



Integrated Transcriptomics and Widely Targeted Metabolomics Analyses Provide Insights Into Flavonoid Biosynthesis in the Rhizomes of Golden Buckwheat (*Fagopyrum cymosum*)

Juan Huang^{1*†}, Luyuan Wang^{2†}, Bin Tang¹, Rongrong Ren¹, Taoxiong Shi¹, Liwei Zhu¹, Jiao Deng¹, Chenggang Liang¹, Yan Wang¹ and Qingfu Chen^{1*}

¹ Research Center of Buckwheat Industry Technology, Guizhou Normal University, Guiyang, China, ² Tunliu District Vocational Senior Middle School, Changzhi, China

OPEN ACCESS

Edited by:

Supaart Sirikantaramas,
Chulalongkorn University, Thailand

Reviewed by:

Ashutosh Pandey,
National Institute of Plant Genome
Research (NIPGR), India
Atsushi Fukushima,
Kyoto Prefectural University, Japan

*Correspondence:

Qingfu Chen
cqf1966@163.com
Juan Huang
huang200669@163.com

[†]These authors have contributed
equally to this work

Specialty section:

This article was submitted to
Plant Metabolism and Chemodiversity,
a section of the journal
Frontiers in Plant Science

Received: 29 October 2021

Accepted: 22 March 2022

Published: 17 June 2022

Citation:

Huang J, Wang L, Tang B, Ren R,
Shi T, Zhu L, Deng J, Liang C, Wang Y
and Chen Q (2022) Integrated
Transcriptomics and Widely Targeted
Metabolomics Analyses Provide
Insights Into Flavonoid Biosynthesis in
the Rhizomes of Golden Buckwheat
(*Fagopyrum cymosum*).
Front. Plant Sci. 13:803472.
doi: 10.3389/fpls.2022.803472

Golden buckwheat (*Fagopyrum cymosum*) is used in Traditional Chinese Medicine. It has received attention because of the high value of its various medicinal and nutritional metabolites, especially flavonoids (catechin and epicatechin). However, the metabolites and their encoding genes in golden buckwheat have not yet been identified in the global landscape. This study performed transcriptomics and widely targeted metabolomics analyses for the first time on rhizomes of golden buckwheat. As a result, 10,191 differentially expressed genes (DEGs) and 297 differentially regulated metabolites (DRMs) were identified, among which the flavonoid biosynthesis pathway was enriched in both transcriptome and metabolome. The integration analyses of the transcriptome and the metabolome revealed a network related to catechin, in which four metabolites and 14 genes interacted with each other. Subsequently, an SG5 R2R3-MYB transcription factor, named *FcMYB1*, was identified as a transcriptional activator in catechin biosynthesis, as it was positively correlated to eight flavonoid biosynthesis genes in their expression patterns and was directly bound to the promoters of *FcLAR2* and *FcF3'H1* by yeast one hybrid analysis. Finally, a flavonoid biosynthesis pathway was proposed in the rhizomes of golden buckwheat, including 13 metabolites, 11 genes encoding 9 enzymes, and 1 MYB transcription factor. The expression of 12 DEGs were validated by qRT-PCR, resulting in a good agreement with the Pearson R ranging from 0.83 to 1. The study provided a comprehensive flavonoid biosynthesis and regulatory network of golden buckwheat.

Keywords: *Fagopyrum cymosum* rhizome, transcriptomics, flavonoid biosynthesis, MYB transcription factor, widely targeted metabolomics

INTRODUCTION

Common buckwheat (*Fagopyrum esculentum* Moench), Tartary buckwheat (*Fagopyrum tataricum* (L.) Gaertn), and golden buckwheat (*Fagopyrum cymosum*, also called *Fagopyrum dibotrys*) are three major species of the genus *Fagopyrum* Mill in the *Polygonum* family (Ohnishi and Matsuoka, 1996; Chen et al., 2018). Among them, common buckwheat and Tartary buckwheat are utilized as

food or feed and mostly found as diploid, whereas golden buckwheat is utilized as medicine or tea and found as diploid or tetraploid (Mazza and Oomah, 2003; Gu et al., 2014; Chen et al., 2018). Golden buckwheat is one of the National Key Protected Wild Plants (Class II) in China, and its rhizome is a traditional Chinese medicine included in the Chinese Pharmacopoeia (2020) with high medicinal value (Chiang et al., 2000). Golden buckwheat is also known as a health food because it has high nutritional value and healthcare function. Therefore, it is utilized to make healthy products such as fermented tea and root slices (Yang et al., 2007; Gu et al., 2014). The rhizome, as the medicinal tissue of golden buckwheat, is rich in various medicinal and hygienical components such as flavonoids (catechin and epicatechin), antioxidant phenolic compounds, γ -aminobutyric acid, and terpenoids (Shao et al., 2005; Wang et al., 2005; Zhang et al., 2016). Among these, flavonoids are some of the most functionally valuable components beneficial to health that are good for gastrointestinal dysfunction; enhance vascular toughness; reduce blood sugar, blood fat, and inflammation; promote tumor inhibition; and improve immunity (Chan, 2003; Li et al., 2021).

Proanthocyanidins are plant polyphenols formed by the condensation of flavane-3-ol monomers such as catechin and epicatechin (Li et al., 2021). They are the last components in flavonoid biosynthesis and essential bioactive components in golden buckwheat. The content of epicatechin is a medicinal indicator for evaluating the quality of the rhizomes of golden buckwheat in the Chinese Pharmacopoeia (2020) (Chiang et al., 2000). Therefore, more and more research has focused on flavonoids in golden buckwheat in recent years. Li et al. (2021) purified proanthocyanidins and analyzed their detailed structure in golden buckwheat rhizome. The contents, antioxidant activity, and antidiabetic activity in golden buckwheat rhizome were higher than those in six other *Polygonaceae* plants (Li et al., 2021). Chen et al. obtained a golden buckwheat mutant that had higher epicatechin content and elucidated the molecular mechanism underlying proanthocyanidin accumulation by transcriptomes in which unigenes in radiation-mediated flavonoid biosynthesis were identified (Jia and Li, 2008; Chen et al., 2012; Chen and Li, 2016). Besides, several genes in the flavonoid biosynthesis of golden buckwheat were cloned and analyzed, including *FdPAL*, *FdFLS*, *FdCHI*, *FdCHS*, *FdDFR*, *FdANS*, *FdLAR*, and *FdMYB* (Liu et al., 2009; Ma et al., 2009, 2012; Meng et al., 2010; Li et al., 2011; Jiang et al., 2013; Luo et al., 2013; Pu et al., 2014).

Although the flavonoid biosynthesis pathway and regulatory network in golden buckwheat are far from clear, they have been thoroughly uncovered in many other plants (Saito et al., 2013; Kelemen et al., 2015; Xu et al., 2015; Tohge et al., 2017). There are more than 6,000 identified flavonoids in plants, which have diverse biological functions, including plant defense, signaling during nodulation, and flower and fruit coloration (Falcone Ferreyra et al., 2012; Tohge et al., 2018; Zhao et al., 2020; Bosse et al., 2021). The biosynthesis pathway of flavonoids initiates from the formation of malonyl-CoA and p-coumaroyl-CoA, among which, malonyl-CoA is catalyzed by acetyl-CoA carboxylase, and p-coumaroyl-CoA is catalyzed by phenylalanine ammonialyase (PAL), cinnamate 4-hydroxylase

(C4H), and 4-coumarate CoA ligase (4CL). Then malonyl-CoA and p-coumaroyl-CoA are converted into flavonoid scaffolds, including dihydroflavonols and the subsequently formed proanthocyanidins, by a complex series of reactions. Dihydroflavonols are catalyzed by chalcone synthase (CHS), chalcone isomerase, flavanone 3-hydroxylase (F3H), flavonoid 3'-hydroxylase (F3'H), flavonoid 3',5'-hydroxylase (F3'5'H), and flavonol synthase (FLS). Proanthocyanidins are catalyzed by dihydroflavonol 4-reductase (DFR), leucoanthocyanidin reductase (LAR), and anthocyanidin reductase (Saito et al., 2013; Chen and Li, 2016; Tohge et al., 2017). Massive flavonoid biosynthesis and regulation are characterized by involving enzyme encoding genes and families of transcription factors, including R2R3-MYB, bHLH, and WD40 proteins, and their interactions (MYB-bHLH-WD40 complex) (Kelemen et al., 2015; Xu et al., 2015; Zhang et al., 2018, 2019; Li et al., 2019c; Sun et al., 2020; Yao et al., 2020; Ding et al., 2021).

Previously, we collected 211 golden buckwheat accessions from different provinces in China. They were divided into three categories based on total flavonoid content (Wang et al., 2019). The current study aimed to make up for the deficiency of flavonoid biosynthesis pathways and regulatory networks in golden buckwheat. We performed transcriptome and widely targeted metabolome analyses using four golden buckwheat accessions that differed in total flavonoid content. Massive differentially expressed genes (DEGs) and differentially regulated metabolites (DRMs) were identified. Integration analyses of the transcriptome and metabolome profiles were carried out, and a flavonoid biosynthesis pathway in the rhizomes of golden buckwheat was proposed. A yeast one-hybrid (Y1H) analysis was conducted to validate the binding of an MYB transcription factor in promoters of flavonoid biosynthesis genes. A quantitative reverse transcription PCR (qRT-PCR) analysis of 13 DEGs was performed to verify the transcriptome results. This study provided a comprehensive flavonoid biosynthesis and regulatory network of golden buckwheat.

MATERIALS AND METHODS

Plant Growth and Sampling

Four golden buckwheat collections were used as plant materials in this study, namely, R1, R2, R3, and R4. Among these, R1 originated from Shangri-la, Yunnan province (longitude 100°05', latitude 27°19', and altitude 3,452 m); R2 originated from Kunming City, Yunnan province (longitude 102°82', latitude 24°89', and altitude 1,916 m); R3 originated from Kaiyang county, Guizhou province (longitude 106°96', latitude 27°06', and altitude 1,228 m); R4 was kept by our laboratory. They were planted by cutting in the experimental field in March 2017, and with normal field management during the growth periods. For transcriptome and widely targeted metabolome analyses, the golden yellow rhizomes were sampled and washed by ddH₂O. The fibrous root was removed, the remaining rhizomes were cut into portions, and quickly subjected to liquid nitrogen for subsequent experiments. Three biological replicates were included for each golden buckwheat collection.

