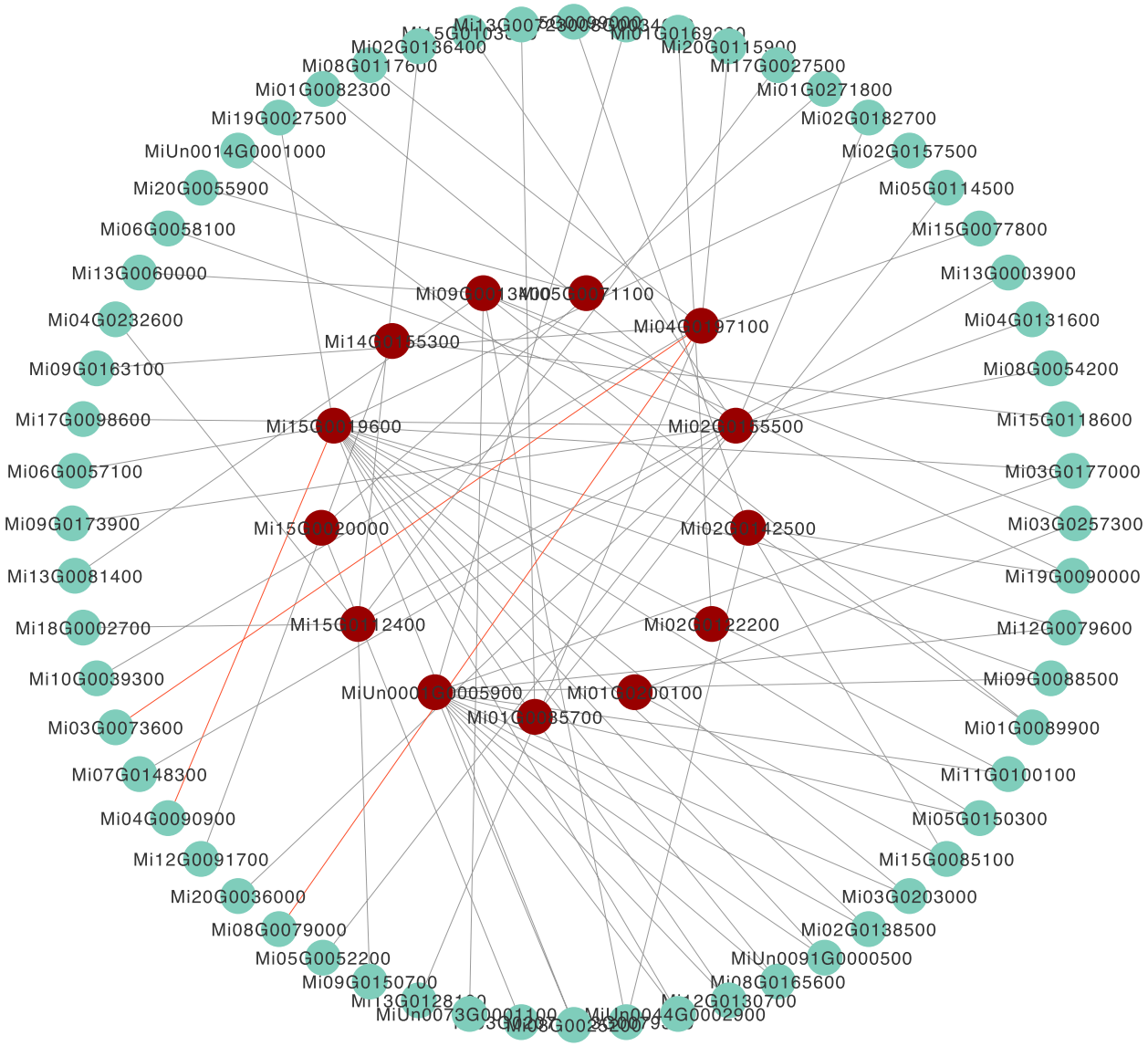

Fig. 7 Co-expression analysis of *MiUGT* genes with *MiMYB* genes. The black line represents a positive correlation between genes, while the red line represents a negative correlation between genes

