Research Square

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

COMT, CRTZ, and F3'H regulate glycyrrhizic acid biosynthesis in Glycyrrhizauralensis hairy roots

Zhixin Zhang
Beijing University of Chinese Medicine
Wenwen Ding
Beijing University of Chinese Medicine
Ziyi Chen
Beijing University of Chinese Medicine
Wenpu Xu
Beijing University of Chinese Medicine
Doudou Wang
Beijing University of Chinese Medicine
Tiangong Lu
Beijing University of Chinese Medicine
Ying Liu (liuyliwd@bucm.edu.cn)
Beijing University of Chinese Medicine https://orcid.org/0000-0001-9084-7894

Research Article

Keywords: Glycyrrhiza uralesis, COMT, CRTZ, F3’H, Gene overexpression, Gene knockout

Posted Date: September 21st, 2022

DOI: https://doi.org/10.21203/rs.3.rs-1509234/v3

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

Version of Record: A version of this preprint was published at Plant Growth Regulation on May 8th, 2023.
See the published version at https://doi.org/10.1007/s10725-023-01005-0.

Abstract

Glycyrrhiza uralensis Fisch. is prescribed as one of the original plants of licorice in Chinese Pharmacopoeia. This herbal medicine possesses numerous important pharmacological activities and has been used in clinic in China since ancient times. Glycyrrhizic acid (GA) is a triterpenoid compound isolated from G. uralensis and also one of the marker components for its quality control. Based on our pervious transcriptome study, three genes, the caffeic acid 3-O-methyltransferase gene (COMT), the β-carotene 3-hydroxylase gene (CRTZ), and the flavonoid 3'-monooxygenase gene (F3'H), were selected as our target genes due to a high correlation of their expression levels with GA biosynthesis. In this study, we investigated the regulatory effects of these genes on GA biosynthesis through gene editing and overexpression in G. uralensis hairy roots. We observed that neither knockout nor overexpression of any of the genes affects the viability of the transgenic hairy roots, indicating that these genes are not essential for survival of hairy roots. However, compared with the wild type and negative control hairy roots, GA contents were significantly lower in hairy roots overexpressing COMT, CRTZ, and F3'H, while higher in those knocking out the three genes. Our findings demonstrate that the three genes, COMT, CRTZ, and F3'H, all negatively regulate the GA biosynthesis.

Key Message

The *COMT, CRTZ, and F3'H* genes negatively regulate glycyrrhizic acid biosynthesis in *Glycyrrhiza uralensis*.

Introduction

Glycyrrhiza uralensis Fisch. is prescribed by *Chinese Pharmacopoeia* as one of the original plants of licorice (National Pharmacopoeia Commission 2020). It is the most commonly-used Chinese herb and has been applied in the treatment of lung diseases (Chen et al. 2018a) and diabetes (Yang et al. 2020) since ancient China. In recent years, modern pharmacological studies have yielded significant insights into the bioactive properties of this ancient Chinese herb. *G. uralensis* has been confirmed to contain many therapeutic activities, such as anti-inflammatory (Huang et al. 2019; Ko et al. 2021), anti-microbial (Yang et al. 2018; Zeng et al. 2021), antiviral (Yi et al. 2022), anticancer (Jain et al. 2022; Wen et al. 2021; Xiao et al. 2011), and immunoregulatory activities (Aipire et al. 2017; Bhattacharjee et al. 2012). Numerous bioactive compounds, including triterpenoids and flavonoids, have been isolated and identified from *G. uralensis*. Among them, glycyrrhizic acid (GA), a pentacyclic triterpenoid, is the most studied bioactive compound of the herb due to its excellent pharmacological activities. It has been used to treat inflammation-associated diseases in clinic (Chu et al. 2020; Xie et al. 2015). Importantly, GA is stipulated by *Chinese Pharmacopoeia* as the quality standard compound in *G. uralensis* (National Pharmacopoeia Commission 2020). *G. uralensis* with a GA content ≥ 2.0% is considered to be qualified as medicinal materials. Thus, uncovering the biosynthetic pathways of GA and elucidating the underlying regulatory mechanisms are important for improving the medicinal quality of this most commonly-used herb.

The biosynthesis of bioactive ingredients in *G. uralensis* depends on a complex metabolic network. As shown in Fig. 1, the primary metabolites derived from glycolysis flow into different pathways to produce the two major bioactive components in *G. uralensis*: terpenoids and flavonoids. These pathways crisscross and interact with each other. Dimethylallyl diphosphate (DMAPP) and isopentenyl diphosphate (IPP), the common precursors of terpenoids, are synthesized through two independent and distinct pathways, the 2-C-methyl-D-erythritol 4-phosphate (MEP) pathway and the mevalonic acid (MVA) pathway (Tian et al. 2022). DMAPP and IPP, the precursors of GA production, are supplied by the MVA pathway (Liao et al. 2016), whereas the equivalent two isomeric five-carbon precursors for the synthesis of tetraterpenoids, such as *β*-carotene and zeaxanthin, are provided by the MEP pathway (Banerjee and Sharkey 2014). In addition, some primary metabolites flow into the shikimate pathway to produce various flavonoids and lignin (Vogt 2010). The aforementioned pathways are intertwined with each other and influence the secondary metabolism in *G. uralensis*. Our previous transcriptome study demonstrated that the glycolysis, the MVA pathway, the MEP pathway, and the phenylpropanoid metabolic pathway were all tightly correlated with GA accumulation, and from these pathways some core differentially expressed genes (DEGs) were identified (Gao et al. 2020) and their functions were investigated (Wang et al. 2021; Yin et al. 2020; Yang et al. 2021; Zhang et al. 2021). In this study, three core DEGs including the *β*-carotene 3-hydroxylase gene (*CRTZ*), the caffeic acid 3-O-methyltransferase gene (*COMT*), and the flavonoid 3'-monooxygenase gene (*F3'H*), were selected as our target genes. Their expression levels correlate inversely with GA contents in our transcriptome study (Gao 2019; Gao et al. 2020). The aim of this study is to evaluate the regulatory effects of these three genes on GA production.

CRTZ, a non heme 2 iron monooxygenase, is a key rate-limiting enzyme involved in the synthesis of tetraterpenoids (Li et al. 2015). It is a multifunctional enzyme and converts *β*-carotene into *β*-cryptoxanthin and *α*-carotene into lutein (Liang et al. 2020). Current studies of CRTZ are mainly focused on its roles in the photoprotection and draught resistance of crops (Davison et al. 2002; Du et al. 2010). COMT mainly involves in the biosynthesis of monolignols. As a multiple-substrate enzyme, it is able to catalyze the methylation of caffealdehyde, caffeyl alcohol, caffeic acid, 5-hydroxmethylyferulic acid ester, 5-hydroxyconiferyl alcohol, and 5-hydroxyconiferaldehyde in the phenylpropanoid metabolic pathway, and eventually influences the composition and content of lignin (Weng et al. 2011). F3'H belongs to cytochrome P450 monooxygenase family and takes part in the hydroxylation of flavonoids (Tanaka and Brugliera 2013). In *Camellia nitidissima* Chi., overexpression of *F3'H* significantly increases the contents of various polyphenol compounds and total flavonoids (Jiang et al. 2021). It also contributes to the adaptation of *Pohlia nutans* in the south pole through regulation of oxidation tolerance and abscisic acid sensitivity (Liu et al. 2021). To our knowledge, none of the three genes has been cloned from *G. uralensis* and their functions in secondary metabolism in *G. uralensis* are still far from clear. Hence potential roles of *CRTZ*, *COMT*, and *F3'H* on GA accumulation is worthy of study.

Hairy root culturing is a technology which combines tissue culturing with genetic transformation. Hairy roots are easily manageable and fast proliferating. Although they are not the real roots of plants, they are able to produce various secondary metabolites. Also, for medicinal plants, hairy roots provide an effectual way to protect their resources by biosynthesizing invaluable active ingredients *in vitro* (Chen et

al. 2018b; Gantait and Mukherjee 2021; Srivastava and Srivastava 2007). In addition, hairy roots are genetically stable than cell suspension cultures. Therefore, they have also been applied as an effective method to study gene expression and function in various plant species (Huber et al. 2009; Ma et al. 2015; Nguyen et al. 2021; Shi et al. 2019). Similar research has also been carried out in *G. uralensis*. In our previous studies, we used hairy roots to identify more than ten functional genes, such as auxin-responsive protein IAA gene (*ARPI*), UDP-Galactose/Glucose-4-epimerase gene (*UGE*), *β*-amyrin synthase gene (*β-AS*), chalcone synthase gene (*CHS*), and chalcone isomerase gene (*CHI*), involved in the secondary metabolism of triterpenoids and flavonoids in *G. uralensis* (Hou et al. 2021; Wang et al. 2021; Yin et al. 2020). In the present work, we used hairy roots to investigate the effects of *CRTZ*, *COMT*, and *F3'H* on GA biosynthesis. The finding of this work will provide new insights on the functions of *CRTZ*, *COMT*, and *F3'H* in the dynamic biosynthetic network of GA in *G. uralensis.*

Materials And Methods

Plant materials

Plump and healthy *G. uralensis* seeds were selected and surface-sterilized with 1% mercury bichloride and 75% ethanol solutions, and then planted on Murashige and Skoog (MS) medium for about 2 weeks. The seedlings were identified based on the internal transcribed spacer (ITS) sequences as described in our previous study (Yang et al. 2017). Next, the correct asepsis seedlings were used as the materials for cloning target genes and also as explants for subsequent infection experiment.

Target genes cloning

Total RNA was extracted from the above asepsis seedlings using an All-In-One DNA/RNA Mini-Preps Kit (Sangon Biotech (Shanghai) Co., Ltd., China). cDNA was converted using a First Strand cDNA Synthesis Kit RT0212-01 (Biomiga. Inc., China). Primer pairs used for amplifying *CRTZ*, *COMT*, or *F3'H* were designed by Primer-BLAST and listed in Table S1. The PCR parameters for amplifying *COMT* were as follows: 94 °C 5 min, 35 cycles (94 °C 30 s, 52 °C 30 s, 72 °C 65 s), and 72 °C 5 min. The PCR programs for amplifying *CRTZ* were 95 °C 5 min, 35 cycles (94 °C 30 s, 54.5 °C 30 s, 72 °C 60 s), and 72 °C 5 min, and for amplifying *F3'H* were 94 °C 5 min, 35 cycles (94 °C 30 s, 52.5 °C 30 s, 72 °C 60 s), and 72 °C 5 min. The amplified fragments were cloned into pMD^{TM} 19-T Cloning Vector (Takara Biomedical Technology (Beijing) Co., Ltd., China), and sequenced by Sangon Biotech (Shanghai) Co., Ltd., China.