Transcriptome Analyses

Total RNA was isolated using a plant RNA purification kit [TianGen Biotech (Beijing) Co., Ltd., China], followed by treatment with Dnase I for 20 min to digest genomic DNA. For each sample, 1- μ g RNA was used for library construction, using NEBNextUltraTM RNA Library Prep Kit for Illumina (NEB, United States) according to the manufacturer's instruction. The library's quality was evaluated with the Agilent Bioanalyzer 2100 system. High-throughput transcriptome sequencing was performed on NOVASEQ 6000 System (Illumina Inc., United States) with a read length of 150 bp and a paired-end method.

Raw reads were filtered by in-house Perl scripts to obtain clean data. Then, the clean reads were mapped to the Tartary buckwheat genome data using Hisat2 (Kim et al., 2015; Zhang et al., 2017). Only reads with a perfect match or one mismatch were further analyzed and annotated based on the reference genome. Gene function annotation was performed by a local BLASTP against NR, Pfam, Swiss-Prot, KO, and GO database ($1e^{-5}$). Afterward, gene expression was normalized to fragments per kilobase of transcript per million fragments mapped (FPKM) as the following:

$$FPKM = \frac{cDNA \text{ Fragments}}{\text{Mapped Fragments (Millions)} * \text{Transcript Length (kb)}}.$$

DEGs were identified with the DESeq2 package in the R language (<https://www.R-project.org/>), with an adjusted *P*-value (padj) < 0.05 and absolute value of \log_2 ratio ≥ 1 as the thresholds (Love et al., 2014). Principal component analysis (PCA), Spearman's correlation coefficient, GO enrichment analyses, and hierarchical cluster were performed and visualized by the libraries in the R language, namely scatterplot3d, ggplot2, clusterProfiler, pheatmap, reshape2, and factoextra. The KEGG pathway was enriched with KOBAS 3.0 (<http://kobas.cbi.pku.edu.cn/kobas3/?t=1>).

Widely Targeted Metabolome Analyses

Metabolite analysis was performed by Metware Biotechnology Co., Ltd. (Wuhan, China) as previously described (Chen et al., 2013; Zhao et al., 2019). The freeze-dried rhizome was crushed using a mixer mill (MM 400; Retsch) with a zirconia bead for 1.5 min at 30 Hz. Then, 100-mg powder for each sample was weighed and treated with 0.6 ml 70% aqueous methanol overnight at 4°C. Following centrifugation at 10,000 g for 10 min, extracts were filtrated with a 0.22- μ m pore size membrane and stored in chromatographic sample bottles.

Then, the sample extracts were analyzed using a UPLC-ESI-MS/MS system. The UPLC (Shim-pack UFLC SHIMADZU CBM30A system) analytical conditions were set as follows: (1) Column, Waters ACQUITY UPLC HSS T3 C18 (1.8 μ m, 2.1 mm*100 mm); (2) mobile phase, solvent A was pure water with 0.04% acetic acid and solvent B was acetonitrile with 0.04% acetic acid; (3) gradient elution: sample measurements were performed with a gradient program that employed the starting conditions of 95% A and 5% B. Within 10 min, a linear gradient was programmed, and a composition of 5% A and 95% B was kept for 1 min. Subsequently, a composition of 95% A and 5% B was adjusted within 0.1 min and kept for 2.9 min; (4) injection volume, 4 μ l; column oven, 40°C; flow rate, 0.35 ml/min.

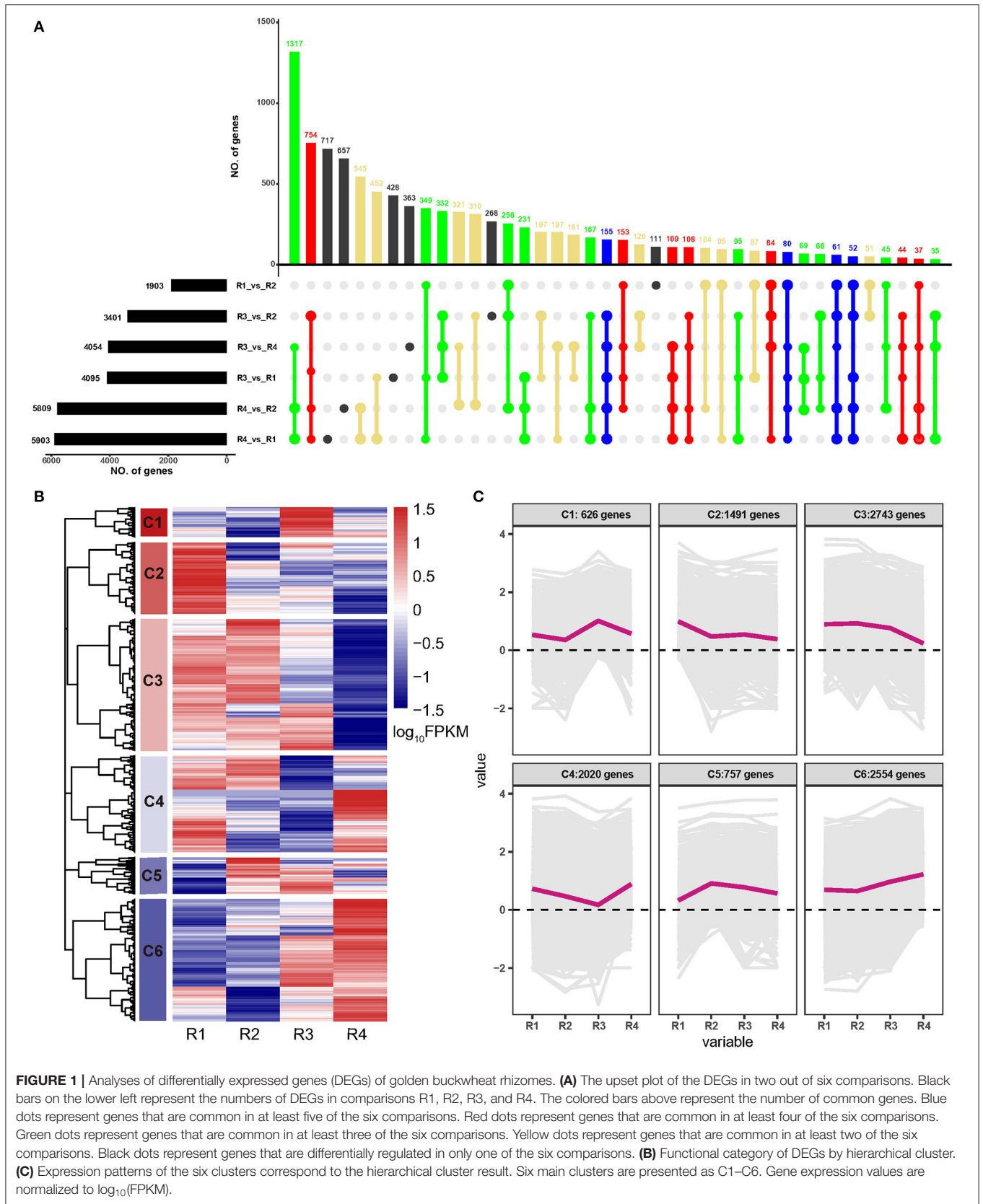
The effluent was alternatively connected to a triple quadrupole-linear ion trap (QTRAP)-MS (Applied Biosystems 4500 QTRAP) for LIT and QQQ scan, equipped with an ESI Turbo Ion-Spray interface operating in positive and negative ion modes and controlled with the Analyst 1.6.3 software (AB Sciex). ESI source operation parameters were as follows: ion source, turbo spray; source temperature 550°C; ion spray voltage 5,500 V (positive ion mode)/-4,500 V (negative ion mode); ion source gas I, gas II, curtain gas were set at 50, 60, and 30 psi, respectively; the collision gas was high. Instrument tuning and mass calibration were performed with 10 and 100 μ mol/L polypropylene glycol solutions in QQQ and LIT modes, respectively. QQQ scans were acquired as MRM experiments with collision gas (nitrogen) set to 5 psi. Declustering potential (DP) and collision energy (CE) for individual MRM transitions were conducted with further DP and CE optimization. A specific set of MRM transitions were monitored for each period according to metabolites eluted within this period (Chen et al., 2013; Zhou et al., 2019). Mass spectrometric data were processed with Analyst 1.6.3. Then, the quality and quantity of metabolites were processed with MultiaQuant based on a local metabolite database constructed by Metware Biotechnology. Quality control was performed by evaluating the repeatability of three technical replicates prepared from a mixed extract of all the samples. Then, the obtained data were log transformed (\log_2) and subjected to the MetaboAnalystR package in the R language for Pearson correlation analysis, PCA analysis, and OPLS-DA analysis (Chong and Xia, 2018). DRMs were identified by variable importance in the projection value (VIP) ≥ 1 and absolute \log_2 ratio ≥ 1 . A permutation test with 200 permutations was performed to avoid overfitting.

Integration of the Transcriptome and Metabolome Profiles

The transcriptome and metabolome data were log transformed (\log_2) for integration analyses. A Pearson correlation coefficient (PCC) analysis was performed in the R language, and significantly correlated genes and metabolites were identified by PCC > 0.8 and a *P*-value of PCC (PCCP) < 0.05 (Jozefczuk et al., 2010). The network of genes and metabolites with PCC > 0.8 was visualized with the software Cytoscape (Shannon et al., 2003). A canonical correlation analysis (CCA) was performed for genes and metabolites in the correlation network using the CCA package in the R language (González et al., 2008).