mango leaves, stems, and peels, commonly used for the defense and treatment of human diseases [8, 45, 46]. Previous studies have shown that using enzymatic glycosylation of mangiferin to produce more soluble mangiferin can solve the problem of poor solubility and low bioavailability of mango [22]. UGT, as a UDP glycyltransferase, catalyzed the glycosylation of glycoside substances [47]. It has been reported in multiple species such as *G. biloba* [48], tobacco [49], *Arabidopsis* [50], rice [51], cabbage [52], and sweet orange [50]. In *G. biloba*, *UGT716A1* has been identified as a multi-substrate UFGT with a broad in vitro flavonoid substrate spectrum [48]. In our study, we identified 221 UGT family members in mango and analyzed the physicochemical properties based on the gene structures. The relative molecular weight of the MiUGT protein ranges from 11.31 KDa (*Mi09G0013700*) to 95.81 KDa (*Mi08G0030900*), and the pI range of the MiUGT protein ranges from 4.67 (*Mi02G015500*) to 8.94 (*Mi04G0074600*). The majority of the MiUGT protein's pI was less than 7. These results indicated that similar to most studies on dicotyledonous plants [53], the MiUGT gene family tends to favor acidic amino acids, and over time, the differentiation of MiUGT gene family members is inconsistent, leading to structural differences [54]. Subcellular localization prediction of 221 *MiUGT* genes revealed that half of the