Construction of plant binary expression vectors and CRISPR/Cas9 vectors

We constructed plant binary vectors expressing *CRTZ*, *COMT*, or *F3'H* as shown in Fig. S1a. Three exogenous genes were inserted into pCAMBIA1305.1 between the restriction enzyme cutting sites *Spe* I and *Bgl* II. The specific primers used for amplifying target genes containing homologous sequences of pCAMBIA1305.1 are listed in Table S2. The PCR parameters were as same as those mentioned in the above section. Target genes were ligated with linearized pCAMBIA1305.1 using a BM effusion kit (Beijing Biomed Medical Technology Co., Ltd., Beijing, China) to yield recombinant plasmids, pCA-COMT, pCA-

CRTZ, and pCA-F3'H. We also constructed plant CRISPR/Cas9 vectors for knocking out *CRTZ*, *COMT*, or *F3'H* as shown in Fig. S1b. The sgRNA sequences were designed and listed in Table S3. They were inserted into linearized pHSE401 between the two *Bsa* I sites by T4 ligase to construct the CRISPR/Cas9 plasmids, pHSE-COMT, pHSE-CRTZ, and pHSE-F3'H. The above recombinant plasmids were all electrotransfered into *Agrobacterium rhizogenes* ATCC15834. Also, the empty pCAMBIA1305.1 and pHSE401 vectors without exogenous genes were introduced into *A. rhizogenes* ATCC15834. Next, the positive colonies were selected on Tryptone-Yeast extract (TY) plates containing kanamycin (50 mg•L^{-1}). PCR and sequencing were performed to identify the correct vectors and strains.

Establishment of *G. uralensis* hairy root lines

The above recombinant *A. rhizogenes* were cultured in liquid TY medium containing 50 mg•L^{-1} kanamycin and collected in logarithmic phase, then resuspended with liquid 6,7-V medium. The *G. uralensis* asepsis seedlings were cut and the hypocotyls and cotyledons were used as explants. They were wounded by scalpel and soaked in the above resuspended cultures for 25 min to complete the infection. Next, the infected explants were cultured on 6,7-V plates at 25 °C for 2 days, and then transferred onto 6,7-V plates containing cefotaxime sodium (Cef) (500 mg·L^{-1}). When hairy roots were ~ 5 cm, they were cut from explants and trans-cultured on fresh 6,7-V plates containing Cef. Totally, we induced nine groups of *G. uralensis* hairy roots using different recombinant *A. rhizogenes* strains. The wild type (WT) was induced by normal *A. rhizogenes* ATCC15834. The negative control hairy root lines were induced by *A. rhizogenes* strains containing empty pHSE401 or pCAMBIA1305.1, referred as NC-PHSE or NC-PCA, respectively. The hairy root lines overexpressing *COMT*, *CRTZ*, or *F3'H*, referred as $COMT^+$, $CRTZ^+$, or $F3'H^+$, were induced by recombinant *A. rhizogenes* strains containing pCA-COMT, pCA-CRTZ, or pCA-F3'H, respectively. The hairy root lines knocking out *COMT*, *CRTZ*, or *F3'H*, referred as $COMT^-$, $CRTZ^-$, or $F3'H^-$, were induced by recombinant *A. rhizogenes* strains containingpHSE-COMT, pHSE-CRTZ, or pHSE-F3'H, respectively. Hairy roots were cultured on 6,7-V plates, and gradient concentration of Cef was used to eliminate the residual *A. rhizogenes*.

Identification of *G. uralensis* hairy root lines

Genomic DNA of *G. uralensis* hairy root lines were extracted by All-In-One DNA/RNA Mini-Preps Kit (Sangon Biotech (Shanghai) Co., Ltd., China) and used as PCR templates. *rol*C gene is a specific marker gene in hairy roots. We amplified *rol*C genes in all the *G. uralensis* hairy root lines as described in our previous study (Wang et al. 2021). For identification of the $COMT^+$, $CRTZ^+$, and $F3'H^+$ hairy root lines, we respectively amplified the corresponding target genes using the above-mentioned PCR primers (Table S1) and programs (in Section "Target genes cloning"). To analyze the gene editing sites in the $COMT^-$, $CRTZ^-$, and $F3'H^-$ hairy root lines, we designed specific primers (Table S4) to respectively amplify the corresponding exons of *COMT*, *CRTZ*, and *F3'H*. The PCR cycling parameters were as follows: 95 °C for 5 min, 32 cycles of 95 °C for 30 s, 55.5 °C for 50 s, 72 °C for 60 s, and 72 °C for 5 min. PCR products were cloned into pMD^{TM} 19-T Cloning Vector (Takara Biomedical Technology (Beijing) Co., Ltd., China) and

selected on LB plates supplemented with ampicillin (50 mg·L^{-1}). Single colonies were randomly picked for sequencing analysis which was performed by Sangon Biotech (Shanghai) Co., Ltd., China.

Gene expression levels analysis

Real-time quantitative PCR (RT-qPCR) was carried out with *β-actin* gene as the internal control to analyze the expression levels of *COMT*, *CRTZ*, or *F3'H* in hairy root lines overexpressing the three genes. Primers for RT-qPCR are listed in Table S5. Total RNA of the *COMT*$^{+}$, *CRTZ*$^{+}$, and *F3'H*$^{+}$ lines were extracted and reversely transcribed into cDNA as mentioned above (in Section "Target genes cloning"). RT-qPCR was carried out using an Agilent StrataGene Mx3005P QPCR (Agilent Technologies, USA) with a program as follows: 95 ℃ 3 min, 45 cycles (95 ℃ 7 sec, 57 ℃ 10 sec, 72 ℃ 15 sec). The expression level differences were evaluated by the $2^{-\Delta\Delta Ct}$ method in triplicate (Annaratone et al. 2013).

Suspension culture of hairy roots and samples preparation

The validated hairy roots were isolated from each line and trans-cultured in liquid 6,7-V medium (250 mL-conical flasks, 110 rpm, 25 °C, dark). After hairy root lines adapted to the liquid culture system and thrived, the healthy branches (2.0 g) were cut in triplicate and suspension cultured in fresh 6,7-V medium at the same conditions for three weeks. Next, the collected hairy roots were dried at 65 °C and powdered. The contents of GA, lignin, and total flavonoids in these powder samples were examined in the next steps.

GA content assay by UPLC

The UPLC method for determination of the GA contents in *G. uralensis* hairy roots established in our previous study was adopted (Wang et al. 2021). The standard compound GA with a purity of 99.45% was purchased from China National Institutes for Food and Drug Control. The stock solution of GA was prepared with 50% methanol at a concentration of 0.0904 mg mL^{-1}. The linear curve was evaluated at seven points by gradient dilution of the GA stock solution to 0.09040, 0.07232, 0.05424, 0.04520, 0.01808, 0.00904, and 0.00452 mg mL^{-1}. 100 mg of each hairy root powder sample was extracted with 50 mL 50% methanol using ultrasonic for 30 min (40 kHz, 500 W), and filtered with 0.45 μm filter membranes. The chromatographic conditions were as follows: chromatographic instrument: Waters Acquity UPLC system (Waters Corporation, Milford, Massachusetts, USA), chromatographic column: Waters UPLC BEH C_{18} column (2.1 mm × 100 mm, 1.7 μm), UV detection wavelength: 250 nm, column temperature: 40 °C, flow rate: 0.3 mL·min^{-1}, and injection volume: 1 μL. The mobile phase A was acetonitrile and B was 0.05% phosphoric acid. The UPLC gradient elution program was shown in Table S6.

Lignin content assay by UV spectrophotometry

To evaluate the effects of *COMT* overexpression or knockout on the lignin metabolism, we used an UV spectrophotometry method to detect the lignin contents in the WT, NC-PCA, NC-PHSE, *COMT*$^{+}$, and *COMT*$^{-}$ hairy root lines. The lignin contents were measured in triplicate using the lignin assay kit (Beijing Solarbio

Science & Technology Co., Ltd, Beijing, China) at 280 nm on an Epoch microplate reader (Bio Tek Instruments Inc, USA).

Total flavonoids determination by UV spectrophotometry

To evaluate the effects of *F3'H* over-expression or knockout on the flavonoid metabolism, we detected the total flavonoid contents in the WT, NC-PCA, NC-PHSE, *F3'H*$^{+}$, and *F3'H*$^{-}$ hairy root lines according to a previous study (Dang et al. 2021). The standard compound liquiritin with a purity of 98.50% was purchased from China National Institutes for Food and Drug Control. The stock solution of liquiritin was prepared with methanol at a concentration of 0.152 mg mL^{-1}. The linear curve was evaluated at six points by gradient dilution of the liquiritin stock solution to 0.0019, 0.0038, 0.0076, 0.0152, 0.0304, and 0.0608 mg mL^{-1}. 20 mg of each hairy root powder sample was extracted with 10 mL methanol using ultrasonic for 80 min (20 kHz, 250 W) and filtered. Then, 0.5 mL of the subsequent filtrate was transferred into a 10 mL volumetric flask, and 1 mL methanol and 0.5 mL 10% KOH were incorporated into the flask for chromogenic reaction at room temperature for 5 min. After the reaction, methanol was added into the flask to a final volume of 10 mL. Next, the absorbance at 334 nm was measured on an Epoch microplate reader (Bio Tek Instruments Inc, USA) with methanol which had been treated accordingly as the blank.

Statistical and bioinformatics analysis

IBM SPSS Statistics 20 was applied to perform the statistical analysis. BLAST (http://www.ncbi.nlm.nih.gov) and DNAMAN 6.0.3.99 were used to analyze the sequencing results. MEGA 11 was applied to construct the phylogenetic tree using Maximum Composite Likelihood method.