Y1H Analysis

The obtaining of promoter sequences (1,500 bp) of candidate genes and scanning of DNA-binding specificities of R2R3-MYB proteins (Kelemen et al., 2015) in the promoters were performed with in-house Python programs. Y1H was performed as previously described (Li et al., 2017). Briefly, the promoters of *FcDFR1* (FtPinG0002371500.01), *FcF3H1* (FtPinG0006662600.01), *FcF3'H1* (FtPinG0002353900.01), *FcLAR2* (FtPinG0002428800.01), and *FcCHS2* (FtPinG0002106500.01) were cloned and linked into pHIS2 with the *EcoRI* and *MluI* restriction sites. The coding sequence of *FcMYB1* (FtPinG0007213700.01) was cloned and linked to pGADT7 with the *EcoRI* and *BamHI* restriction sites. Yeast Y187



was co-transformed with pGADT7-MYB and five constructs containing corresponding promoters. The co-transformed yeast was selected in an SD/–Leu/–Trp medium to test whether the co-transformations were successful. Then, an SD/–His/–Leu/–Trp medium was used to test the expression of the HIS3 reporter; 90, 110, and 150 mM 3-AT were used to inhibit the auto-activation of pHIS2 constructs with five promoters. The primer sequences described above are listed in **Supplementary Table 1**.

qRT-PCR Analysis

The transcriptome results were verified by qRT-PCR. A total of 13 DEGs related to flavonoid biosynthesis were selected, namely, *FcCHS1* (FtPinG0000551600.01), *FcCHS2*, *FcCHS3* (FtPinG0008806400.01), *FcCHI1* (FtPinG0003061400.01), *FcF3H1*, *FcF3'H1*, *FcF3'5'H1* (FtPinG0006940000.01), *FcF3'5'H2* (FtPinG0006132400.01), *FcDFR1*, *FcFLS1* (FtPinG0006907000.01), *FcLAR1* (FtPinG0000053800.01), *FcLAR2*, and *FcMYB1*. *Actin* was used as the inner reference gene. Primer3Plus (<http://www.primer3plus.com/cgi-bin/dev/primer3plus.cgi>) was used to pick gene-specific primers (**Supplementary Table 1**). qRT-PCR was performed on an ABI 7500 Fast Real-Time PCR system (Applied Biosystems, United States) using an SYBR[®] Premix Ex Taq[™] II kit (Takara Biomedical Technology (Beijing) Co., Ltd., China). The procedure was performed in accordance with the manufacturers' instructions, with three technical replicates. qRT-PCR results were calculated using the $2^{-\Delta\Delta C_t}$ method.

RESULTS

Transcriptome Analysis of Golden Buckwheat Rhizomes

High-throughput RNA-Seq was performed on 12 RNA samples from the rhizomes of four golden buckwheat collections (R1, R2, R3, and R4), with three biological replicates for each collection. As a result, we obtained 41,034,456–52,182,636 clean reads and 6,111,106,094–7,791,174,252 clean bases for each library. GC content ranged from 46.17–49.63%. Among these, 63.6–72.36% of clean reads were uniquely mapped to the predicted coding sequences of the genome data of Tartary buckwheat. Meanwhile, 34.3–36.98% of clean reads were mapped to the “+” strand of the coding sequences, whereas 34.04–37.01% were mapped to the “–” strand of the coding sequences (**Supplementary Table 2**). After mapping, 28,492 genes were identified in all the 12 libraries. Besides, new genes that were not included in the reference genome data were also identified, resulting in 1,665 new genes.

In addition, all of the 12 samples were clustered, and the repeatability of biological replicates was evaluated. As shown in **Supplementary Figure 1**, samples of the same biological replicates were clustered closer, with correlations ranging from 0.8503 to 0.9546, whereas samples out of the biological replicates were not closely related, with correlations ranging from 0.6777 to 0.8799, indicating our data had high reliability and reproducibility.

A total of 10,191 DEGs were identified in all the samples, among which 1,903, 3,401, 4,054, 4,095, 5,809, and 5,903

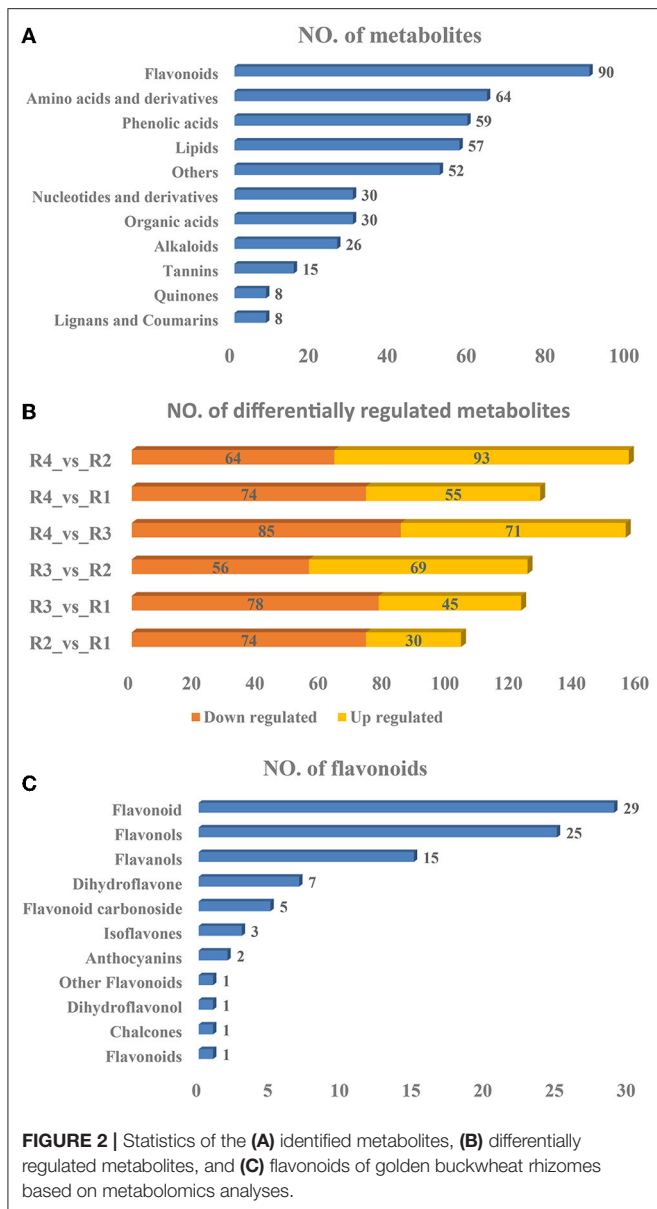
were identified in R4_vs._R1, R4_vs._R2, R3_vs._R1, R3_vs._R4, R3_vs._R2, and R1_vs._R2, respectively (**Supplementary Table 3** and **Figure 1A**). Of these, 155 DEGs were common in five comparisons except for R1_vs._R2; 80 DEGs were common in five comparisons except for R3_vs._R2; 61 DEGs were common in five comparisons except for R3_vs._R1, and 52 DEGs were common in five comparisons except for R3_vs._R4 (**Figure 1A**). Besides, a hierarchical cluster analysis based on gene expression pattern was performed and six clusters, namely, C1 to C6, were obtained (**Figures 1B,C**). Among these, C1 included 626 genes that were upregulated in R3; C2 included 1,492 genes that were upregulated in R1; C3 included 2,743 genes that showed higher expression levels in R1, R2, and R3 but lower expression levels in R4; C4 included 2,020 genes that showed higher expression level in R1, R2, and R4 but lower expression level in R3; C5 included 757 genes that showed higher expression level in R2, R3, and R4 but lower expression level in R1; and C6 included 2,554 genes that showed higher expression level in R3 and R4 but lower expression level in R1 and R2.

A GO analysis was performed, and DEGs in the six comparisons were classified into three component categories. The top 30 GO terms are presented in **Supplementary Figure 2**. Interestingly, some biological processes related to the pathway of flavonoid biosynthesis and metabolism, such as “flavonoid biosynthetic process,” “flavonol biosynthetic process,” “flavone biosynthetic process,” and “flavonoid glucuronidation,” were enriched in at least two comparisons. At the same time, we analyzed the metabolic pathways based on KEGG annotation. Among the top 20 most enriched pathways in the six comparisons (**Supplementary Figure 3**), “flavonoid biosynthesis,” “flavone and flavonol biosynthesis,” and “phenylpropanoid biosynthesis” also took place. These implied that flavonoids might show some difference among four golden buckwheat rhizomes.

Metabolomics Analyses of Golden Buckwheat Rhizomes

Widely targeted metabolomics analyses were subsequently performed to identify the metabolites existing in golden buckwheat rhizomes, resulting in 439 metabolites being identified (**Supplementary Table 4**). These included 90 flavonoids, 64 amino acids and derivatives, 59 phenolic acids, 57 lipids, 52 others, 30 organic acids, 30 nucleotides and derivatives, 26 alkaloids, 15 tannins, 8 lignans and coumarins, and 8 quinones (**Figure 2A**). The PCA analysis showed that samples of the same biological replicates were less variable than those of the biological replicates, suggesting that the metabolomics data had high reliability and reproducibility (**Supplementary Figure 4**).

A total of 297 DRMs were identified in all the samples, among which 93, 55, 71, 69, 45, and 30 were upregulated in the comparisons R4_vs._R2, R4_vs._R1, R4_vs._R3, R3_vs._R2, R3_vs._R1, and R2_vs._R1, respectively, whereas 64, 74, 85, 56, 78, and 74 were downregulated in the comparisons R4_vs._R2, R4_vs._R1, R4_vs._R3, R3_vs._R2, R3_vs._R1, and R2_vs._R1, respectively (**Figure 2B** and **Supplementary Table 5**). All the DRMs were clustered into nine subclasses based on K



means cluster analysis (Supplementary Figure 5). Metabolites in different subclasses showed different regulation patterns: subclass 1 included 58 metabolites that were upregulated in R4, subclass 2 included 37 metabolites that were upregulated in R1, subclass 3 included 23 metabolites that were upregulated in R2 and R3, subclass 4 included 26 metabolites that were downregulated in R3 but were upregulated in R4, subclass 5 included 21 metabolites that were upregulated in R2, R3, and R4, subclass 6 included 32 metabolites that were upregulated in R3, subclass 7 included 26 metabolites that were upregulated in R3 but were downregulated in R4, subclass 8 included 52 metabolites that were upregulated in R2. Subclass 8 also included 52 metabolites that were downregulated in R2. In addition, KEGG metabolic pathways based on the DRMs were analyzed. Similar to the transcriptomics result, “flavonoid biosynthesis,” “flavone and

flavonol biosynthesis,” and “phenylpropanoid biosynthesis” were also listed in the top 20 significantly enriched pathways in all of the six comparisons (Supplementary Figure 6). These indicated that flavonoids might be the most essential compounds in golden buckwheat rhizomes.

It is well-known that some kind of flavonoids, such as epicatechin, is the most essential index for evaluating whether golden buckwheat rhizomes are up to the standard of Chinese medicinal material (Chiang et al., 2000). In all kinds of identified metabolites, flavonoids accounted for the most abundant one, so we subsequently analyzed the type of flavonoids. Among the 90 flavonoids, 29 flavonoids, 25 flavonols, 15 flavanols, 7 dihydroflavones, 5 flavonoid carbonosides, 3 isoflavones, 2 anthocyanins, 1 flavonoid, 1 chalcone, 1 dihydroflavonol, and 1 other flavonoid were included (Figure 2C).