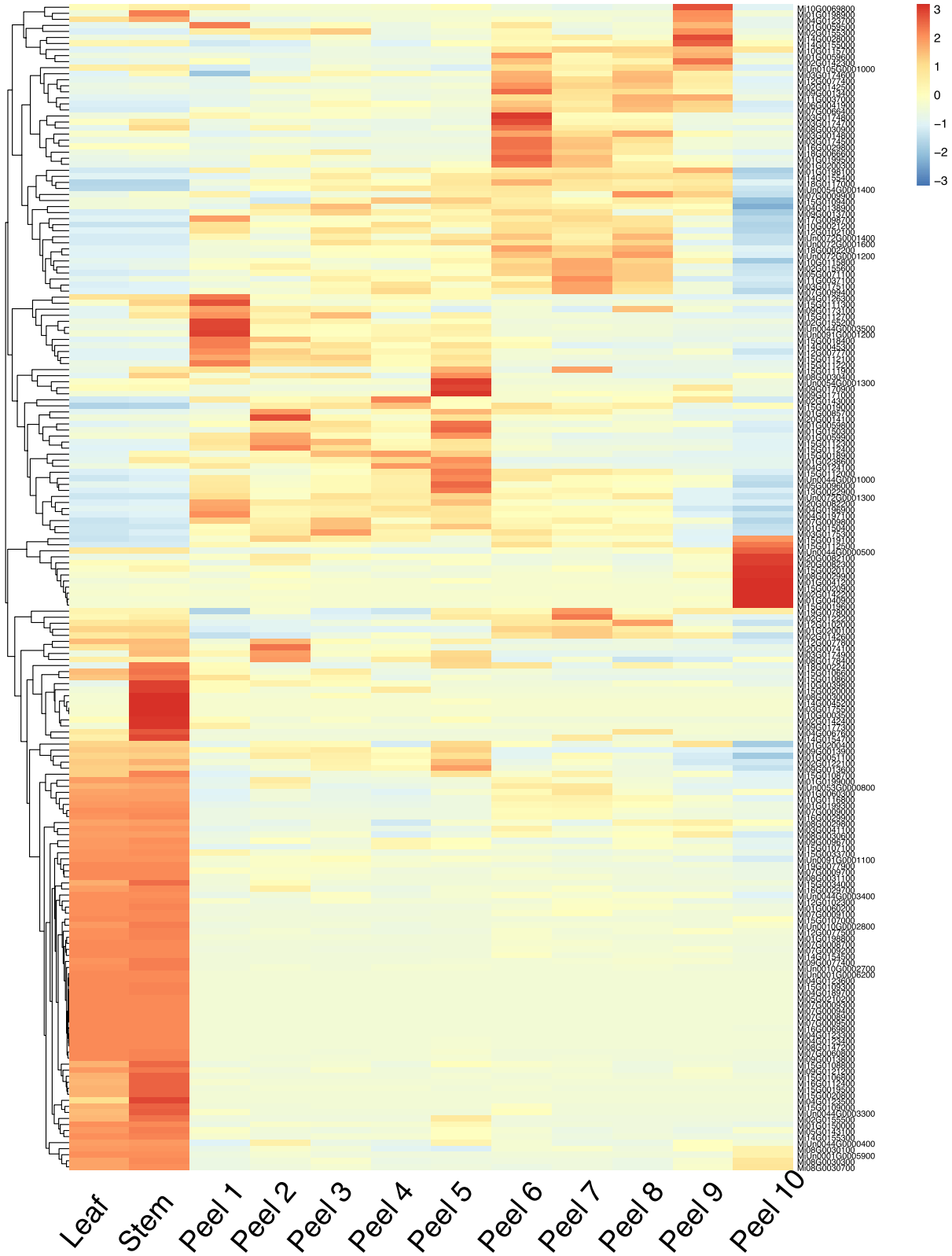

Fig. 8 Heatmap of *MiUGT* genes expression in mango peel, stem, and leaf