Results

Identification of three target genes

As shown in Fig. S2a, three fragments, the length of which were about 1200, 1000, and 1000 bp, were obtained by PCR. Sequencing results confirmed that the exact full length of the three PCR products were 1140, 930, and 993 bp, respectively. The 1140 bp-PCR product had 86.71% identity with the *Abrus precatorius COMT* (GenBank accession No. XM_027512828.1), the 930 bp-PCR product had 89.82 % identity with the *Glycine max CRTZ* (NM_001254504.1), and the 993 bp-PCR product had 81.82% identity with *Cajanus cajan F3'H* (XM_020357268.2), indicating that the *COMT*, *CRTZ*, and *F3'H* sequences were obtained from *G. uralensis*. We registered these sequences in GenBank with the accession numbers as follows: *COMT* (MZ169549), *CRTZ* (MZ169550), and *F3'H* (MZ169551). It is the first time that the *COMT*, *CRTZ*, and *F3'H* genes are identified from *G. uralensis.* To evaluate the evolution of these genes, we then established phylogenetic trees based on COMT, CRTZ, and F3'H homologous sequences registered in NCBI, respectively. As shown in Fig. 2, COMT and F3'H homologous sequences both clustered into three major independent clades, Dicotyledoneae, Monocotyledoneae, and Pteridophyta, while CRTZ orthologs clustered into four, Dicotyledoneae, Monocotyledoneae, Pteridophyta, and Bryophyta. These findings suggest that the evolution of the three enzymes is conserved.

Construction of recombinant vectors

As shown in Fig. S2b, three fragmentswere amplified from the plant binary expression vectors, pCA-COMT, pCA-CRTZ, and pCA-F3'H, respectively. Sequencing results further confirmed that these PCR products were completely identical to *COMT* (MZ169549), *CRTZ* (MZ169550), and *F3'H* (MZ169551)cloned from *G. uralensis*. These findings demonstrated that the plant binary expression vectors for overexpressing *COMT*, *CRTZ*, or *F3'H* were correct. As shown in Fig. S2c, several 400 bp PCR products were obtained from the plant CRISPR/Cas9 vectors, pHSE-COMT, pHSE-CRTZ, and pHSE-F3'H. Sequencing results confirmed that these fragments contained the corresponding sgRNA sequences of *COMT*, *CRTZ*, or *F3'H* genes, indicating that the CRISPR/Cas9 vectors for knocking out these genes were constructed successfully.

Identification of *G. uralensis* hairy root lines

We generated nine groups of *G. uralensis* hairy root lines, including WT, NC-PCA, NC-PHSE, $COMT^{+}$, $CRTZ^{+}$, $F3'H^{+}$, $COMT^{-}$, $CRTZ^{-}$, and $F3'H^{-}$. Fig. 3 shows the growth situation of these hairy roots after the onset of induction for 10, 20, and 30 days. The roots of the transgenic lines grew at a rate similar to that of WT. To verify these hairy root lines, we firstly cloned the *rol*C genes, the signature gene in hairy roots, from all lines and got 600 bp-fragments by PCR (Fig. S3a), which were identified to have 100% identity with *rol*C (GenBank accession No. DQ160187.1). Also, the fragments amplified from the $COMT^{+}$, $CRTZ^{+}$, and $F3'H^{+}$ lines shown in Fig. S3b were confirmed with sequence analyses. Fig. S3c and S3d display the third exons of *COMT* and *CRTZ*amplified from the $COMT^{-}$ and $CRTZ^{-}$ lines respectively, and Fig. S3e shows the first exons of *F3'H* amplified from the $F3'H^{-}$ lines. These fragments, as described in next section, will be used to analyze the exact editing sites of target genes in hairy root lines of $COMT^{-}$, $CRTZ^{-}$, and $F3'H^{-}$ through further cloning and sequencing. In the end, one WT, five $COMT^{+}$, five $CRTZ^{+}$, seven $F3'H^{+}$, three $COMT^{-}$, five $CRTZ^{-}$, four $F3'H^{-}$,one NC-PCA, and one NC-PHSE lines were obtained.

Analyses of gene editing and expression levels in transgenic *G. uralensis* hairy root lines

Further cloning and sequencing analyses confirmed that the *COMT* gene was edited in four hairy root lines ($COMT^{-}$-2,-5,-6, and -8) out of nine with a gene editing efficiency of 44.4%. The gene editing details are shown in Fig. 4a and the mutation sites present in COMT amino acid sequences are listed in Table 1. We observed a synonymous mutation in the $COMT^{-}$-2 line and missense mutations in the other three lines. The most efficient gene edition was a fragment deletion in the $COMT^{-}$-5 line, which results in a termination codon "TAA" in this position.

CRTZ gene was edited in eight lines ($CRTZ^{-}$-1, -2,-3,-4,-5,-6,7, and-8) out of ten with an editing efficiency of 80%. Among them, homozygous mutations were present in four lines, $CRTZ^{-}$-1,-3,-4 and-6 (Fig. 4b), while heterozygous mutations in the other four (Fig. S4b). Frameshifts were observed in all lines. The mutation sites in the CRTZ amino acid sequences are listed in Table 2. The CRTZ sequences in the $CRTZ^{-}$-1

and *CRTZ*$^{-}$-6 lines are the same, both containing two mutation sites. The CRTZ sequence in the *CRTZ*$^{-}$-3 line is identical to that in the *CRTZ*$^{-}$-4 line with a single mutation site. All mutation sites in these four lines were located within the functional domain of CRTZ.

Sequencing analyses showed that the *F3'H* gene was edited in seven lines (*F3'H*$^{-}$-2, -4, -5, -7, -8, -9, and -10) out of eleven with a gene editing efficiency of 63.6%. Four hairy root lines, *F3'H*$^{-}$-2, -7, -8, and -10, were homozygous mutants (Fig. 4c), while the other three lines were heterozygous mutants (Fig. S4a). The mutation sites present in the F3'H amino acid sequence are listed in Table 3. The F3'H sequence in the *F3'H*$^{-}$-2 line is identical to that in the *F3'H*$^{-}$-10 line, containing 16 mutations. All these mutations are located within the functional domain of F3'H.

We next examined the gene expression levels of the three genes in newly constructed *COMT*$^{+}$, *CRTZ*$^{+}$, and *F3'H*$^{+}$ hairy root lines. As illustrated in Fig. 4d, the relative expression levels of *COMT*, *CRTZ*, and *F3'H* in the *COMT*$^{+}$, *CRTZ*$^{+}$, and *F3'H*$^{+}$ hairy root lineswere all remarkably higher than that in WT. In particular, samples *COMT*$^{+}$-4, *CRTZ*$^{+}$-4, and *F3'H*$^{+}$-6 showed the highest expression level in their respective groups.

GA content assay by UPLC in *G. uralensis* hairy root samples

Fig. 5a shows the hairy root samples cultured in liquid 6,7-V medium. Fig. 5b shows the collected hairy root samples, which were prepared for UPLC analyses. It is worth noting that the roots of the *F3'H*$^{-}$ and *CRTZ*$^{+}$ lines had a whitish appearance and those of the *CRTZ*$^{-}$ lines appeared reddish. The hairy roots of other lines were whitish with a yellow undertone. The UPLC chromatograms of GA in the WT, NC-PCA, NC-PHSE, *COMT*$^{+}$, *CRTZ*$^{+}$, *F3'H*$^{+}$, *COMT*$^{-}$, *CRTZ*$^{-}$, and *F3'H*$^{-}$ lines as well as the GA standard are shown in Fig. 6a. The UPLC retention time of GA was 6.532 min and the standard curve was expressed as: Y = 2,683,455.37 X - 1,660.17 (R^2 = 0.9999, X: the GA content (mg·mL^{-1}), Y: the peak area (mAU*s)). The GA contents in all of the hairy root lines are calculated and listed in Table S7. The differences of GA contents among different lines are shown in Fig. 6b and those among different groups are shown in Fig. 6c. We found that the GA contents in negative control (NC-PCA and NC-PHSE) were comparable to that in WT. Interestingly, we noticed an inverse correlation between the expression levels of the three genes and the GA contents in the hairy root lines. For example, the GA contents in all the seven *F3'H*$^{+}$ lines were significantly lower than those in the WT and NC-PCA lines, while the GA contents in the *F3'H*$^{-}$ lines (*F3'H*$^{-}$-2, -8, and -10), except *F3'H*$^{-}$-7, were all significantly higher than those in the WT and NC-PHSE lines. Similar variations in GA contents were also observed in the *COMT*$^{+}$, *CRTZ*$^{+}$, *COMT*$^{-}$, and *CRTZ*$^{-}$ groups. As shown in Table S7, the GA contents in the *F3'H*$^{+}$ (average value: 2.3654 mg·g^{-1} dry weight (DW)), *CRTZ*$^{+}$ (2.2050 mg·g^{-1} DW), and *COMT*$^{+}$ (2.3509 mg·g^{-1} DW) groups were significantly lower than those in both WT (3.3242 mg·g^{-1} DW) and NC-PCA (3.5188 mg·g^{-1} DW). While the GA contents in the *F3'H*$^{-}$ (5.4744 mg·g^{-1} DW), *CRTZ*$^{-}$ (5.3651 mg·g^{-1} DW) and *COMT*$^{-}$ (4.2746 mg·g^{-1} DW) groups were significantly higher

than those in WT and NC-PHSE (3.3459 mg·g^{-1} DW). These findings suggest that the *COMT*, *CRTZ*, and *F3'H* genes negatively regulate GA production.

The lignin and total flavonoids contents analyses

The lignin contents in the hairy roots of the WT, negative controls, *COMT*$^{+}$, and *COMT*$^{-}$ lines were examined using a lignin assay kit. As shown in Fig. 7a and Table S8, we found that the lignin contents in all the three *COMT*$^{-}$ lines were significantly lower than those in the WT and NC-PHSE lines, whereas the lignin contents in all the five *COMT*$^{+}$ lines were significantly higher than those in the WT and NC-PCA lines. The lignin contents in the *COMT*$^{-}$ group (average value: 37.23 mg·g^{-1} DW) were remarkably lower than those in the *COMT*$^{+}$ group (72.23 mg·g^{-1} DW). These findings suggest that *COMT* positively regulates the biosynthesis of lignin.

Fig. 7b and Table S9 show the total flavonoid contents in the hairy roots of the WT, negative controls, *F3'H*$^{+}$, and *F3'H*$^{-}$ lines. We constructed a standard curve expressed as: $Y = 3.9146\ X - 0.0003$ (R^2 = 0.9979) using *X* axis for the total flavonoid contents (mg·mL^{-1}) and *Y* axis for the relative absorbance value. We found that the total flavonoid contents in three *F3'H*$^{-}$ lines (*F3'H*$^{-}$-2, -8, and -10) were significantly lower than those in the WT and NC-PHSE lines, whereas the flavonoid contents in all the seven *F3'H*$^{+}$ lines were significantly higher than those in the WT and NC-PCA lines. The total flavonoid contents in the *F3'H*$^{-}$ group (average value: 31.70 mg·g^{-1} DW) were significantly lower than those in the *F3'H*$^{+}$ group (51.00 mg·g^{-1} DW). These findings indicate that overexpression of *F3'H* promotes whereas knockout of *F3'H* inhibits the biosynthesis of flavonoids.