The top 20 abundant flavonoids were compared, among which 18 were identified as DRMs. These included 6 flavonols (6-hydroxykaempferol-7-o-glucoside, syringetin 3-o-hexoside, quercetin 3-o-rhanosylgalactoside, quercetin-3-o-glucoside-7-o-rhamnoside, kaempferol 3,7-dirhamnoside(kaempferitrin), and bioquercetin), 5 flavonoids (nepetin, tricrin 7-o-hexoside, luteolin 8-c-hexosyl-o-hexoside, chrysoeriol, and quercetin-p-coumaroyl hexose), 4 flavanols (catechin, gallic acid, (-)-epicatechin gallate, and (-)-catechin gallate), 2 dihydroflavones (hesperidin and eriodictyol c-hexoside), and 1 isoflavone (pratensein) (Table 1).

Integration of the Transcriptome and Metabolome Profiles Related to Flavonoid Biosynthesis

Integration analysis of the transcriptome and metabolome profiles revealed tens of thousands of correlations between genes and metabolites. What we focused on was the correlations in flavonoid biosynthesis (KEGG pathway: ko00941), in which a total of 32 genes were correlated to 15 metabolites (Supplementary Table 6). As catechin is the most abundant flavonoid, we subsequently examined whether some genes or metabolites interacted with each other. Interestingly, we found that 14 genes interacted with four metabolites, namely, catechin (mws0054), myricetin (mws0032), and two chlorogenic acids (mws0178 and pmp000544) (Figure 3A). Genes in this interaction network included 1 *CHS* gene (*FcCHS2*), 3 hydroxycinnamoyl-coenzyme A quinate transperases (HQT, FtPinG0001508400.01, FtPinG0009375700.01, and FtPinG0008704000.01), 2 spermidine hydroxycinnamoyl transferases (SHT, FtPinG0003453100.01 and FtPinG0009113300.01), 3 resveratrol synthases (RS, FtPinG0003701300.01, FtPinG0003701500.01, and FtPinG0003710800.01), 2 *C4H* genes (FtPinG0001861600.01 and FtPinG0005329600.01), 1 *LAR* gene (*FcLAR2*), 1 *FLS* gene (FtPinG0006907100.01), and one caffeoyl-CoA O-methyltransferase (CCoAOMT, FtPinG0005958500.01).

Based on their correlation with the metabolites, the 14 genes were divided into four classes (K1 to K4) based on CCA analysis (Figure 3B). K1 included two genes (*FcCHS2* and FtPinG0001508400.01) that were

TABLE 1 | Top 20 Flavonoids identified in golden buckwheat rhizomes.

Index	Retention time (min)	Molecular weight (Da)	Formula	Ionization model	Compounds	Class II	Log ₂ (R2/R1)	log ₂ (R3/R1)	log ₂ (R4/R1)	log ₂ (R3/R2)	log ₂ (R4/R2)	log ₂ (R4/R3)
mws0054	3.5	290.1	C ₁₅ H ₁₄ O ₆	[M-H] ⁻	Catechin	Flavanols	-0.1	-1.0	-0.9	-0.9	-0.8	0.1
mws0024	2.2	170.0	C ₇ H ₆ O ₅	[M-H] ⁻	Gallic acid	Flavanols	-1.2	-3.6	-0.7	-2.4	0.5	2.9
mws0036	4.5	610.2	C ₂₈ H ₃₄ O ₁₅	[M-H] ⁻	Hesperidin	Dihydroflavone	1.2	1.8	1.9	0.6	0.7	0.1
pmp000003	5.7	316.0	C ₁₆ H ₁₂ O ₇	[M+H] ⁺	Nepetin	Flavonoid	2.3	2.3	-0.5	-0.1	-2.9	-2.8
pme0460	3.8	290.1	C ₁₅ H ₁₄ O ₆	[M+H] ⁺	L-Epicatechin	Flavanols	0.0	-0.1	0.4	0.0	0.4	0.5
pmp001309	4.2	464.1	C ₂₁ H ₂₀ O ₁₂	[M+H] ⁺	6-Hydroxykaempferol-7-O-glucoside	Flavanols	1.6	2.2	1.9	0.6	0.3	-0.3
pmb0565	3.9	508.1	C ₂₃ H ₂₄ O ₁₃	[M+H] ⁺	Syringetin 3-O-hexoside	Flavanols	3.7	2.0	1.3	-1.7	-2.4	-0.7
mws0183	3.0	154.0	C ₇ H ₆ O ₄	[M-H] ⁻	Protocatechuic acid	Flavanols	0.6	0.4	0.9	-0.2	0.3	0.5
pmb0736	4.7	492.1	C ₂₃ H ₂₄ O ₁₂	[M+H] ⁺	Tricin 7-O-hexoside	Flavonoid	0.3	0.0	-6.3	-0.3	-6.6	-6.4
Li512117	4.4	610.1	C ₂₇ H ₃₀ O ₁₆	[M+H] ⁺	Quercetin 3-O-rhanosylgalactoside	Flavanols	1.6	2.1	2.2	0.5	0.6	0.1
GQ512006	4.1	610.1	C ₂₇ H ₃₀ O ₁₆	[M+H] ⁺	Quercetin-3-O-glucoside-7-O-rhamnoside	Flavanols	1.6	2.1	2.2	0.5	0.6	0.1
pme2493	4.3	578.1	C ₂₇ H ₃₀ O ₁₄	[M+H] ⁺	Kaempferol 3,7-dirhamnoside(Kaempferitrin)	Flavanols	0.6	0.5	1.2	-0.2	0.6	0.8
pmb0665	4.2	610.1	C ₂₇ H ₃₀ O ₁₆	[M+H] ⁺	Luteolin 8-C-hexosyl-O-hexoside	Flavonoid	1.3	2.0	2.0	0.6	0.7	0.1
pmn001583	4.1	610.1	C ₂₇ H ₃₀ O ₁₆	[M-H] ⁻	Bioquercetin	Flavanols	1.3	2.0	2.0	0.7	0.7	0.0
pmp000548	6.4	300.1	C ₁₆ H ₁₂ O ₆	[M+H] ⁺	Pratensein	Isoflavones	4.9	3.7	2.4	-1.3	-2.5	-1.3
pmp001127	6.7	300.1	C ₁₆ H ₁₂ O ₆	[M+H] ⁺	Chrysoeriol	Flavonoid	4.9	3.6	2.4	-1.3	-2.5	-1.2
Zmjp003072	4.3	610.1	C ₃₀ H ₂₆ O ₁₄	[M+H] ⁺	Quercetin-p-coumaroyl hexose	Flavonoids	1.5	2.0	2.3	0.6	0.8	0.2
pmb3023	3.8	450.1	C ₂₁ H ₂₂ O ₁₁	[M-H] ⁻	Eriodictyol C-hexoside	Dihydroflavone	-2.2	-1.1	-1.6	1.1	0.6	-0.5
mws1397	4.3	442.1	C ₂₂ H ₁₈ O ₁₀	[M-H] ⁻	(-)-Epicatechin gallate	Flavanols	-0.2	-1.9	-0.5	-1.7	-0.3	1.4
mws0355	4.2	442.1	C ₂₂ H ₁₈ O ₁₀	[M-H] ⁻	(-)-Catechin gallate	Flavanols	-0.2	-2.0	-0.4	-1.8	-0.2	1.5