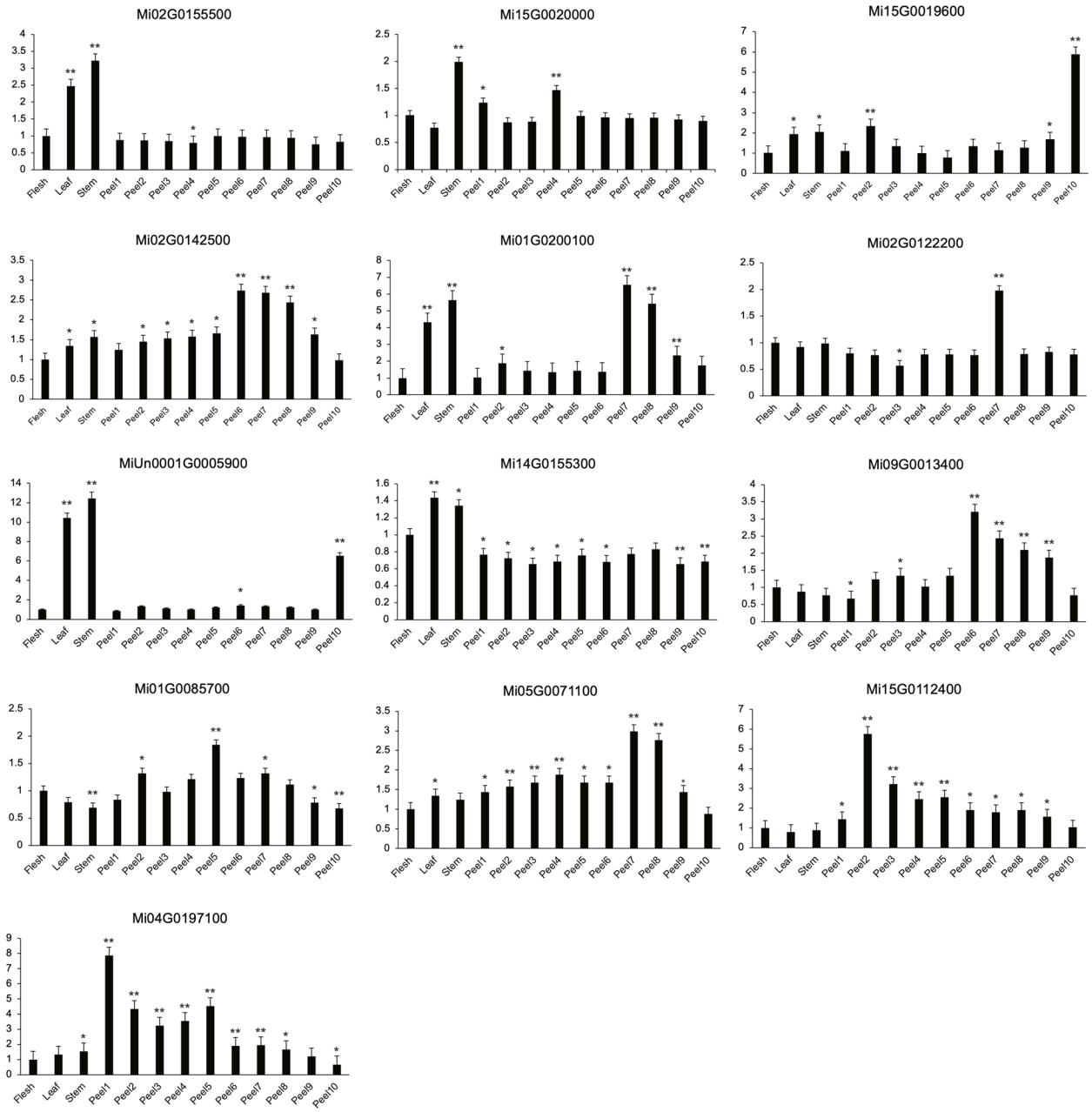

Fig. 9 QRT-PCR analysis of the expression of *MiUGT* genes in stems, leaves, and fruit peels