Discussion And Conclusion

In the present study, we established transgenic *G. uralensis* hairy root lines to investigate the effects of the *CRTZ*, *COMT*, and *F3'H* genes on GA accumulation by overexpression or knockout of these genes. By UPLC analyses, we found that in comparison with WT and negative control, the GA contents in the hairy roots overexpressing anyone of the three genes were all significantly lower, while the GA contents in those with anyone of the genes knocked out were higher, indicating negative effects of *CRTZ*, *COMT*, and *F3'H* on GA biosynthesis.

The major secondary metabolites in *G. uralensis* include terpenoids, flavonoids, and lignin. As shown in Fig. 1, GA, a triterpenoid, is biosynthesized through the MVA pathway. Carotenoids which belong to tetraterpenoids are produced by the MEP pathway. Lignin and flavonoids are downstream products of the phenylpropanoid metabolic pathway. CRTZ, COMT, and F3'H are key enzymes involved in the synthesis of carotenoids (Li et al. 2015), lignin (Saluja et al. 2021), and flavonoids (Tanaka and Brugliera 2013) in *G. uralensis*, respectively. There is a competition among the three pathways for production of these secondary metabolites in *G. uralensis*. Several lines of evidence have demonstrated that overexpression of *CRTZ* increases the utilization of *β*-carotene (Pollmann et al. 2017), while knockout of *CRTZ* has an opposite effect (Tomlekova et al. 2021). It is thus expected that an accumulation of *β*-carotene by the

CRTZ knockout leads to a feedback inhibition of the MEP pathway, which in turn promotes the MVA pathway and the production of GA in *G. uralensis*. This notion is supported by the color changes of the hairy roots of the *CRTZ*$^{-}$ and *CRTZ*$^{+}$ lines. We find that the color of the hairy roots of the *CRTZ*$^{-}$ lines looks reddish, indicating a high content of *β*-carotene, whereas the color of the hairy roots of the *CRTZ*$^{+}$ lines appears pale, suggesting a low content of *β*-carotene. It has also been observed in *Brassica rapa* and *Oncidium hybridum* that down-regulation of *CRTZ* causes the accumulation of *β*-carotene and makes the flower color to turn from yellow to orange (Chiou et al. 2010; Zhang et al. 2019). It is thus likely that knocking out *CRTZ* increases the GA contents through promoting the accumulation of *β*-carotene in *G. uralensis* hairy roots.

COMT is an essential enzyme which catalyzes the methylation of multiple substrates in the biosynthesis of lignin (Saluja et al. 2021). Since the intermediate products are unstable and their contents are constantly changing, we analyzed the content of the end metabolite of the pathway. We found that over-expression of *COMT* promoted whereas knockout of *COMT* inhibited the production of lignin in *G. uralensis* hairy roots. Since acetyl-CoA (the substrate for producing GA) and phosphoenolpyruvate (the substrate for producing lignin) are both derived from the metabolites of glycolysis, the increase of lignin is bound to suppress the carbon flux in the MVA pathway, which may count for the negative effect of *COMT* on GA accumulation.

F3’H is an important multiple-substrate enzyme involved in the biosynthetic pathway of flavonols (Tanaka and Brugliera 2013). In order to evaluate the function of F3’H, we detected the contents of total flavonoids in hairy root samples overexpressing or knocking out *F3’H*. We found that over-expression of *F3’H* promoted whereas knockout of *F3’H* inhibited the production of total flavonoids in *G. uralensis* hairy roots. The increase of total flavonoids caused by over-expression of *F3’H* consumes more phosphoenolpyruvate, leading to the reduction of acetyl-CoA, the substrate of the MVA pathway. The reduction of acetyl-CoA may count for the negative effect of *F3’H* on GA accumulation.

Although the GA contents differ greatly between different groups, within the same group, the contents in different lines also vary. This phenomenon is mainly observed in groups knocking out the three genes. For example, in the *F3’H*$^{-}$ group the GA contents in three lines (*F3’H*$^{-}$-2, -8, and − 10) are all significantly higher than those in WT and NC-PHSE. However, the *F3’H*$^{-}$-7 line is an exception. We believe that the variations in the GA contents among the *F3’H*$^{-}$ lines are caused by the different gene editing effects. As shown in Table 3, there is only one mutation at the 56th position causing a substitution of valine to isoleucine in the F3’H sequence of *F3’H*$^{-}$-7 line, while multiple sites are mutated in the other three lines, which may aggravate the dysfunction of F3’H. This notion is verified by the total flavonoid contents analysis. We observed that the total flavonoid contents in the *F3’H*$^{-}$-2, -8, and − 10 lines were significantly decreased, but that in the *F3’H*$^{-}$-7 line was unaffected (Fig. 7b). Also, in the *COMT*$^{-}$ group the GA content of the *COMT*$^{-}$-5 line was significantly higher than those in WT and NC-PHSE. In the other two lines (*COMT*$^{-}$-6 and 8), although the GA contents were also higher, the differences were not remarkable. We analyzed the gene editing in these three lines and found that the mutation in *COMT* cDNA sequence of

the *COMT*$^{-}$-5 line created a premature termination codon, which makes the translation of COMT stop at the 160th position, leading to a complete loss of COMT function (Fig. 4a and Table 1). In the COMT sequence of the *COMT*$^{-}$-6 line, there is only one mutation at the 174th position causing a substitution of methionine to histidine and in that of the *COMT*$^{-}$-8 line only one mutation at the 219th position causing a substitution of valine to alanine. These mutations are expected to affect the function of COMT, but not as serious as the mutation in the *COMT*$^{-}$-5 line that causes a premature termination of the COMT translation. This notion is supported by the lignin content analysis of the *COMT*$^{-}$ lines. As shown in Fig. 7a, although the lignin contents in the three *COMT*$^{-}$ lines are all reduced, that in the *COMT*$^{-}$-5 line is the lowest. Also, there is a significant difference in lignin contents between the *COMT*$^{-}$-5 line and the other two lines ($p < 0.001$). In addition, the GA content differences among the *CRTZ*$^{-}$ lines might also due to the variations in gene editing effects (Fig. 4b and Table 2).

To sum up, this is the first time that *CRTZ*, *COMT*, and *F3'H* genes are identified from *G. uralensis.* Our findings demonstrate that the three genes all negatively regulate GA production in *G. uralensis*. However, the relationships of these genes in secondary metabolic pathways are still far from clear and deserve further investigation. In addition, we will continue evaluate the functions of other DEGs to establish a gene regulatory network responsible for secondary metabolism in *G. uralensis.*

Abbreviations

ARPI, Auxin-responsive protein IAA; ***β*-AS**, *β*-amyrin synthase; **Cef**, cefotaxime sodium; **CHI**, Chalcone isomerase; **CHS**, Chalcone synthase; ***COMT***, Caffeic acid 3-O-methyltransferase gene; ***CRTZ***, *β*-carotene 3-hydroxylase gene; **DEGs**, Differentially expressed genes; **DMAPP**, Dimethylallyl diphosphate; **WD**, dry weight; ***F3'H***, Flavonoid 3'-monooxygenase gene; **GA**, Glycyrrhizic acid; **IDI**: Isopentenyl diphosphate isomerase; **IPP**, Isopentenyl diphosphate; **ITS**, Internal transcribed spacer; **M**, Missense mutations; **MEP**, 2-C-methyl-D-erythritol 4-phosphate; **MS**, Murashige and Skoog; **MVA**, Mevalonic acid; **RT-qPCR**, Real-time quantitative PCR; **S**, Synonymous mutations; **TY**, Tryptone-Yeast extract; **UGE**, UDP-Galactose/Glucose-4-epimerase; **WT**, Wild type

Declarations

Funding: Not applicable.

Conflicts of interest: The authors declare that there is no conflict of interest regarding the publication of this paper.

Availability of data and material: Not applicable.

Code availability: Not applicable.

Authors' contributions: YL conceived and designed the experiments. ZXZ, WWD and ZYC performed the experiments. WPX and DDW helped in analyzing the relevant data and searching literature. YL and TGL wrote the manuscript.

Ethics approval: Not applicable.

Consent to participate: Not applicable.

Consent for publication: Not applicable.

Acknowledgement: We would like to thank Dr. Yu Jiang at University of Pittsburgh for critically reading the manuscript.

References

1. Aipire A, Li JY, Yuan PF, He J, Hu YL, Liu L, Feng XL, Li YJ, Zhang FC, Yang JH, Li JY (2017) *Glycyrrhiza uralensis* water extract enhances dendritic cell maturation and antitumor efficacy of HPV dendritic cell-based vaccine. Sci Rep 7:43796. https://doi.org/10.1038/srep43796
2. Annaratone L, Volante M, Asioli S, Rangel N, Bussolati G (2013) Characterization of neuroendocrine tumors of the pancreas by real-time quantitative polymerase chain reaction. A methodological approach. Endocr Pathol 24(2):83–91. https://doi.org/10.1007/s12022-013-9246-y
3. Banerjee A, Sharkey TD (2014) Methylerythritol 4-phosphate (MEP) pathway metabolic regulation. Nat Prod Rep 31(8):1043–1055. https://doi.org/10.1039/c3np70124g
4. Bhattacharjee S, Bhattacharjee A, Majumder S, Majumdar SB, Majumdar S (2012) Glycyrrhizic acid suppresses Cox-2-mediated anti-inflammatory responses during *Leishmania donovani* infection. J Antimicrob Chemother 67(8):1905–1914. https://doi.org/10.1093/jac/dks159
5. Chen CL, Shenoy AK, Padia R, Fang DD, Jing Q, Yang P, Su SB, Huang S (2018a) Suppression of lung cancer progression by isoliquiritigenin through its metabolite 2, 4, 2', 4'-Tetrahydroxychalcone. J Exp Clin Cancer Res 37(1):243. https://doi.org/10.1186/s13046-018-0902-4
6. Chen IJ, Lee MS, Lin MK, Ko CY, Chang WT (2018b) Blue light decreases tanshinone IIA content in *Salvia miltiorrhiza* hairy roots via genes regulation. J Photochem Photobiol B 183:164–171. https://doi.org/10.1016/j.jphotobiol.2018.04.013
7. Chiou CY, Pan HA, Chuang YN, Yeh KW (2010) Differential expression of carotenoid-related genes determines diversified carotenoid coloration in floral tissues of *Oncidium* cultivars. Planta 232(4):937–948. https://doi.org/10.1007/s00425-010-1222-x
8. Chu SF, Niu ZQ, Guo QX, Bi HZ, Li XY, Li FF, Zhang Z, He WB, Cao P, Chen NH, Sun XY (2020) Combination of monoammonium glycyrrhizinate and cysteine hydrochloride ameliorated lipopolysaccharide/galactosamine-induced acute liver injury through Nrf2/ARE pathway. Eur J Pharmacol 882:173258. https://doi.org/10.1016/j.ejphar.2020.173258