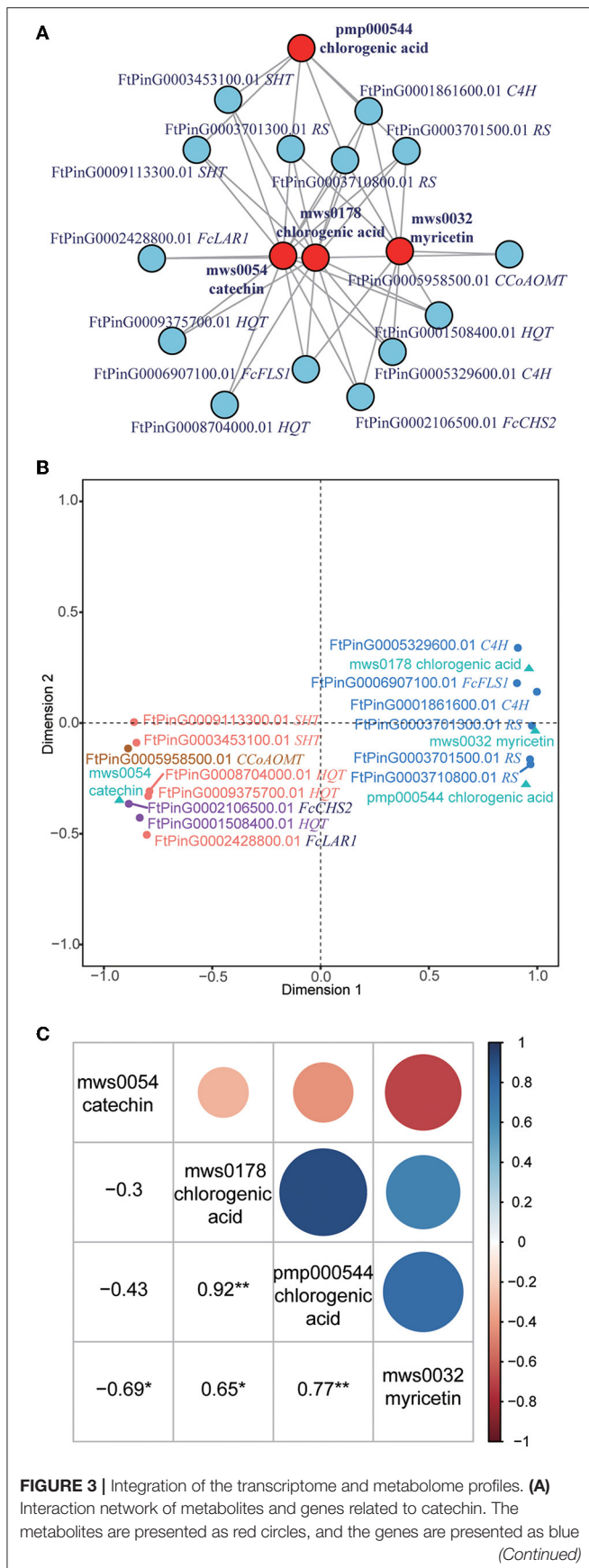


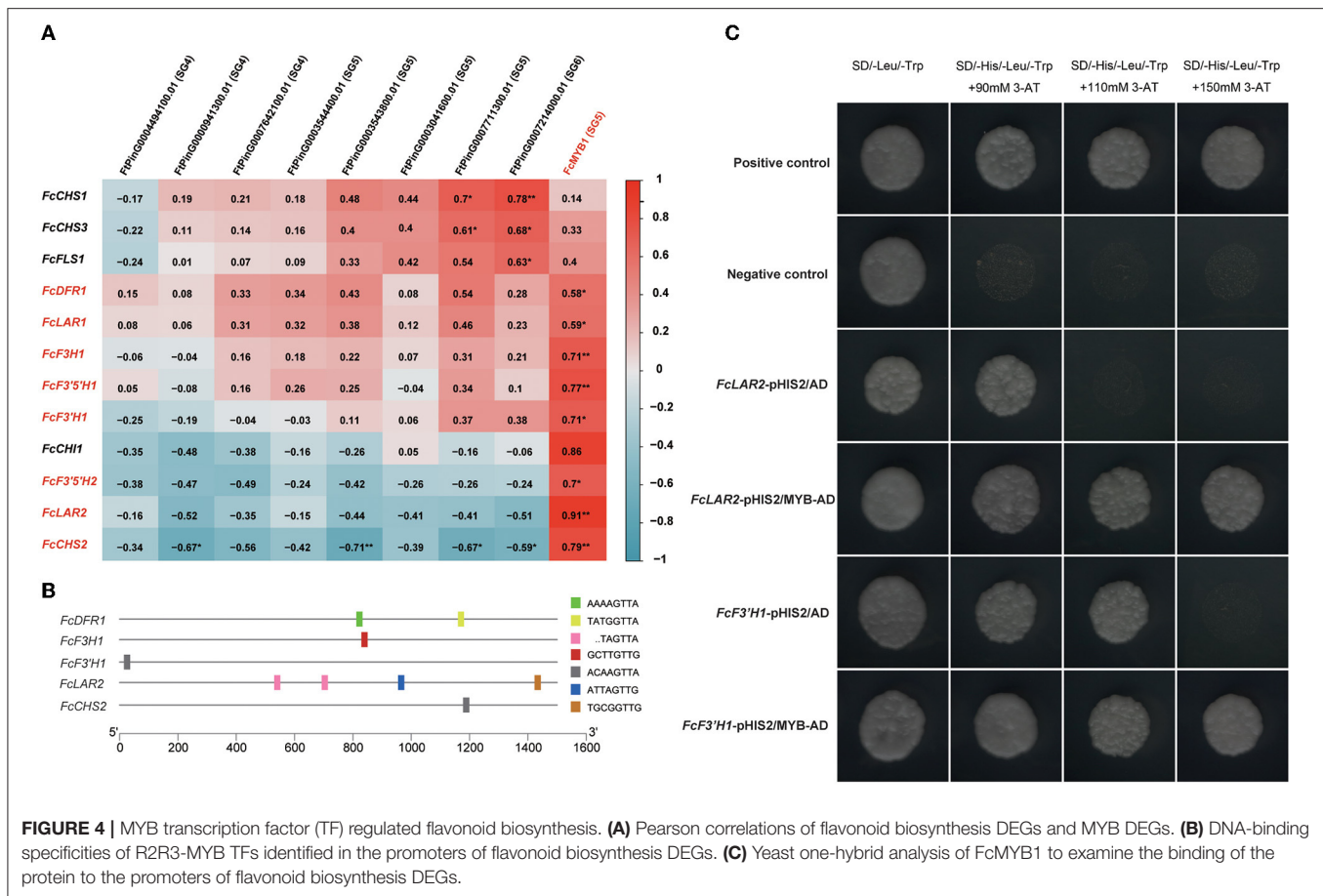
FIGURE 3 | circles. CHS, chalcone synthases; HQT, hydroxycinnamoyl-coenzyme A quinate transferase; SHT, spermidine hydroxycinnamoyl transferase; RS, resveratrol synthase; C4H, cinnamate 4-hydroxylase; LAR, leucoanthocyanidin reductase; FLS, flavonol synthase; CCoAOMT, caffeoyl-CoA O-methyltransferase. **(B)** Canonical correlation analysis (CCA) of metabolites and genes related to catechin. The metabolites are presented as green triangles. Genes in K1 are presented as purple dots. Genes in K2 are presented as orange dots. Genes in K3 are presented as blue dots. Genes in K4 are presented as brown dots. **(C)** Pearson correlation of four metabolites in the related network. *Indicates that the correlation reached a significant level of 0.05. **Indicates that the correlation reached a significant level of 0.01.

positively correlated to catechin but negatively correlated to myricetin and chlorogenic acids. K2 included five genes (FtPinG0009375700.01, FtPinG0008704000.01, FtPinG0003453100.01, FtPinG0009113300.01, and *FcLAR2*) that were positively correlated to catechin but negatively correlated to chlorogenic acids. K3 included six genes (FtPinG0003710800.01, FtPinG0006907100.01, FtPinG0003701500.01, FtPinG0001861600.01, FtPinG0005329600.01, and FtPinG0003701300.01) that were negatively correlated to catechin but positively correlated to myricetin and chlorogenic acids. K4 included one gene (FtPinG0005958500.01) that was positively correlated to myricetin and chlorogenic acids.

The Pearson correlation of four metabolites in the same network was also analyzed (Figure 3C). Notably, catechin (mws0054) was significantly negatively correlated to myricetin (mws0032) ($R = -0.69$). Myricetin (mws0032) was significantly positively correlated to two chlorogenic acids (mws0178 and pmp000544) ($R = 0.65$ and 0.77) but negatively correlated to catechin (mws0054). The correlation of two chlorogenic acids (mws0178 and pmp000544) was highest, with a Pearson R-value of 0.92.

MYB Transcription Factor Regulated Flavonoid Biosynthesis

It is widely reported that R2R3-MYB transcription factors play essential roles in the regulation of flavonoid biosynthesis (Xu et al., 2015). In Tartary buckwheat, SG4, SG5, SG6, SG7 subfamily MYBs are responsible for this process (Zhang et al., 2017, 2018, 2019; Li et al., 2019c; Yao et al., 2020; Ding et al., 2021). In order to learn more about the genes in the regulatory network of flavonoid biosynthesis, we collected all DEGs annotated to flavonoid biosynthesis genes and SG4, SG5, SG6, SG7 subfamily MYBs and resulting in 12 flavonoid biosynthesis DEGs and nine MYB DEGs (Figure 4A). Pearson correlations of the flavonoid biosynthesis DEGs and MYB DEGs were calculated. The results showed that an SG5 subfamily MYB, *FcMYB1*, was positively correlated to eight flavonoid biosynthesis genes, including two *LAR* homologs (*FcLAR2*, Pearson $R = 0.91$, and *FcLAR1*, Pearson $R = 0.59$), a *DFR* homolog (*FcDFR1*, Pearson $R = 0.58$), an *F3H* homolog (*FcF3H1*, Pearson $R = 0.71$), two *F3'5'H* homologs (*FcF3'5'H1*, Pearson $R = 0.77$, and *FcF3'5'H2*, Pearson $R = 0.7$), an *F3'H* homolog (*FcF3'H1*, Pearson $R = 0.71$), and a *CHS* homolog (*FcCHS2*, Pearson $R = 0.79$) (Figure 4A). The highly



correlated expression patterns suggested that *FcMYB1* might be an essential regulator in flavonoid biosynthesis.

Previous studies have exhaustively characterized the DNA-binding specificities of R2R3-MYB proteins in relation to flavonoid biosynthesis. Nine DNA-binding specificities could bind to most R2R3-MYB proteins related to flavonoid biosynthesis in plants, namely, TACTGTTG, TGCGGTTG, AAAAGTTA, GTCAGTTA, ACAAGTTA, ATTAGTTG, GCTTGTG, TATGGTTA, and TAGTTA (Kelemen et al., 2015). Therefore, we obtained and scanned the 1,500 bp promoter sequences from the translation initiation site of the eight DEGs in flavonoid biosynthesis that were highly correlated to *FcMYB1* to determine whether these DNA-binding specificities were included in them. Notably, *FcLAR2* had four R2R3-MYB DNA-binding specificities on its promoter; *FcDFR1* had two R2R3-MYB DNA-binding specificities on its promoter; *FcCHS2*, *FcF3'H1*, and *FcF3H1* had one R2R3-MYB DNA-binding specificity on each of their promoters (Figure 4B). These suggested that *FcMYB1* might directly bind to the promoter of these five genes to regulate flavonoid biosynthesis.

Yeast one-hybrid technology was performed to determine the interactions of the FcMYB1 protein and the promoter of five flavonoid biosynthesis genes, namely, *FcCHS2*, *FcLAR2*, *FcDFR1*, *FcF3'H1*, and *FcF3H1*. The constructs of pHIS2-promoter and pGADT7-MYB were co-transformed into yeast strain Y187. The