MiUGT genes were located on chloroplast. Previous studies have shown that light affects the transcription of chloroplast localization genes and regulates the related functions of photosynthesis [55]. These results indicated that the *MiUGT* genes play a crucial role in plant growth, development, and adaptation to environmental changes. Based on phylogenetic tree analysis of *UGT* genes in *Arabidopsis*, Chinese bayberry, and mango, 221 *MiUGT* genes were divided into 13 subfamilies. The genes in all 13 subfamilies have conserved UDP-glucosyl transferase domains. The conserved motifs among the same subfamily are relatively similar, for instance, motif 15 (Glycogen Phosphorylase B) mainly exists in the UGT88 subfamily. As a key enzyme in the first step of glycogen reaction, glycogen phosphatase can provide sufficient energy for the glycogen metabolism process [56], this may enable members of the MiUGT88 subfamily to play a unique role in plant development.

During the evolutionary process, plants undergo various gene duplication events, which result in the differentiation of different family members [57]. The chromosome mapping results showed that 221 members of the MiUGT gene family were unevenly distributed on 20 chromosomes of mango, with the most *MiUGT* genes distributed on Chr 15. Tandem duplication usually leads to the development of gene clusters [58]. We observed that the *MiUGT* genes only contain tandem duplication events, involving 72 pairs of genes distributed on 16 chromosomes, indicating that tandem duplication is the main cause of the expansion of the mango UGT gene family (Table S2). This result was similar to previous research, the *BpUGT* genes were involved in segment duplication and tandem duplication events in *B. papyrifera*. Tandem duplication has been proven to be the main driving force for the expansion of the BpUGT family [59]. The Ka/Ks ratio indicates the selection pressure during plant evolution [25]. The Ka/Ks > 1 of duplicated genes indicated the necessity of the *UGT* genes in *E. pubescens* survival. In this study, the Ka/Ks results indicated that the evolution of *MiUGT* genes is mainly influenced by purification selection.

The DNA sequence upstream of the gene coding region is the gene promoter, which contains multiple *cis*-acting elements [60]. We identified 10 types of homeostatic elements in the *MiUGT* genes. Interestingly, the presence of MBSI elements was found in 20 *MiUGT* genes. In previous studies, the *MYB* genes are involved in synthesizing flavonoids [61]. For example, in gerbera, overexpression of GhMYB1a can promote the accumulation of anthocyanins [62]. The *FhMYB21L2* gene activates the *FhFLS2* gene to promote the accumulation of flavonols in flowers in *Freesia hybrida* [63]. To further verify the relationship between *MiUGT* genes and *MiMYB* genes, we constructed a co-expression network of candidate *MiUGT* genes and *MiMYB* genes, with 74 positive and 3 negative correlations. The transcription factors of MYB have been reported to regulate the biosynthesis and accumulation of anthocyanins in chrysanthemum [64] and celery [65]. Combined with previous research results, *MiMYB* genes may be involved in the synthesis of flavonoids in mango by regulating *MiUGT* genes.

The *UGT* genes have been proven to be involved in plant growth, development, and stress resistance [21]. We performed transcriptome sequencing on the stems, leaves, and peel of mango at different stages of development to elucidate the expression of the *MiUGT* genes in different tissues and developmental stages. In this study, the *MiUGT* genes were highly expressed specifically in stems, leaves, and fruit peels at different developmental stages, indicating its involvement in the growth and development of mangoes. In addition, we conducted qRT-PCR analysis on *MiUGT* genes that may be regulated by *MiMYB* genes. We found that the *MiUGT* genes that can bind to *MiMYB* genes are highly expressed in leaves, stems, and various epidermal development stages. This result indicated that different *MiMYB-MiUGT* pathways were involved in the biosynthesis of mangiferin in other tissues.

Conclusions

Mangiferin has many pharmacological properties such as anti-inflammatory, antipyretic and analgesic, and antibacterial. In actual production, due to the poor solubility of mangiferin, the low bioavailability limits the clinical application of mangiferin. Therefore, the identification of mangiferin biosynthesis and glycosylation genes is beneficial to improve the utilization of mangiferin. In this study, mango was used as the experimental material to identify the UGT gene family and evolution process of mangiferin synthesis, and the *MiUGT* genes that interact with the *MiMYB* genes were excavated by cis-element analysis. Combined with transcriptome and qRT-PCR, the expression of candidate the *MiUGT* genes in different tissues and peels of mango at different developmental stages of fruit was revealed. It laid a solid theoretical foundation for further revealing the biosynthetic pathway of mangiferin.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12864-024-10998-5.