9. Dang HL, Zhang T, Wang ZK, Li GF, Zhao WQ, Lv XH, Zhuang L (2021) Differences in the endophytic fungal community and effective ingredients in root of three *Glycyrrhiza* species in Xinjiang. China PeerJ 9:e11047. https://doi.org/10.7717/peerj.11047
10. Davison PA, Hunter CN, Horton P (2002) Overexpression of beta-carotene hydroxylase enhances stress tolerance in *Arabidopsis*. Nature 418(6894):203–206. https://doi.org/10.1038/nature00861
11. Du H, Wang NL, Cui F, Li XH, Xiao JH, Xiong LZ (2010) Characterization of the beta-carotene hydroxylase gene *DSM2* conferring drought and oxidative stress resistance by increasing xanthophylls and abscisic acid synthesis in rice. Plant Physiol 154(3):1304–1318. https://doi.org/10.1104/pp.110.163741
12. Gantait S, Mukherjee E (2021) Hairy root culture technology: applications, constraints and prospect. Appl Microbiol Biotechnol 105(1):35–53. https://doi.org/10.1007/s00253-020-11017-9
13. Gao ZQ (2019) Using RNA-Seq and X-ray irradiation to study the molecular mechanism of triterpenoid biosynthesis in *Glycyrrhiza glabra*. Dissertation, Beijing University of Chinese Medicine
14. Gao ZQ, Tian SK, Hou JM, Zhang ZX, Yang L, Hu T, Li WD, Liu Y (2020) RNA-Seq based transcriptome analysis reveals the molecular mechanism of triterpenoid biosynthesis in *Glycyrrhiza glabra*. Bioorg Med Chem Lett 30(10):127102. https://doi.org/10.1016/j.bmcl.2020.127102
15. Hou JM, Yin YC, Tian SK, Zhang ZX, Yang L, Li WD, Liu Y (2021) Overexpressing of chalcone isomerase(*CHI*) gene enhances flavonoid accumulation in *Glycyrrhiza uralensis* hairy roots. Acta Pharm Sin 56(01):319–327. https://doi.org/10.16438/j.0513-4870.2020-0951
16. Huang WC, Liu CY, Shen SC, Chen LC, Yeh KW, Liu SH, Liou CJ (2019) Protective effects of Licochalcone A improve airway hyper-responsiveness and oxidative stress in a mouse model of asthma. Cells 8(6):617. https://doi.org/10.3390/cells8060617
17. Huber C, Bartha B, Harpaintner R, Schroder P (2009) Metabolism of acetaminophen (paracetamol) in plants–two independent pathways result in the formation of a glutathione and a glucose conjugate. Environ Sci Pollut Res Int 16(2):206–213. https://doi.org/10.1007/s11356-008-0095-z
18. Jain R, Hussein MA, Pierce S, Martens C, Shahagadkar P, Munirathinam G (2022) Oncopreventive and oncotherapeutic potential of licorice triterpenoid compound glycyrrhizin and its derivatives: Molecular insights. Pharmacol Res 178:106138. https://doi.org/10.1016/j.phrs.2022.106138
19. Jiang LN, Fan ZQ, Tong R, Yin HF, Li JY, Zhou XW (2021) Flavonoid 3'-hydroxylase of *Camellia nitidissima* Chi. promotes the synthesis of polyphenols better than flavonoids. Mol Biol Rep 48(5):3903–3912. https://doi.org/10.1007/s11033-021-06345-6
20. Ko HM, Lee SH, Jee W, Jung JH, Kim KI, Jung HJ, Jang HJ (2021) Gancaonin N from *Glycyrrhiza uralensis* attenuates the inflammatory response by downregulating the NF-kappaB/MAPK pathway on an acute pneumonia in vitro model. Pharmaceutics 13(7):1028. https://doi.org/10.3390/pharmaceutics13071028
21. Li XR, Tian GQ, Shen HJ, Liu JZ (2015) Metabolic engineering of *Escherichia coli* to produce zeaxanthin. J Ind Microbiol Biotechnol 42(4):627–636. https://doi.org/10.1007/s10295-014-1565-6

22. Liang MH, Xie H, Chen HH, Liang ZC, Jiang JG (2020) Functional identification of two types of carotene hydroxylases from the green alga *Dunaliella bardawil* rich in lutein. ACS Synth Biol 9(6):1246–1253. https://doi.org/10.1021/acssynbio.0c00070
23. Liao P, Hemmerlin A, Bach TJ, Chye ML (2016) The potential of the mevalonate pathway for enhanced isoprenoid production. Biotechnol Adv 34(5):697–713. https://doi.org/10.1016/j.biotechadv.2016.03.005
24. Liu HW, Liu SH, Wang HJ, Chen KS, Zhang PY (2021) The flavonoid 3'-hydroxylase gene from the Antarctic moss *Pohlia nutans* is involved in regulating oxidative and salt stress tolerance. Biotechnol Appl Biochem 69(2):676–686. https://doi.org/10.1002/bab.2143
25. Ma XH, Ma Y, Tang JF, He YL, Liu YC, Ma XJ, Shen Y, Cui GH, Lin HX, Rong QX, Guo J, Huang LQ (2015) The biosynthetic pathways of Tanshinones and Phenolic acids in *Salvia miltiorrhiza*. Molecules 20(9):16235–16254. https://doi.org/10.3390/molecules200916235
26. National Pharmacopoeia Commission (2020) Chinese Pharmacopoeia, 2020 edn. China Medical Science and Technology Press, Beijing, pp 88–89
27. Nguyen DV, Hoang TT, Le NT, Tran HT, Nguyen CX, Moon YH, Chu HH, Do PT (2021) An efficient hairy root system for validation of plant transformation vector and CRISPR/Cas construct activities in Cucumber (*Cucumis sativus* L.). Front Plant Sci 12:770062. https://doi.org/10.3389/fpls.2021.770062
28. Pollmann H, Breitenbach J, Sandmann G (2017) Engineering of the carotenoid pathway in *Xanthophyllomyces dendrorhous* leading to the synthesis of zeaxanthin. Appl Microbiol Biotechnol 101(1):103–111. https://doi.org/10.1007/s00253-016-7769-0
29. Saluja M, Zhu FY, Yu HF, Walia H, Sattler SE (2021) Loss of COMT activity reduces lateral root formation and alters the response to water limitation in sorghum *brown midrib* (*bmr*) *12* mutant. New Phytol 229(5):2780–2794. https://doi.org/10.1111/nph.17051
30. Shi M, Huang FF, Deng CP, Wang Y, Kai GY (2019) Bioactivities, biosynthesis and biotechnological production of phenolic acids in *Salvia miltiorrhiza*. Crit Rev Food Sci Nutr 59(6):953–964. https://doi.org/10.1080/10408398.2018.1474170
31. Srivastava S, Srivastava AK (2007) Hairy root culture for mass-production of high-value secondary metabolites. Crit Rev Biotechnol 27(1):29–43. https://doi.org/10.1080/07388550601173918
32. Tanaka Y, Brugliera F (2013) Flower colour and cytochromes P450. Philos Trans R Soc Lond B Biol Sci 368(1612):20120432. https://doi.org/10.1098/rstb.2012.0432
33. Tian SK, Wang DD, Yang L, Zhang ZX, Liu Y (2022) A systematic review of 1-Deoxy-D-xylulose-5-phosphate synthase in terpenoid biosynthesis in plants. Plant Growth Regul 96(2):221–235. https://doi.org/10.1007/s10725-021-00784-8
34. Tomlekova N, Spasova-Apostolova V, Pantchev I, Sarsu F (2021) Mutation associated with orange fruit color increases concentrations of beta-carotene in a sweet pepper variety (*Capsicum annuum* L.). Foods 10(6):1225. https://doi.org/10.3390/foods10061225