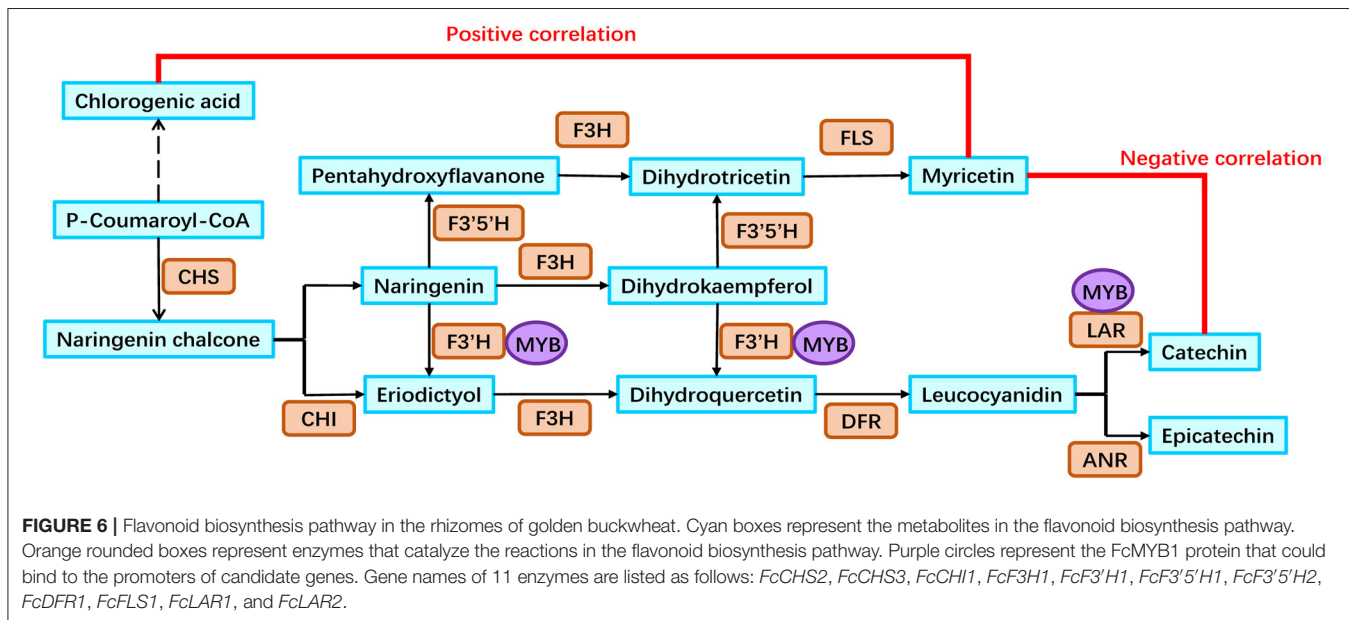
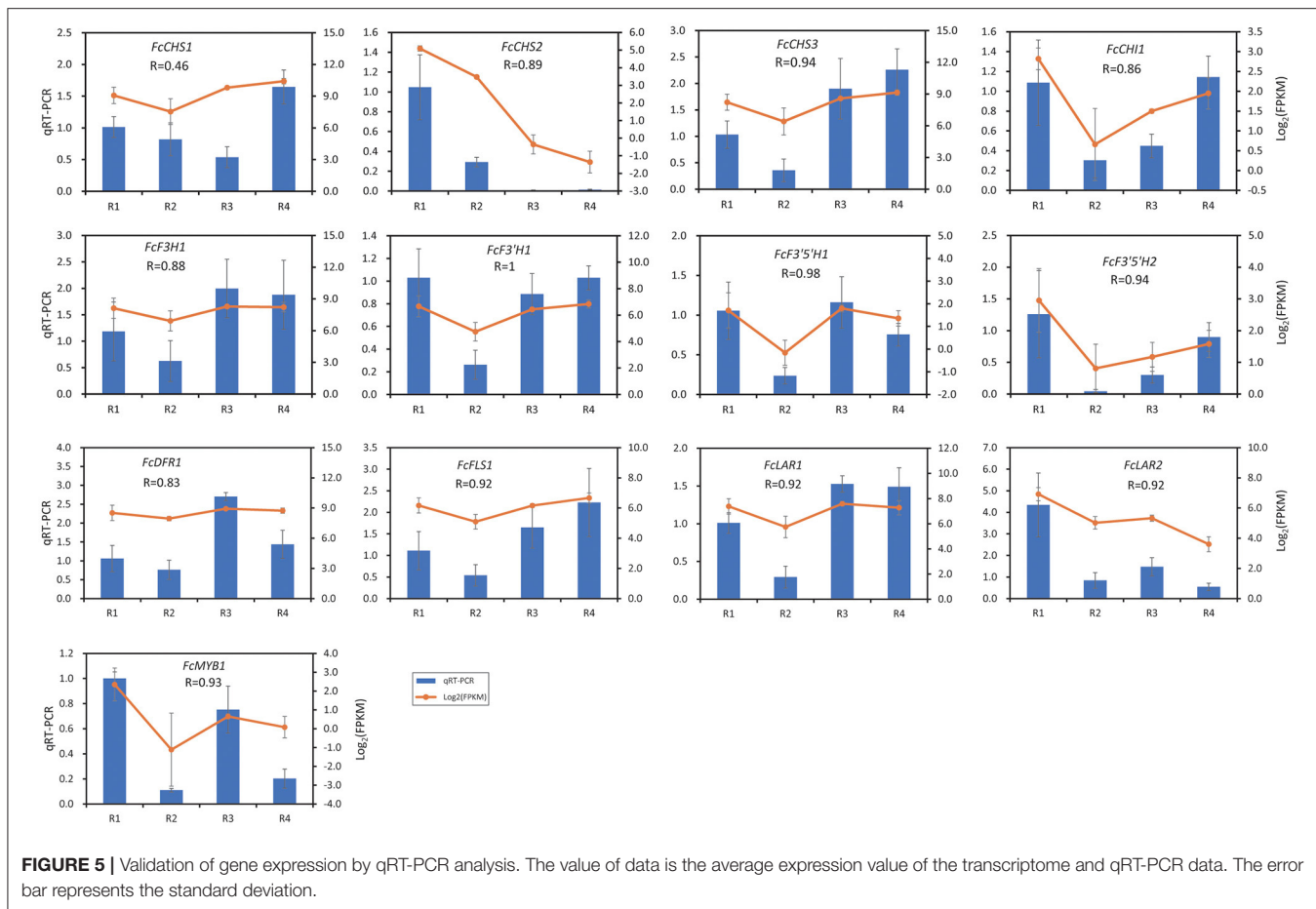
results showed that MYB could bind to the promoter of *FcLAR2* and *FcF3'H1* (Figure 4C) but could not bind to the promoter of the other three genes. This suggested that MYB had a DNA-binding activity and could directly regulate the expression of *FcLAR2* and *FcF3'H1*.

Validation of Gene Expression by qRT-PCR Analysis

The expression of 13 DEGs was validated by qRT-PCR, in which 12 flavonoid biosynthesis DEGs and *FcMYB1* that was mentioned in Figure 4 were included. As shown in Figure 5, the expressions of 12 out of the 13 DEGs were well correlated by qRT-PCR, with a Pearson R ranging from 0.83 to 1. The Pearson R of *FcMYB1* was 0.93. The Pearson R of *FcLAR2* and *FcF3'H1*, whose promoter could be bound by the FcMYB1 protein, was 0.92 and 1, respectively. One CHS gene, namely, *FcCHS1*, showed a lower correlation between the transcriptome and qRT-PCR (Pearson $R = 0.46$), so it was eliminated in the subsequent analysis.

Flavonoid Biosynthesis Pathway in the Rhizomes of Golden Buckwheat

Based on our analyses, a putative flavonoid biosynthesis pathway in the rhizomes of golden buckwheat was proposed (Figure 6).



This included 13 metabolites, 11 genes that were annotated to encode eight enzymes, and one MYB transcription factor.

The expression patterns of genes in this pathway are shown in **Figure 5** and validated by qRT-PCR.

DISCUSSION

First Integrative Transcriptomics and Metabolomics Analyses of Golden Buckwheat

Among the three buckwheat species, Tartary buckwheat is the most thoroughly studied species in molecular biology. This is largely due to its high content of flavonoids, which are known as antioxidants with anti-inflammatory, anticancer, and vascular-protective properties (Fabjan et al., 2003). Since its genome sequencing was finished years ago, flavonoid biosynthesis and its regulatory mechanisms have been extensively reported (Zhang et al., 2017, 2019; Li et al., 2019c; Yao et al., 2020; Ding et al., 2021). Meanwhile, types and amounts of Tartary buckwheat flavonoids were identified and measured (Fabjan et al., 2003; Ma et al., 2019; Li et al., 2019a,b; Yang et al., 2020; Hou et al., 2021). By contrast, only a few studies related to the above areas were reported on golden buckwheat, as its genome is complex and not sequenced yet, which limited its molecular biology development. To date, only one transcriptome analysis of underlying proanthocyanidin accumulation was performed by *de novo* assembly, and eight coding sequences in flavonoid biosynthesis were cloned in golden buckwheat (Liu et al., 2009; Ma et al., 2009, 2012; Meng et al., 2010; Li et al., 2011; Jiang et al., 2013; Luo et al., 2013; Pu et al., 2014; Chen and Li, 2016). In the current transcriptome, we tried to map the clean reads to Tartary buckwheat genome data, in view of *F. tataricum* being very close to *F. cymosum* (Ohnishi and Matsuoka, 1996), which showed the expected results, with a unique mapping percentage ranging from 63.6–72.36%. A total of 28,492 genes based on the genome data of Tartary buckwheat and 1,665 new genes not included in the reference genome data were identified in the rhizomes of golden buckwheat, among which 10,191 genes were DEGs. This would enrich the gene sequence resources of golden buckwheat.

According to the metabolome analysis, 439 metabolites were identified, including 90 flavonoids. The number of flavonoids in golden buckwheat was much less than that in Tartary buckwheat (Ma et al., 2019; Zhao et al., 2019; Li et al., 2019a,b). A recent study determined nine flavonoids from rhizomes of golden buckwheat, namely, catechin, epicatechin, epicatechin gallate, gallic acid, epigallocatechin, catechin gallate, epicatechin gallate, afzelechin, and epiafzelechin, of which catechin and epicatechin together constituted about 90% (Li et al., 2021). Compared with our study, eight of the nine flavonoids were included in our metabolome except for afzelechin. All the other 82 flavonoids were identified in golden buckwheat for the first time. Catechin, gallic acid, hesperidin, nepetin, and L-epicatechin were the top five flavonoids identified in this study, which together constituted about 57% of the total flavonoids. This difference in the percentage of catechin and epicatechin is largely due to the difference in methods. Our metabolome was a large-scale method that included more than 1,000 flavonoids in the constructed flavonoid metabolite database (Chen et al., 2013). The most abundant flavonoid in Tartary buckwheat was rutin, which belongs

to the flavonol subclass and accounts for 50–80% of total flavonoids (Fabjan et al., 2003; Li et al., 2019a; Ma et al., 2019). In our golden buckwheat metabolome data, rutin was 23rd.

To our knowledge, this is the first metabolome study, and the first to integrate metabolome and transcriptome data in golden buckwheat. Based on our analyses, we enriched the expressed genes and the existing metabolites and proposed a flavonoid biosynthesis pathway in golden buckwheat rhizomes.