Supplementary Material 1.

Supplementary Material 2.

Acknowledgements
Not applicable.

Authors' contributions
FQ, HC, and YB designed the experiments. YB and XH performed the experiments. YB, XH, RY and MMZ analyzed the data. YB and WSC wrote the paper. All authors read and approved the final manuscript.

Funding
The research was supported by the Major Science and Technology plan of Hainan Province (ZDKJ2021012), the Hainan Provincial Natural Science Foundation of China (322MS133), the Open Project of Guangxi Key Laboratory of Biology for Mango (GKLBM02205), and Central Public-interest Scientific Institution Basal Research Fund for Chinese Academy of Tropical Agricultural Sciences (No.1630032024023).

Data availability
The RNA-seq data were deposited to the Sequence Read Archive (SRA) at NCBI under the accession number PRJNA1120973 (https://www.ncbi.nlm.nih.gov/bioproject/PRJNA1120973).

Declarations

Ethics approval and consent to participate
Not applicable.

Consent for publication
Not applicable.

Competing interests
The authors declare no competing interests.

Received: 7 June 2024 Accepted: 5 November 2024
Published online: 12 November 2024

References

1. Tahir F, Fatima F, Fatima R, Ali E. Fruit peel extracted polyphenols through ultrasonic assisted extraction: A review. Agrobiological Records. 2024;15:01–12.
2. Rehman I, Nazim R, Meraj A, Ameer A, Arshad F, Ali S. Regulation and improvement of carotenoid metabolism in tomato. Int J Agri Biosci. 2023;12:199–207.
3. Tahir Z, Khan MI, Ashraf U, Adan IRDN, Mubarik U. Industrial application of orange peel waste; a review. Int J Agri Biosci. 2023;12(2):71–6.
4. Ji Y, Li B, Qiao M, Li J, Xu H, Zhang L, Zhang X. Advances on the in vivo and in vitro glycosylations of flavonoids. Appl Microbiol Biotechnol. 2020;104(15):6587–600.
5. Anwar A, Ali E, Nisar W, Ashraf S, Javed N, Anwar L, Zulfiqar A, Nazir M, Tahir F. Development and quality evaluation of functional carbonated pomegranate orange nectar. Agrobiological Records. 2023;15:52–8.
6. Andualem M. Nutritional and anti-nutritional characteristics of okra (Abelmoschus esculents (L.) Moench) accessions grown in Pawe District, Northwestern Ethiopia. Int J Agri Biosci. 2023;12(1):18–23.
7. Campa C, Mondolot L, Rakotondravao A, Bidel LP, Gargadennec A, Couturon E, La Fisca P, Rakotomalala JJ, Jay-Allemand C, Davis AP. A survey of mangiferin and hydroxycinnamic acid ester accumulation in coffee (Coffea) leaves: biological implications and uses. Ann Bot. 2012;110(3):595–613.
8. Septiana I, Nguyen TTH, Lim S, Lee S, Park B, Kwak S, Park S, Kim SB, Kim D. Enzymatic synthesis and biological characterization of a novel mangiferin glucoside. Enzyme Microb Technol. 2020;134: 109479.
9. Mubashir A, Ghani A, Mubashar A. Common medicinal plants effective in peptic ulcer treatment: a nutritional review. Int J Agri Biosci. 2022;11(2):70–4.
10. Abdel-Mageed WM, Bayoumi SA, Chen C, Vavricka CJ, Li L, Malik A, Dai H, Song F, Wang L, Zhang J, et al. Benzophenone C-glucosides and gallotannins from mango tree stem bark with broad-spectrum anti-viral activity. Bioorg Med Chem. 2014;22(7):2236–43.
11. Li XJ, Du ZC, Huang Y, Liu BM, Hu WJ, Lu WJ, Deng JG. Synthesis and hypoglycemic activity of esterified-derivatives of mangiferin. Chin J Nat Med. 2013;11(3):296–301.
12. Carvalho AC, Guedes MM, de Souza AL, Trevisan MT, Lima AF, Santos FA, Rao VS. Gastroprotective effect of mangiferin, a xanthonoid from Mangifera indica, against gastric injury induced by ethanol and indomethacin in rodents. Planta Med. 2007;73(13):1372–6.
13. Rashid S, Ali L, Saleem A, Ummer K, Nisar I, Ahmed F, Zubair M. Fatima 1: Studies on the seasonal variation of major elements during fruit development stages and evaluation of post-harvest changes in mango germplasm. Agrobiological Records. 2024;15:75–90.
14. Lamichhaney S, Fuentes-Pardo AP, Rafati N, Ryman N, McCracken GR, Bourne C, Singh R, Ruzzante DE, Andersson L. Parallel adaptive evolution of geographically distant herring populations on both sides of the North Atlantic Ocean. Proc Natl Acad Sci USA. 2017;114(17):E3452-e3461.
15. Liang DM, Liu JH, Wu H, Wang BB, Zhu HJ, Qiao JJ. Glycosyltransferases: mechanisms and applications in natural product development. Chem Soc Rev. 2015;44(22):8350–74.
16. Meech R, Miners JO, Lewis BC, Mackenzie PI. The glycosidation of xenobiotics and endogenous compounds: versatility and redundancy in the UDP glycosyltransferase superfamily. Pharmacol Ther. 2012;134(2):200–18.