35. Vogt T (2010) Phenylpropanoid biosynthesis. Mol Plant 3(1):2–20. https://doi.org/10.1093/mp/ssp106
36. Wang DD, Zhang ZX, Yang L, Tian SK, Liu Y (2021) *ARPI*, beta-*AS*, and *UGE* regulate glycyrrhizin biosynthesis in *Glycyrrhiza uralensis* hairy roots. Plant Cell Rep 40(7):1285–1296. https://doi.org/10.1007/s00299-021-02712-6
37. Wen Y, Chen HS, Zhang L, Wu ML, Zhang F, Yang D, Shen JG, Chen JP (2021) Glycyrrhetinic acid induces oxidative/nitrative stress and drives ferroptosis through activating NADPH oxidases and iNOS, and depriving glutathione in triple-negative breast cancer cells. Free Radic Biol Med 173:41–51. https://doi.org/10.1016/j.freeradbiomed.2021.07.019
38. Weng JK, Akiyama T, Ralph J, Chapple C (2011) Independent recruitment of an *O*-methyltransferase for syringyl lignin biosynthesis in *Selaginella moellendorffii*. Plant Cell 23(7):2708–2724. https://doi.org/10.1105/tpc.110.081547
39. Xiao XY, Hao M, Yang XY, Ba Q, Li M, Ni SJ, Wang LS, Du X (2011) Licochalcone A inhibits growth of gastric cancer cells by arresting cell cycle progression and inducing apoptosis. Cancer Lett 302(1):69–75. https://doi.org/10.1016/j.canlet.2010.12.016
40. Xie CF, Li XT, Wu JS, Liang ZF, Deng FF, Xie W, Zhu MM, Zhu JY, Zhu WW, Geng SS, Zhong CY (2015) Anti-inflammatory activity of magnesium isoglycyrrhizinate through inhibition of phospholipase A_2/arachidonic acid pathway. Inflammation 38(4):1639–1648. https://doi.org/10.1007/s10753-015-0140-2
41. Yang G, Lee HE, Yeon SH, Kang HC, Cho YY, Lee HS, Zouboulis CC, Han SH, Lee JH, Lee JY (2018) Licochalcone A attenuates acne symptoms mediated by suppression of NLRP3 inflammasome. Phytother Res 32(12):2551–2559. https://doi.org/10.1002/ptr.6195
42. Yang L, Jiang Y, Zhang ZX, Hou JM, Tian SK, Liu Y (2020) The anti-diabetic activity of licorice, a widely used Chinese herb. J Ethnopharmacol 263:113216. https://doi.org/10.1016/j.jep.2020.113216
43. Yang L, Wang DD, Tian SK, Zhang ZX, Hou JM, Xiao Y, Liu Y (2021) The role of licorice *DXS* knockout and overexpression in glycyrrhizic acid biosynthesis. Acta Pharm Sin 56(07):2025–2032. https://doi.org/10.16438/j.0513-4870.2021-0225
44. Yang R, Li WD, Ma YS, Zhou S, Xue YT, Lin RC, Liu Y (2017) The molecular identification of licorice species and the quality evaluation of licorice slices. Acta Pharm Sin 52(02):318–326. https://doi.org/10.16438/j.0513-4870.2016-0602
45. Yi Y, Li JH, Lai XY, Zhang M, Kuang Y, Bao YO, Yu R, Hong W, Muturi E, Xue H, Wei HP, Li T, Zhuang H, Qiao X, Xiang KH, Yang H, Ye M (2022) Natural triterpenoids from licorice potently inhibit SARS-CoV-2 infection. J Adv Res 36:201–210. https://doi.org/10.1016/j.jare.2021.11.012
46. Yin YC, Hou JM, Tian SK, Yang L, Zhang ZX, Li WD, Liu Y (2020) Overexpressing chalcone synthase (CHS) gene enhanced flavonoids accumulation in *Glycyrrhiza uralensis* hairy roots. Bot Lett 167(2):219–231. https://doi.org/10.1080/23818107.2019.1702896

47. Zeng F, Weng ZB, Zheng HL, Xu MM, Liang XF, Duan JN (2021) Preparation and characterization of active oxidized starch films containing licorice residue extracts and its potential against methicillin-resistant *S. aureus*. Int J Biol Macromol 187:858–866. https://doi.org/10.1016/j.ijbiomac.2021.07.179
48. Zhang N, Ma XM, Li R, Xue YH, Sun YS, Nie SS, Zhang LG (2019) Transcriptome-based analysis of carotenoid accumulation-related gene expression in petals of Chinese cabbage (*Brassica rapa* L.). 3 Biotech 9(7):274. https://doi.org/10.1007/s13205-019-1813-6
49. Zhang ZX, Wang DD, Yang L, Tian SK, Xiao Y, Liu Y (2021) Ferulate 5-hydroxylase gene (*F5H*) regulation of glycyrrhizic acid biosynthesis determined by gene overexpression and knockout. Acta Pharm Sin 56(06):1719–1726. https://doi.org/10.16438/j.0513-4870.2021-0196

Tables

Table 1 Mutation sites in the COMT sequences of the *COMT*$^{-}$ hairy root lines

Lines	Registered sequence MZ169549 in GenBank	Sequencing results	Mutation sites	Synonymous / Missense mutations
COMT-2	~NNAYGMT~	~NNAYGMT~	None	Synonymous (S)
COMT-5	~LPFN~	~LP.	160	Missense (M)
COMT-6	~AYGMTAF~	~AYGTTAF~	174	M
COMT-8	~VGSVVDVG~	~VGSVADVG~	219	M

Note: The mutation sites are marked in red; ~ represents the abbreviated amino acid sequences without mutations.

Table 2 Mutation sites in the CRTZ sequences of the *CRTZ*$^{-}$ hairy root lines

Lines	Registry sequence MZ169550 in GenBank	Sequencing results	Mutation sites	Synonymous / Missense mutations
CRTZ-1 CRTZ-5-2 CRTZ-6 CRTZ-8-1	~INAAPAITLLS~	~INAVPAIALLS~	200 204	M
CRTZ-2-1	~INAAPA~	~INAVPA~	200	M
CRTZ-2-2	~HESHH~ ~INAAPA~	~HEPHH~ ~NAVPA~	179 200	M
CRTZ-3 CRTZ-4 CRTZ-5-1 CRTZ-7-1 CRTZ-7-2 CRTZ-8-2	~PAITLLS~	~PAIALLS~	204	M

Note: The mutation sites are marked in red; ~ represents the abbreviated amino acid sequences without mutations.

Table 3 is available in the Supplementary Files section.

Figures

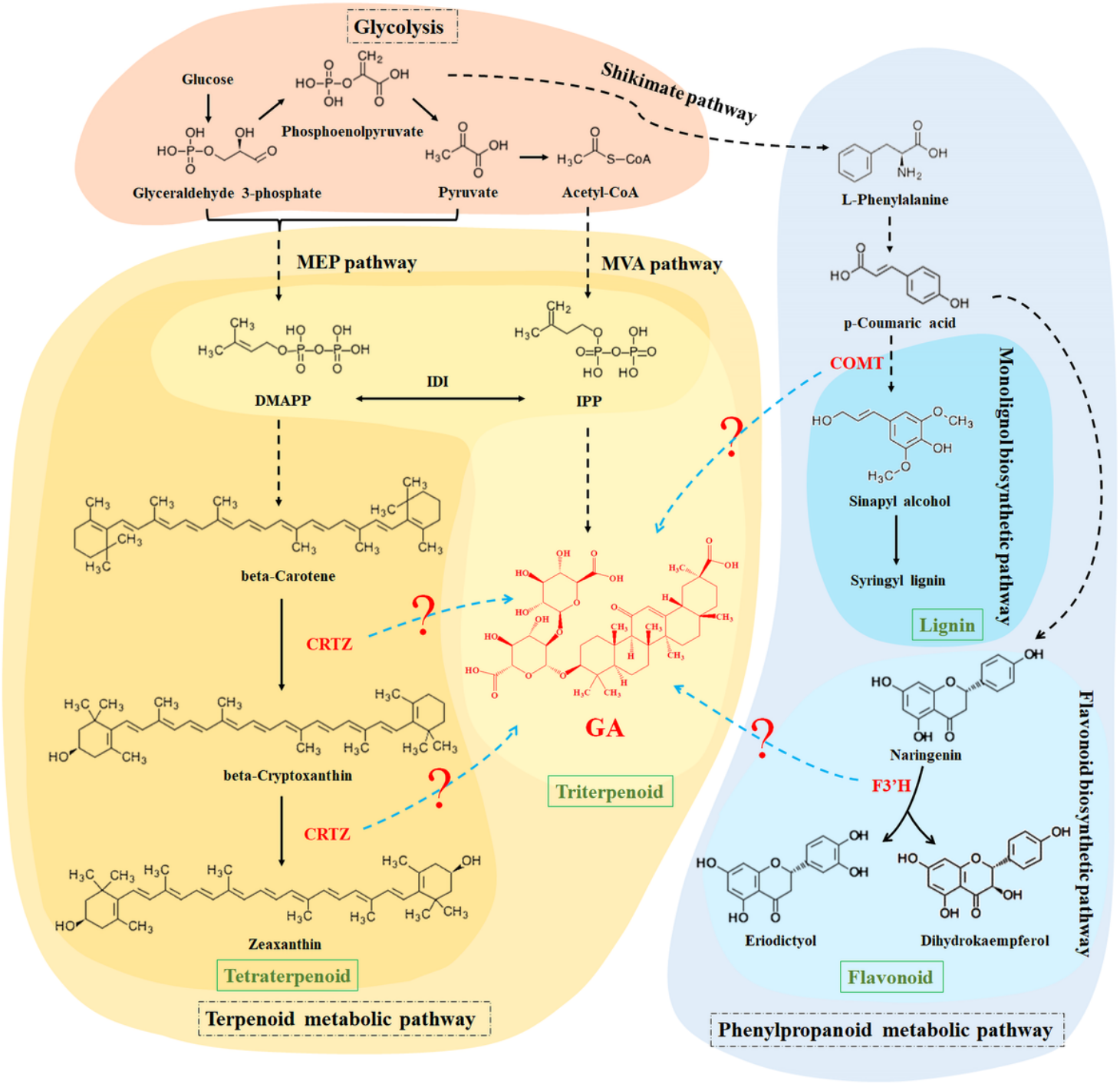


Figure 1

Metabolic pathways in *G. uralensis*. Glycolysis, the primary metabolic pathway, is marked in orange, the terpenoid metabolic pathway is marked in yellow, and the phenlpropanoid metabolic pathway for flavonoid and lignin synthesis is marked in blue. The primary metabolites, such as pyruvate, acetyl-CoA, and phosphoenolpyruvate, flow into different pathways to produce various secondary metabolites. CRTZ, COMT, and F3'H are key enzymes involved in the tetraterpenoid, monolignol, and flavonol biosynthetic pathways, respectively. Their effects on the biosynthesis of GA are far from clear. The dotted arrows indicate multi-step reactions, and the straight arrows indicate single-step reactions. MEP, 2-C-methl-D-

erythritol-4-phospate; MVA, Mevalonate; DMAPP, Dimethylallyl diphosphate; IPP, Isopentenyl pyrophosphate; IDI, Isopentenyl diphosphate isomerase; CRTZ, Beta-carotene 3-hydroxylase; F3'H, Flavonoid 3'-monooxygenase; COMT, Caffeic acid 3-O-methyltransferase