Integration Analysis of the Transcriptome and Metabolome Revealed a Catechin Integrative Network in the Rhizomes of Golden Buckwheat

A catechin integrative network was identified based on the transcriptome and metabolome integration analyses, in which 14 genes and four metabolites were included. The Pearson correlation of four metabolites indicated that catechin (mws0054) was negatively correlated to myricetin (mws0032) and chlorogenic acids (mws0178 and pmp000544), which implied two competitive synthesis reactions during catechin biosynthesis. The first competitive synthesis reaction occurred during the biosynthesis of chlorogenic acids and flavonoids where they competed for the same substrate, p-coumaroyl-CoA. The formation of p-coumaroyl-CoA by 4CL was a common step for both chlorogenic acid biosynthesis and flavonoid biosynthesis (Kim et al., 2013; Saito et al., 2013; Tohge et al., 2017). After this step, HQT catalyzed p-coumaroyl-CoA to p-coumaroyl quinic acid or hydroxycinnamoyl-CoA shikimate/quinate hydroxycinnamoyl transferase catalyzed p-coumaroyl-CoA to p-coumaroyl shikimic, which started the biosynthesis pathway of chlorogenic acids (Kim et al., 2013). Meanwhile, CHS catalyzed p-coumaroyl-CoA to naringenin chalcone, which started the biosynthesis pathway of flavonoids (Saito et al., 2013; Tohge et al., 2017). Catechin was the most abundant flavonoid, so it was negatively correlated to chlorogenic acids. The second competitive synthesis reaction occurred during the biosynthesis of catechin and myricetin. Although catechin is a kind of flavanol and myricetin is a flavonol, naringenin and dihydrokaempferol are their common upstream substrates. For catechin biosynthesis, naringenin and dihydrokaempferol were first catalyzed by F3'H to form dihydroquercetin; then dihydroquercetin was catalyzed by DFR to form leucocyanidin; finally, leucocyanidin was catalyzed by LAR to form catechin. For myricetin biosynthesis, naringenin and dihydrokaempferol were first catalyzed by F3'5'H to form dihydrotricetin; then, dihydrotricetin was catalyzed by FLS to form myricetin (Saito et al., 2013; Tohge et al., 2017). They competed with the same substrates during their biosynthesis, so they were negatively correlated to each other.

It is worth mentioning that the enzymes in the network were proposed based on the gene expression in transcriptome data. Although there are some reports that construct the flavonoid biosynthesis pathway based on transcriptome data (Chen and Li, 2016; Chen et al., 2020; Zhang et al., 2022), we do not have experimental evidence for the enzymatic activity of gene

products. Therefore, enzyme activity should be determined for further analyses.

MYB Transcription Factor Positively Regulated Catechin Biosynthesis

In plants, R2R3-MYB transcription factors are involved in the regulation of flavonoid biosynthesis by binding to DNA-binding specificities on promoters of flavonoid biosynthesis genes (Kelemen et al., 2015; Xu et al., 2015). In Tartary buckwheat, massive genes have been identified and, thus, form a comprehensive flavonoid biosynthesis pathway, in which some R2R3-MYB transcription factors are included (Zhang et al., 2017, 2018, 2019; Li et al., 2019c; Yao et al., 2020; Ding et al., 2021). After the release of the complete Tartary buckwheat genome, genes involved in rutin biosynthesis and regulation pathway were comprehensively identified (Zhang et al., 2017). Afterward, several FtMYBs were found to directly regulate flavonoid biosynthesis, among which FtMYB11, FtMYB13, FtMYB14, FtMYB15, and FtMYB16 could directly bind to the MBS site on the promoter of *FtPAL* and repress rutin biosynthesis, FtMYB116 could bind directly to the promoter region of *FtF3'H* and increase the content of rutin and quercetin, and FtMYB6 promoted the activity of *FtF3H* and *FtFLS1* promoters and significantly increased the accumulation of flavonols (Zhang et al., 2018, 2019; Li et al., 2019c; Yao et al., 2020; Ding et al., 2021). However, there has been no reported regulator involved in the flavonoid biosynthesis of golden buckwheat so far. We identified a novel MYB transcription factor, *FcMYB1*, whose expression patterns were positively correlated to eight out of 12 flavonoid biosynthesis genes significantly. The Y1H analysis showed that the FcMYB1 protein could directly bind to the promoters of *FcLAR2* and *FcF3'H1*. F3'H is an essential enzyme that promotes the formation of flavanols in the competitive biosynthesis of flavonols and flavanols, and LAR is the key enzyme catalyzing leucocyanidin to form catechin (Saito et al., 2013; Tohge et al., 2017). This suggested that FcMYB1 might be a transcriptional activator in catechin biosynthesis by directly binding to the promoters of *FcLAR2* and *FcF3'H1*. This is the first identified regulator in the flavonoid biosynthesis of golden buckwheat. It is a very good beginning for further identification of flavonoid regulatory factors including R2R3-MYB, bHLH, WD40 proteins, and the interaction of MYB-bHLH-WD40 complex in golden buckwheat.

CONCLUSION

In conclusion, for the first time, this study performed transcriptome and widely targeted metabolome analyses using four golden buckwheat accessions that differed in total flavonoid content. A total of 10,191 DEGs and 297 DRMs were identified, among which the flavonoid biosynthesis pathway was enriched in both transcriptome and metabolome. Integration analyses of the transcriptome and metabolome profiles revealed a network related to catechin in which 4 metabolites and 14 genes interacted with each other. An SG5 R2R3-MYB transcription factor, *FcMYB1*, was found as a transcriptional activator in catechin

biosynthesis for the first time, as it was positively correlated to eight flavonoid biosynthesis genes in their expression patterns and could directly bind to the promoters of *FcLAR2* and *FcF3'H1*. Finally, a flavonoid biosynthesis pathway in the rhizomes of golden buckwheat was proposed, including 13 metabolites, 11 genes encoding 9 enzymes, and 1 MYB transcription factor. The result provided a comprehensive flavonoid biosynthesis and regulatory network of golden buckwheat.

DATA AVAILABILITY STATEMENT

The raw sequence data are deposited in the Genome Sequence Archive in the National Genomics Data Center, Chinese Academy of Sciences, under accession number CRA003372 which is publicly accessible at <https://ngdc.cnca.ac.cn/gsa>.

AUTHOR CONTRIBUTIONS

JH and QC designed and coordinated the study. LW, BT, RR, TS, LZ, and CL planted the materials, took the samples, and carried out the experiments. JH, JD, and LW analyzed the transcriptomics and metabolomics data, made the figures and wrote the draft of the manuscript. All authors approved the final version of the manuscript.

FUNDING

This research was funded by the Joint Project of the National Science Foundation of China and Guizhou Provincial Government Karst Science Research Center (U1812401), National Natural Science Foundation of China (32060508), Science and Technology Foundation of Guizhou Province (QianKeHeJiChu-ZK[2021] General 109), and National Key Research and Development Program of China (2019YFD1001300/2019YFD1001304).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpls.2022.803472/full#supplementary-material>

Supplementary Figure 1 | Sample cluster of the transcriptome. Each square indicates the Pearson's correlation coefficient of a pair of samples.

Supplementary Figure 2 | Top 30 enriched Gene Ontology (GO) items of the transcriptome. The results are summarized in three main categories: biological process, cellular component, and molecular function.

Supplementary Figure 3 | Top 20 enriched Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways of the transcriptome. The rich factor is the ratio of the number of DEGs to that of all genes annotated to a pathway term. A higher rich factor indicates greater intensity.

Supplementary Figure 4 | PCA analysis of the metabolome.

Supplementary Figure 5 | K means a cluster of the differentially regulated metabolites. The x-axis corresponds to the samples and the y-axis to standardized intensity.

Supplementary Figure 6 | Top 20 enriched KEGG pathways of the metabolome. The rich factor is the ratio of the number of DEGs to that of all genes annotated to a pathway term. A higher rich factor indicates greater intensity.

Supplementary Table 1 | List of all the primers used in this study.

Supplementary Table 2 | Summary statistics of RNA sequencing results of golden buckwheat rhizomes.

Supplementary Table 3 | Differentially expressed genes (DEGs) identified among the four golden buckwheat rhizomes. DESeq2 was used to identify the DEGs, and

thresholds were set as the absolute value of $\log_2(\text{ratio}) \geq 1$ and $P\text{-value (padj)} < 0.05$.

Supplementary Table 4 | All metabolites identified in golden buckwheat rhizomes.

Supplementary Table 5 | DRMs identified in golden buckwheat rhizomes. MetaboAnalystR was used to identify the DRMs by variable importance in the projection value (VIP) ≥ 1 and absolute $\log_2(\text{ratio}) \geq 1$.

Supplementary Table 6 | Correlations between genes and metabolites in flavonoid biosynthesis (KEGG pathway ko00941).