17. Le Roy J, Huss B, Creach A, Hawkins S, Neutelings G. Glycosylation is a major regulator of phenylpropanoid availability and biological activity in plants. Front Plant Sci. 2016;7:735.
18. Yu X, Yu J, Liu S, Liu M, Wang K, Zhao M, Wang Y, Chen P, Lei J, Wang Y, et al. Transcriptome-wide identification and integrated analysis of a UGT gene involved in ginsenoside ro biosynthesis in Panax ginseng. Plants (Basel). 2024;13(5):604.
19. Yang C, Li C, Wei W, Wei Y, Liu Q, Zhao G, Yue J, Yan X, Wang P, Zhou Z. The unprecedented diversity of UGT94-family UDP-glycosyltransferases in Panax plants and their contribution to ginsenoside biosynthesis. Sci Rep. 2020;10(1):15394.
20. Dong NQ, Sun Y, Guo T, Shi CL, Zhang YM, Kan Y, Xiang YH, Zhang H, Yang YB, Li YC, et al. UDP-glucosyltransferase regulates grain size and abiotic stress tolerance associated with metabolic flux redirection in rice. Nat Commun. 2020;11(1):2629.
21. Gharabli H, Della Gala V, Welner DH. The function of UDP-glycosyltransferases in plants and their possible use in crop protection. Biotechnol Adv. 2023;67: 108182.
22. Chen D, Chen R, Wang R, Li J, Xie K, Bian C, Sun L, Zhang X, Liu J, Yang L, et al. Probing the catalytic promiscuity of a regio- and stereospecific C-glycosyltransferase from *Mangifera indica*. Angew Chem Int Ed Engl. 2015;54(43):12678–82.
23. Mackenzie PI, Owens IS, Burchell B, Bock KW, Bairoch A, Bélanger A, Fournel-Gigleux S, Green M, Hum DW, Iyanagi T, et al. The UDP glycosyltransferase gene superfamily: recommended nomenclature update based on evolutionary divergence. Pharmacogenetics. 1997;7(4):255–69.
24. Zhou Y, Fu WB, Si FL, Yan ZT, Zhang YJ, He QY, Chen B. UDP-glycosyltransferase genes and their association and mutations associated with pyrethroid resistance in *Anopheles sinensis* (Diptera: Culicidae). Malar J. 2019;18(1):62.
25. Yao Y, Gu J, Luo Y, Wang Y, Pang Y, Shen G, Guo B. Genome-wide analysis of UGT gene family identified key gene for the biosynthesis of bioactive flavonol glycosides in Epimedium pubescens Maxim. Synth Syst Biotechnol. 2022;7(4):1095–107.
26. Barreto JC, Trevisan MT, Hull WE, Erben G, de Brito ES, Pfundstein B, Würtele G, Spiegelhalder B, Owen RW. Characterization and quantitation of polyphenolic compounds in bark, kernel, leaves, and peel of mango (*Mangifera indica* L.). J Agric Food Chem. 2008;56(14):5599–610.
27. Kumar M, Saurabh V, Tomar M, Hasan M, Changan S, Sasi M, Maheshwari C, Prajapati U, Singh S, Prajapat RK. Mango (*Mangifera indica* L.) leaves: Nutritional composition, phytochemical profile, and health-promoting bioactivities. Antioxidants. 2021;10(2):299.
28. Jhaumeer Laulloo S, Bhowon M, Soyfoo S, Chua L. Nutritional and biological evaluation of leaves of *Mangifera indica* from Mauritius. J Chem. 2018;2018(1):6869294.
29. Diso S, Ali M, Mukhtar S, Garba M. Antibacterial activity and phytochemical screening of *Mangifera indica* (Mango) stem and leaf extracts on clinical isolates of methicillin resistant Staphylococcus aureus. J Adv Med Pha Sci. 2017;13(1):1–6.
30. Wang P, Luo Y, Huang J, Gao S, Zhu G, Dang Z, Gai J, Yang M, Zhu M, Zhang H, et al. The genome evolution and domestication of tropical fruit mango. Genome Biol. 2020;21(1):60.
31. Ren C, Cao Y, Xing M, Guo Y, Li J, Xue L, Sun C, Xu C, Chen K, Li X. Genome-wide analysis of UDP-glycosyltransferase gene family and identification of members involved in flavonoid glucosylation in Chinese bayberry (*Morella rubra*). Front Plant Sci. 2022;13: 998985.
32. Letunic I, Khedkar S, Bork P. SMART: recent updates, new developments and status in 2020. Nucleic Acids Res. 2021;49(D1):458–60.
33. Nguyen LT, Schmidt HA, von Haeseler A, Minh BQ. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. Mol Biol Evol. 2015;32(1):268–74.
34. Duvaud S, Gabella C, Lisacek F, Stockinger H, Ioannidis V, Durinx C. Expasy, the swiss bioinformatics resource portal, as designed by its users. Nucleic Acids Res. 2021;49(W1):216–27.
35. Horton P, Park KJ, Obayashi T, Fujita N, Harada H, Adams-Collier CJ, Nakai K. WoLF PSORT: protein localization predictor. Nucleic Acids Res. 2007;35:585–7.
36. Hu B, Jin J, Guo AY, Zhang H, Luo J, Gao G. GSDS 2.0: an upgraded gene feature visualization server. Bioinformatics. 2015;31(8):1296–7.
37. Bailey TL, Johnson J, Grant CE, Noble WS. The MEME Suite. Nucleic Acids Res. 2015;43(W1):39–49.
38. Voorrips RE. MapChart: software for the graphical presentation of linkage maps and QTLs. J Hered. 2002;93(1):77–8.
39. Qiao X, Li Q, Yin H, Qi K, Li L, Wang R, Zhang S, Paterson AH. Gene duplication and evolution in recurring polyploidization-diploidization cycles in plants. Genome Biol. 2019;20(1):38.