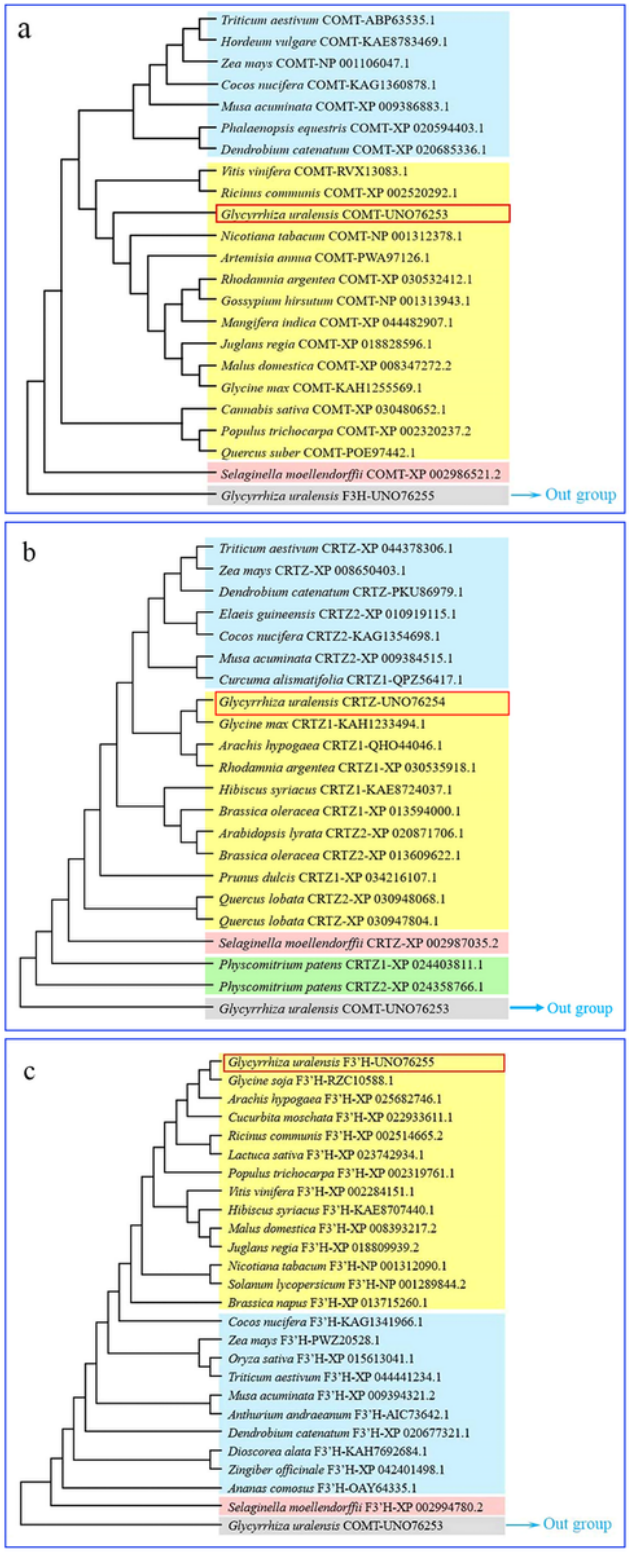


Figure 2

The phylogenetic trees of COMT (a), CRTZ (b), and F3’H (c). Species from Monocotyledoneae, Dicotyledoneae, Pteridophyta, and Bryophyta are marked in blue, yellow, red, and green, respectively. The COMT, CRTZ, and F3’H sequences identified from *G. uralensis* are marked in the red boxes. The outgroups are marked in gray.

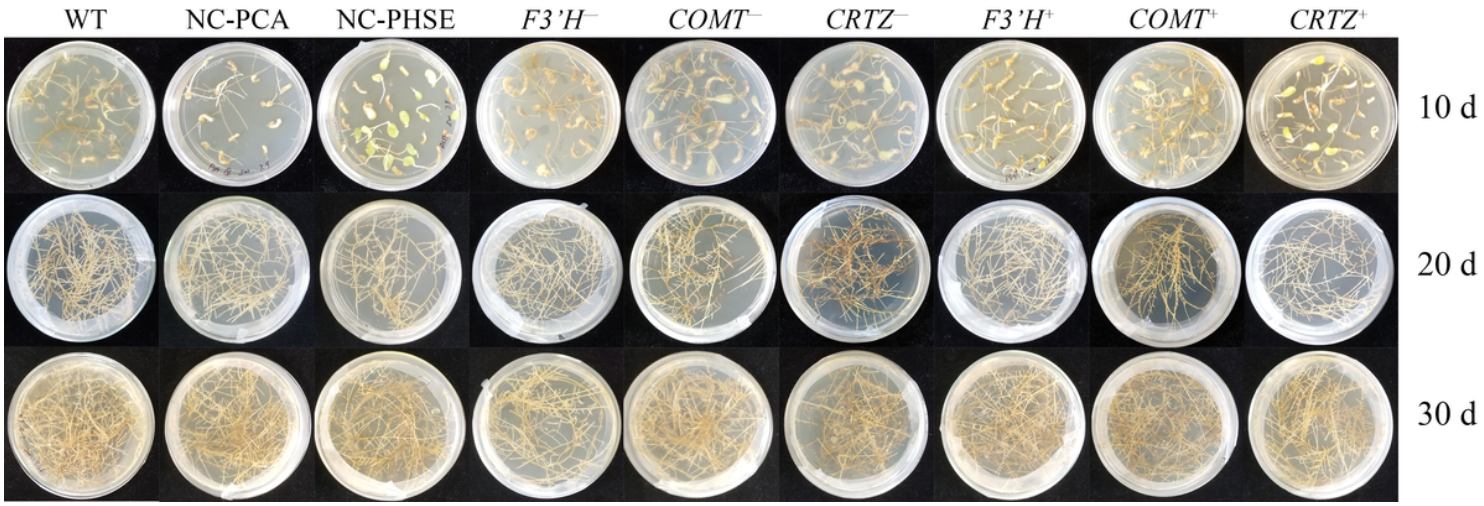


Figure 3

The induction and culture of *G. uralensis* hairy roots. The photos of 10-, 20-, and 30-day old hairy root lines after the onset of induction in solid 6,7-V medium. WT represents the wild type hairy roots. NC-PCA and NC-PHSE represent the negative control hairy root lines containing the empty pCAMBIA1305.1 and pHSE401 vectors, respectively. *F3’H-*, *COMT-*, and *CRTZ-* represent the hairy root lines knocking out the *F3’H*, *COMT*, and *CRTZ* genes, respectively. *F3’H+*, *COMT+*, and *CRTZ+* represent the hairy root lines overexpressing the *F3’H*, *COMT*, and *CRTZ* genes, respectively.

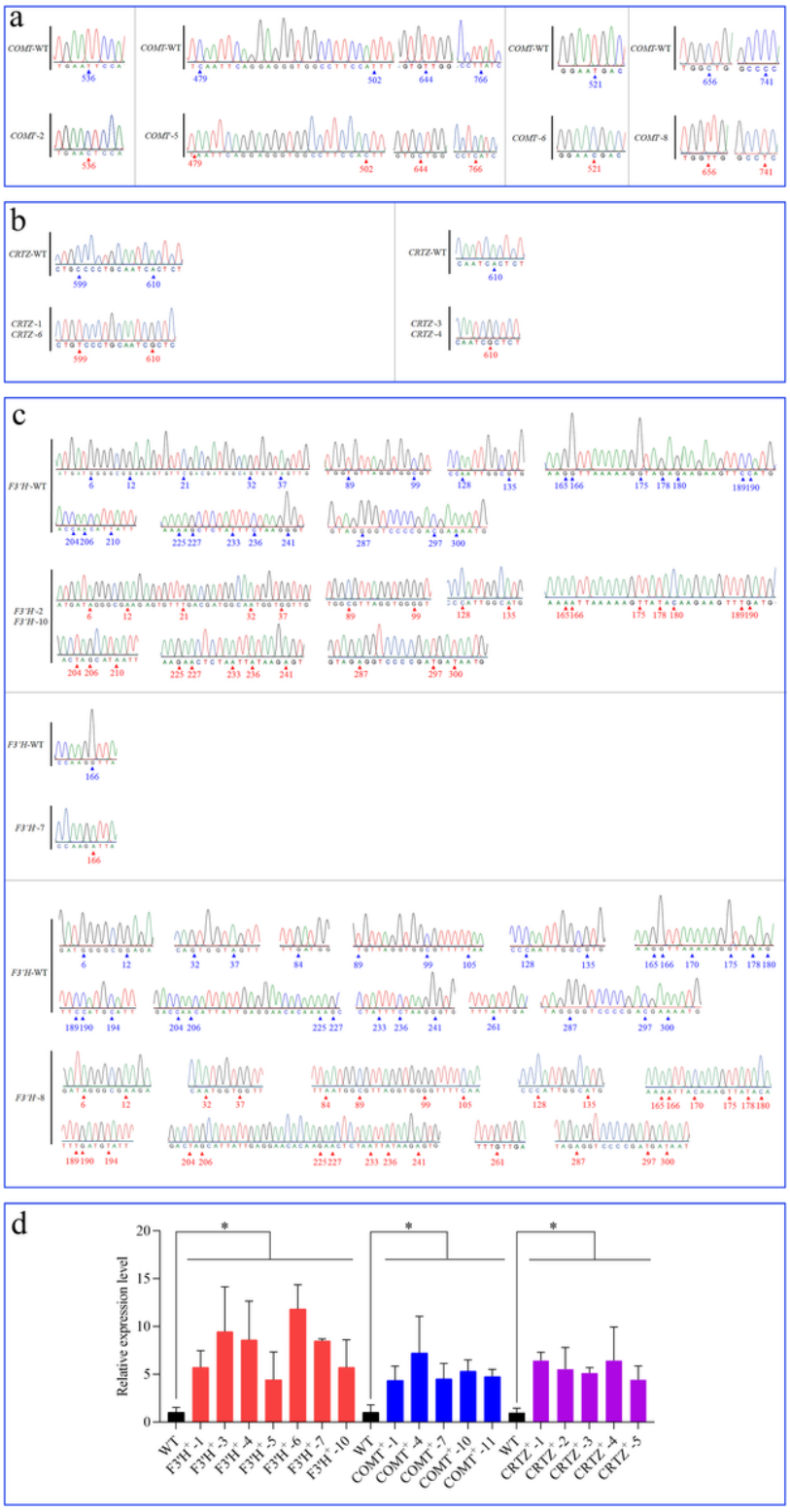


Figure 4

Analyses of gene editing sites and expression levels in hairy root lines. a, b, and **c** show the homozygous mutations in the *CRTZ*-, *COMT*-, and *F3'H*- hairy root lines, respectively. The mutation sites are marked by the red triangles and the corresponding sites in the wild type sequences are marked by the blue ones. **d** shows the relative expression levels of the three genes in the *F3'H*+, *COMT*+, and *CRTZ*+hairy root lines. * represents $P < 0.05$ (vs. WT).