40. Kumar S, Stecher G, Li M, Knyaz C, Tamura K. MEGA X: Molecular evolutionary genetics analysis across computing platforms. Mol Biol Evol. 2018;35(6):1547–9.
41. Lescot M, Déhais P, Thijs G, Marchal K, Moreau Y, Van de Peer Y, Rouzé P, Rombauts S. PlantCARE, a database of plant cis-acting regulatory elements and a portal to tools for in silico analysis of promoter sequences. Nucleic Acids Res. 2002;30(1):325–7.
42. Zhao Y, Li MC, Konaté MM, Chen L, Das B, Karlovich C, Williams PM, Evrard YA, Doroshow JH, McShane LM. TPM, FPKM, or normalized counts? A comparative study of quantification measures for the analysis of RNA-seq data from the NCI patient-derived models repository. J Transl Med. 2021;19(1):269.
43. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B, Ideker T. Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res. 2003;13(11):2498–504.
44. Rao X, Huang X, Zhou Z, Lin X. An improvement of the 2^(-delta delta CT) method for quantitative real-time polymerase chain reaction data analysis. Biostat Bioinforma Biomath. 2013;3(3):71–85.
45. Singh V, Dhyani S, Sharma S, Kumar SS. S S, Sinha R: Phytochemical and pharmacological investigation on mangiferin. Herba Polonica. 2009;55:126–39.
46. Tayana Ngampuk, Inthakusol Wichayasith, Duangdee Nongnaphat, Chewchinda Savita, Pandith Hataichanok. Mangiferin content in different parts of mango tree (*Mangifera indica* L.) in Thailand. Environ Sci Med. 2019;41(3):522–8.
47. Rahimi S, Kim J, Mijakovic I, Jung KH, Choi G, Kim SC, Kim YJ. Triterpenoid-biosynthetic UDP-glycosyltransferases from plants. Biotechnol Adv. 2019;37(7): 107394.
48. Su X, Shen G, Di S, Dixon RA, Pang Y. Characterization of UGT716A1 as a Multi-substrate UDP: Flavonoid glucosyltransferase gene in ginkgo biloba. Front Plant Sci. 2017;8:2085.
49. Yang Q, Zhang Y, Qu X, Wu F, Li X, Ren M, Tong Y, Wu X, Yang A, Chen Y, et al. Genome-wide analysis of UDP-glycosyltransferases family and identification of UGT genes involved in abiotic stress and flavonol biosynthesis in *Nicotiana tabacum*. BMC Plant Biol. 2023;23(1):204.
50. Liu X, Lin C, Ma X, Tan Y, Wang J, Zeng M. Functional characterization of a flavonoid glycosyltransferase in sweet orange (*Citrus sinensis*). Front Plant Sci. 2018;9:166.
51. Kantharaj V, Ramasamy NK, Yoon YE, Cheong MS, Kim YN, Lee KA, Kumar V, Choe H, Kim SY. Auxin-glucose conjugation protects the rice (*Oryza sativa* L.) seedlings against hydroxyurea-induced phytotoxicity by activating UDP-glucosyltransferase enzyme. Front Plant Sci. 2021;12.
52. Yu J, Hu F, Dossa K, Wang Z, Ke T. Genome-wide analysis of UDP-glycosyltransferase super family in *Brassica rapa* and *Brassica oleracea* reveals its evolutionary history and functional characterization. BMC Genomics. 2017;18(1):474.
53. Lai D, Yao X, Yan J, Gao A, Yang H, Xiang D, Ruan J, Fan Y, Cheng J. Genome-wide identification, phylogenetic and expression pattern analysis of GATA family genes in foxtail millet (*Setaria italica*). BMC Genomics. 2022;23(1):549.
54. Xiong R, Peng Z, Zhou H, Xue G, He A, Yao X, Weng W, Wu W, Ma C. Genome-wide identification, structural characterization and gene expression analysis of the WRKY transcription factor family in pea (*Pisum sativum* L.). BMC Plant Biol. 2024;24(1):113.
55. Cackett L, Luginbuehl LH, Schreier TB, Lopez-Juez E, Hibberd JM. Chloroplast development in green plant tissues: the interplay between light, hormone, and transcriptional regulation. New Phytol. 2022;233(5):2000–16.
56. Figueroa CM, Asencion Diez MD, Ballicora MA, Iglesias AA. Structure, function, and evolution of plant ADP-glucose pyrophosphorylase. Plant Mol Biol. 2022;108(4–5):307–23.
57. Vision TJ, Brown DG, Tanksley SD. The origins of genomic duplications in *Arabidopsis*. Science. 2000;290(5499):2114–7.
58. Graham GJ. Tandem genes and clustered genes. J Theor Biol. 1995;175(1):71–87.
59. Wang F, Su Y, Chen N, Shen S. Genome-wide analysis of the UGT gene family and identification of flavonoids in *Broussonetia papyrifera*. Molecules. 2021;26(11):3449.
60. Hernandez-Garcia CM, Finer JJ. Identification and validation of promoters and cis-acting regulatory elements. Plant Sci. 2014;217–218:109–19.
61. Feng D, Zhang H, Qiu X, Jian H, Wang Q, Zhou N, Ye Y, Lu J, Yan H, Tang K. Comparative transcriptomic and metabonomic analysis revealed the relationships between biosynthesis of volatiles and flavonoid metabolites in *Rosa rugosa*. Ornamental Plant Research. 2021;1(1):1–10.
62. Zhong C, Tang Y, Pang B, Li X, Yang Y, Deng J, Feng C, Li L, Ren G, Wang Y, et al. The R2R3-MYB transcription factor GhMYB1a regulates flavonol and anthocyanin accumulation in *Gerbera hybrida*. Hortic Res. 2020;7:78.
63. Shan X, Li Y, Yang S, Yang Z, Qiu M, Gao R, Han T, Meng X, Xu Z, Wang L, et al. The spatio-temporal biosynthesis of floral flavonols is controlled by differential phylogenetic MYB regulators in *Freesia hybrida*. New Phytol. 2020;228(6):1864–79.
64. Ahmad S, Yuan C, Cong T, Yang Q, Yang Y, Zhang Q. Transcriptome and chemical analyses identify candidate genes associated with flower color shift in a natural mutant of *Chrysanthemum* × *morifolium*. Ornamental Plant Res. 2022;2(1):1–12.
65. Feng K, Xing G, Liu J, Wang H, Tan G, Wang G, Xu Z, Xiong A. AgMYB1, an R2R3-MYB factor, plays a role in anthocyanin production and enhancement of antioxidant capacity in celery. Vegetable Res. 2021;1(1):1–12.

Publisher's Note