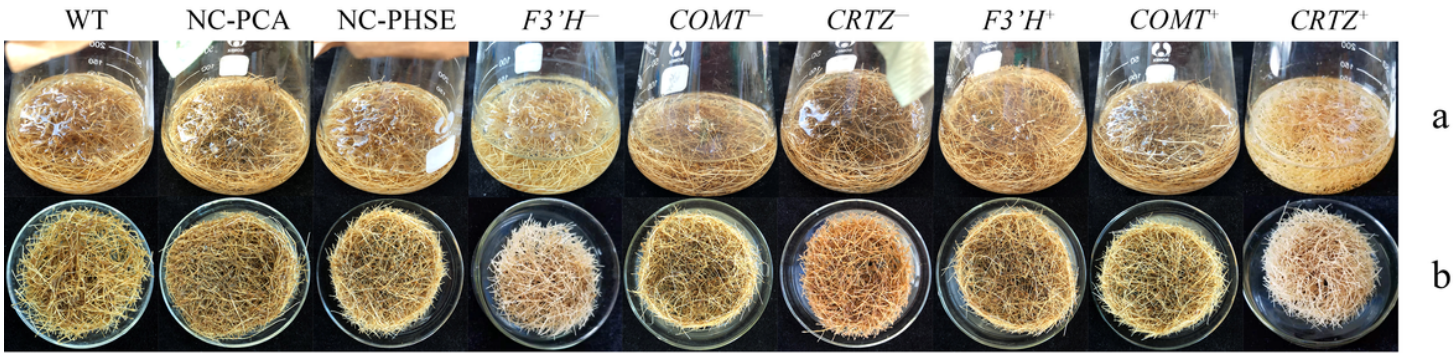


Figure 5

UPLC sample preparation. **a**shows the 3-week old hairy root lines cultured in liquid 6,7-V medium. **b**shows the hairy root samples collected for UPLC analyses.

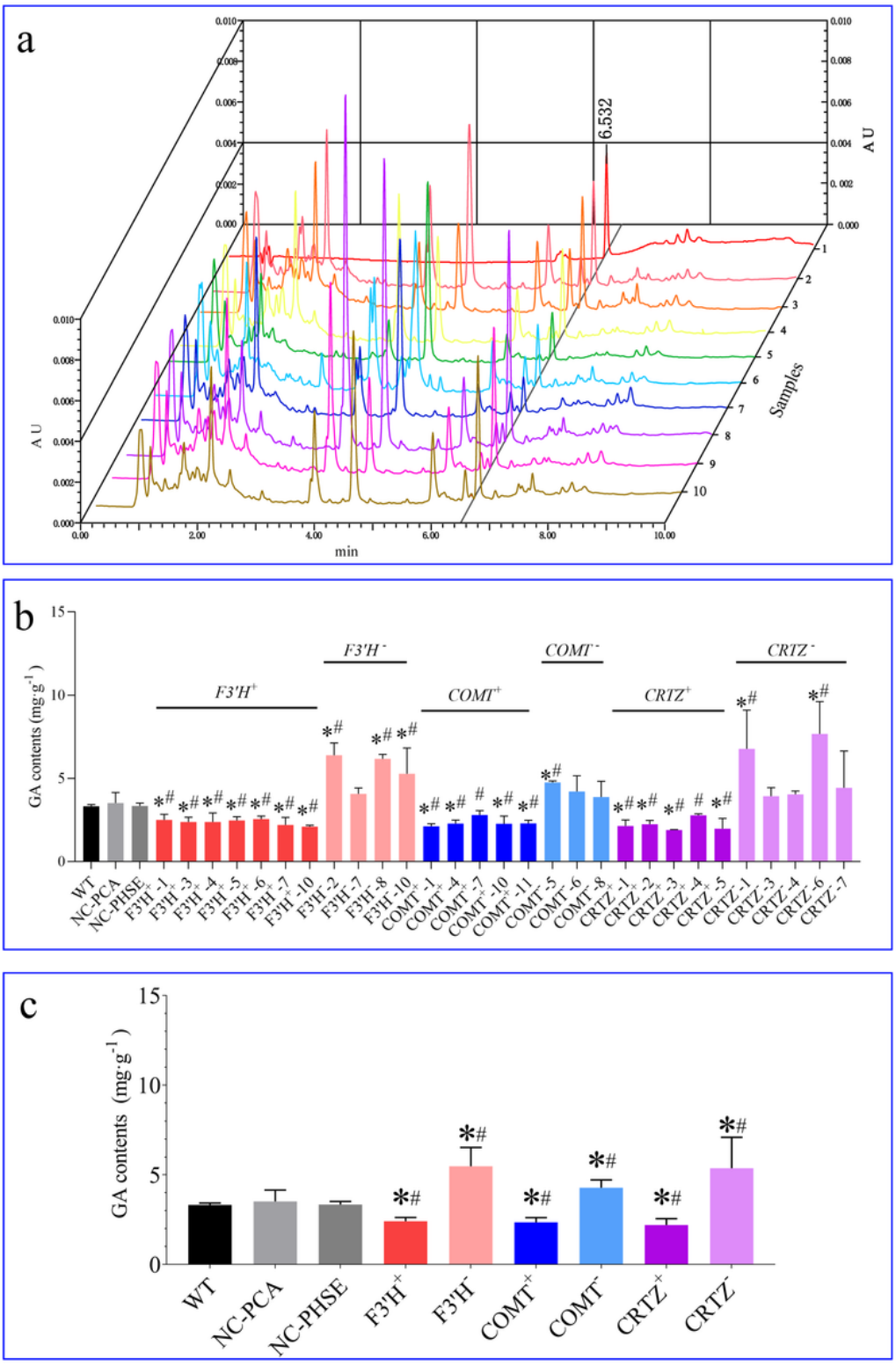


Figure 6

GA content analyses in hairy root lines. **a** shows the UPLC chromatograms, line 1 is the UPLC chromatogram of the reference substance GA, lines 2~10 are UPLC chromatograms of the WT, NC-PCA, NC-PHSE, *F3'H*+, *COMT*+, *CRTZ*+, *F3'H*-, *COMT*-, and *CRTZ*- samples, respectively. **b** shows the GA content in each hairy root line. **c** shows the GA contents in the nine groups. * represents $P < 0.05$ (vs. WT). **#**represents $P < 0.05$ (vs. NC-PCA or NC-PHSE).

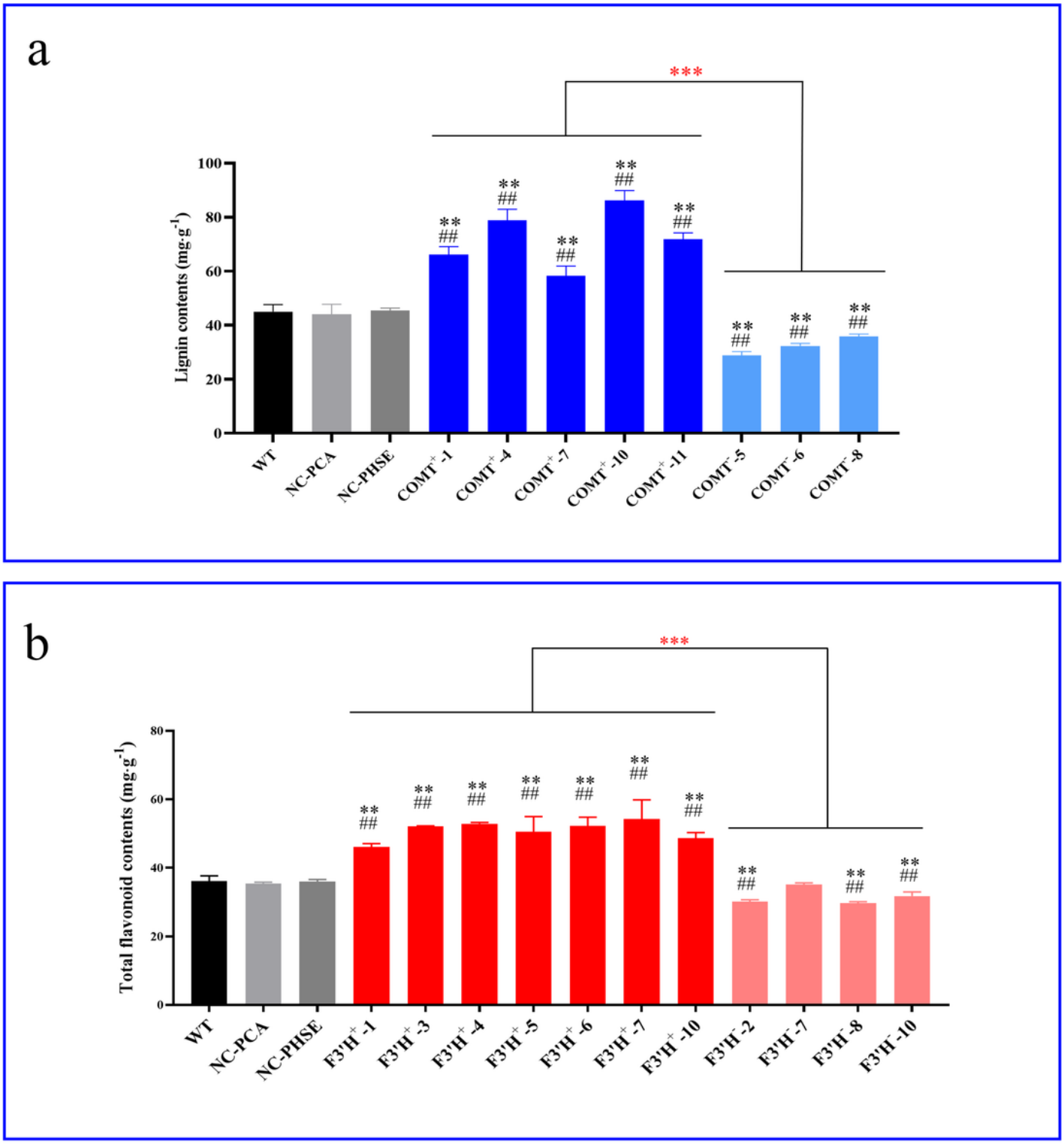


Figure 7

Lignin and total flavonoid content analyses in hairy root lines. **a** shows the lignin contents in the WT, NC-PCA, NC-PHSE, *COMT*+, and *COMT*- lines. **b** shows the total flavonoid contents in the WT, NC-PCA, NC-PHSE, *F3'H*+, and *F3'H*- lines.* and **represents $P < 0.05$ and $P < 0.01$ (vs. WT). **#** and ## represents $P < 0.05$ and $P < 0.01$ (vs. NC-PCA or NC-PHSE). ***represents $P < 0.001$.

Supplementary Files

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

- Supplementalmaterials.pdf
- Table3.docx