REFERENCES

- Bosse, M. A., Silva, M. B. D., Oliveira, N. G. R. M., Araujo, M. A., Rodrigues, C., Azevedo, J. P., et al. (2021). Physiological impact of flavonoids on nodulation and ureide metabolism in legume plants. *Plant Physiol. Biochem.* 166, 512–521. doi: 10.1016/j.plaphy.2021.06.007
- Chan, P. K. (2003). Inhibition of tumor growth *in vitro* by the extract of fagopyrum cymosum (fago-c). *Life Sci.* 72, 1851–1858. doi: 10.1016/S0024-3205(03)00013-4
- Chen, C., Lan, J., Xie, S., Ui, S., and Li, A. (2012). *In vitro* propagation and quality evaluation of Long-term micro-propagated and conventionally grown *Fagopyrum dibotrys* Hara mutant, an important medicinal plant. *J. Med. Plants Res.* 5, 3003–3012. doi: 10.5897/JMPR11.1508
- Chen, C., and Li, A. (2016). Transcriptome analysis of differentially expressed genes involved in proanthocyanidin accumulation in the rhizomes of fagopyrum dibotrys and an irradiation-induced mutant. *Front. Physiol.* 7:100. doi: 10.3389/fphys.2016.00100
- Chen, J., Wang, J., Wang, R., Xian, B., Ren, C., Liu, Q., et al. (2020). Integrated metabolomics and transcriptome analysis on flavonoid biosynthesis in safflower (*Carthamus tinctorius* L.) under MeJA treatment. *BMC Plant Biol.* 20, 353. doi: 10.1186/s12870-020-02554-6
- Chen, Q.-F., Huang, X.-Y., Li, H.-Y., Yang, L.-J., and Cui, Y.-S. (2018). Recent progress in perennial buckwheat development. *Sustainability* 10:536. doi: 10.3390/su10020536
- Chen, W., Gong, L., Guo, Z., Wang, W., Zhang, H., Liu, X., et al. (2013). A novel integrated method for large-scale detection, identification, and quantification of widely targeted metabolites: application in the study of rice metabolomics. *Mol. Plant* 6, 1769–1780. doi: 10.1093/mp/sst080
- Chiang, Y. C., Yang, C. K., Shih, J. C., and Lee, C. N. (2000). The Pharmacopoeia of the People's Republic of China. (English edition). *Can. J. Behav. Ence* 26, 534–550.
- Chong, J., and Xia, J. (2018). MetaboAnalystR: an R package for flexible and reproducible analysis of metabolomics data. *Bioinformatics* 34, 4313–4314. doi: 10.1093/bioinformatics/bty528
- Ding, M., Zhang, K., He, Y., Zuo, Q., Zhao, H., He, M., et al. (2021). FtBPM3 modulates the orchestration of FtMYB11-mediated flavonoids biosynthesis in Tartary buckwheat. *Plant Biotechnol. J.* 19, 1285–1287. doi: 10.1111/pbi.13587
- Fabjan, N., Rode, J., Kosir, I. J., Wang, Z., Zhang, Z., and Krefl, I. (2003). Tartary buckwheat (*Fagopyrum tataricum* Gaertn.) as a source of dietary rutin and quercitrin. *J. Agric. Food Chem.* 51, 6452–6455. doi: 10.1021/jf034543e
- Falcone Ferreyra, M. L., Rius, S., and Casati, P. (2012). Flavonoids: biosynthesis, biological functions, and biotechnological applications. *Front. Plant Sci.* 3:222. doi: 10.3389/fpls.2012.00222
- González, I., Déjean, S., Martin, P. G. P., and Baccini, A. (2008). CCA: an R package to extend canonical correlation analysis. *J. Stat. Softw.* 23, 1–14. doi: 10.18637/jss.v023.i12
- Gu, L., Huang, X., Li, Y., and Chen, Q. (2014). Production technology optimization on fermented tea of golden buckwheat. *Guizhou Agric. Sci.* 42, 169–171+174. doi: 10.3969/j.issn.1001-3601.2014.04.043
- Hou, S., Du, W., Hao, Y., Han, Y., Li, H., Liu, L., et al. (2021). Elucidation of the regulatory network of flavonoid biosynthesis by profiling the metabolome and transcriptome in tartary buckwheat. *J. Agric. Food Chem.* 69, 7218–7229. doi: 10.1021/acs.jafc.1c00190
- Jia, C. F., and Li, A. L. (2008). Effect of gamma radiation on mutant induction of *Fagopyrum dibotrys* Hara. *Photosynthetica* 46, 363–369. doi: 10.1007/s11099-008-0066-0
- Jiang, J., Bai, Y. C., Li, C. L., Chen, H., and Wu, Q. (2013). Cloning of flavonol synthase gene from *Fagopyrum dibotrys* and its expression in *Escherichia coli*. *Chinese Tradit. Herb. Drugs* 44, 1974–1978. doi: 10.7501/j.issn.0253-2670.2013.14.020
- Jozefczuk, S., Klie, S., Catchpole, G., Szymanski, J., Cuadros-Inostroza, A., Steinhäuser, D., et al. (2010). Metabolomic and transcriptomic stress response of *Escherichia coli*. *Mol. Syst. Biol.* 6:364. doi: 10.1038/msb.2010.18
- Kelemen, Z., Sebastian, A., Xu, W., Grain, D., Salsac, F., Avon, A., et al. (2015). Analysis of the DNA-binding activities of the arabidopsis R2R3-MYB transcription factor family by one-hybrid experiments in yeast. *PLoS ONE* 10:e0141044. doi: 10.1371/journal.pone.0141044
- Kim, D., Langmead, B., and Salzberg, S. L. (2015). HISAT: a fast spliced aligner with low memory requirements. *Nat. Methods* 12, 357–360. doi: 10.1038/nmeth.3317
- Kim, Y. B., Thwe, A. A., Kim, Y. J., Li, X., Kim, H. H., Park, P. B., et al. (2013). Characterization of genes for a putative hydroxycinnamoyl-coenzyme A quinate transferase and p-coumarate 3'-hydroxylase and chlorogenic acid accumulation in tartary buckwheat. *J. Agric. Food Chem.* 61, 4120–4126. doi: 10.1021/jf4000659
- Li, C., Feng, Z., Bai, Y., Hui, C., and Qi, W. (2011). Molecular cloning and prokaryotic expression of phenylalanine ammonia-lyase gene FdPAL from *Fagopyrum dibotrys*. *Zhongguo Zhong Yao Za Zhi* 36, 3238–3243. doi: 10.4268/cjcm20112305
- Li, H., Lv, Q., Ma, C., Qu, J., Cai, F., Deng, J., et al. (2019a). Metabolite profiling and transcriptome analyses provide insights into the flavonoid biosynthesis in the developing seed of tartary buckwheat (*Fagopyrum tataricum*). *J. Agric. Food Chem.* 67, 11262–11276. doi: 10.1021/acs.jafc.9b03135
- Li, H., Xiao, Q., Zhang, C., Du, J., Li, X., Huang, H., et al. (2017). Identification and characterization of transcription factor ZmEREB94 involved in starch synthesis in maize. *J. Plant Physiol.* 216, 11–16. doi: 10.1016/j.jplph.2017.04.016
- Li, J., Yang, P., Yang, Q., Gong, X., Ma, H., Dang, K., et al. (2019b). Analysis of flavonoid metabolites in buckwheat leaves using UPLC-ESI-MS/MS. *Molecules* 24:1310. doi: 10.3390/molecules24071310
- Li, J., Zhang, K., Meng, Y., Li, Q., Ding, M., and Zhou, M. (2019c). FtMYB16 interacts with Ftimportin- α 1 to regulate rutin biosynthesis in tartary buckwheat. *Plant Biotechnol. J.* 17, 1479–1481. doi: 10.1111/pbi.13121
- Li, X., Liu, J., Chang, Q., Zhou, Z., Han, R., and Liang, Z. (2021). Antioxidant and antidiabetic activity of proanthocyanidins from *Fagopyrum dibotrys*. *Molecules* 26:2417. doi: 10.3390/molecules26092417
- Liu, G. D., Lei, X. H., Zhu, Q. L., Zhong, G. C., Guo, T. Y., Yan-Dong, L. I., et al. (2009). Cloning and Sequence Analysis of a DFR Gene from *Fagopyrum dibotrys* (D. Don) Hara. *Sci. Agric. Sin.* 1, 55–63. doi: 10.3864/j.issn.0578-1752.2009.01.007
- Love, M. I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* 15:550. doi: 10.1186/s13059-014-0550-8
- Luo, X. P., Bai, Y. C., Gao, F., Li, C. L., and Wu, Q. (2013). Gene cloning and expression level of chalcone isomerase during florescence and content of flavonoids in *Fagopyrum dibotrys*. *Chin Tradit. Herb. Drugs* 44, 1481–1485. doi: 10.7501/j.issn.0253-2670.2013.11.023
- Ma, J., Wang, B., Dai, Y., Sui, S. Z., and Li, M. Y. (2012). Cloning and expression analysis of leucoanthocyanidin reductase gene in *Fagopyrum dibotrys*. *Yao Xue Xue Bao* 47, 953–961.

- Ma, J., Zhu, Q., Guo, T., Liu, G., and Li, M. (2009). Cloning and molecular characteristics analysis of a MYB gene from *Fagopyrum dibotrys*. *China J. Chin. Mater. Medica* 34, 2155–2159. doi: 10.3321/j.issn:1001-5302.2009.17.005
- Ma, W., Kim, J. K., Jia, C., Yin, F., Kim, H. J., Akram, W., et al. (2019). Comparative transcriptome and metabolic profiling analysis of buckwheat (*Fagopyrum Tataricum* (L.) Gaertn.) under salinity stress. *Metabolites* 9:225. doi: 10.3390/metabo9100225
- Mazza, G., and Oomah, B. D. (2003). *BUCKWHEAT - Encyclopedia of Food Sciences and Nutrition (Second Edition)*. doi: 10.1016/B0-12-227055-X/00132-2
- Meng, H., Cheng-Lei, L. I., Qi, W. U., Shao, J. R., and Chen, H. (2010). Cloning and sequence analysis of the chalcone synthase gene (CHS) from *Fagopyrum dibotrys*. *Acta Prataculturae Sin.* 19, 162–169.
- Ohnishi, O., and Matsuoka, Y. (1996). Search for the wild ancestor of buckwheat II. Taxonomy of *Fagopyrum* (Polygonaceae) species based on morphology, isozymes and cpDNA variability. *Genes Genet. Syst.* 71, 383–390. doi: 10.1266/ggs.71.383
- Pu, X. X., Luo, X. P., Bai, Y. C., Li, C. L., and Wu, Q. (2014). Gene cloning of anthocyanin synthase in *Fagopyrum dibotrys* and correlation between its expression level and anthocyanin content. *Chin. Tradit. Herb. Drugs* 45, 985–989. doi: 10.7501/j.issn.0253-2670.2014.07.017
- Saito, K., Yonekura-Sakakibara, K., Nakabayashi, R., Higashi, Y., Yamazaki, M., Tohge, T., et al. (2013). The flavonoid biosynthetic pathway in *Arabidopsis*: structural and genetic diversity. *Plant Physiol. Biochem.* 72, 21–34. doi: 10.1016/j.plaphy.2013.02.001
- Shannon, P., Markiel, A., Ozier, O., Baliga, N. S., Wang, J. T., Ramage, D., et al. (2003). Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res.* 13, 2498–2504. doi: 10.1101/gr.1239303
- Shao, M., Yang, Y. H., Gao, H. Y., Bin, W. U., Wang, L. B., and Li-Jun, W. U. (2005). Studies on the chemical constituents of *Fagopyrum dibotrys* (D. Don) Hara. *J. Shenyang Pharm. Univ.* 02, 100–102+160. doi: 10.3969/j.issn.1006-2858.2005.02.005
- Sun, W., Ma, Z., and Liu, M. (2020). Cytochrome P450 family: genome-wide identification provides insights into the rutinsynthesis pathway in Tartary buckwheat and the improvement of agricultural product quality. *Int. J. Biol. Macromol.* 164, 4032–4045. doi: 10.1016/j.ijbiomac.2020.09.008
- Tohge, T., de Souza, L. P., and Fernie, A. R. (2017). Current understanding of the pathways of flavonoid biosynthesis in model and crop plants. *J. Exp. Bot.* 68, 4013–4028. doi: 10.1093/jxb/erx177
- Tohge, T., Perez de Souza, L., and Fernie, A. R. (2018). On the natural diversity of phenylacylated-flavonoid and their in planta function under conditions of stress. *Phytochem. Rev.* 17, 279–290. doi: 10.1007/s11101-017-9531-3
- Wang, K. J., Zhang, Y. J., and Yang, C. R. (2005). Antioxidant phenolic constituents from *Fagopyrum dibotrys*. *J. Ethnopharmacol.* 99, 259–264. doi: 10.1016/j.jep.2005.02.029
- Wang, L., Rong, Y., Huang, J., and Chen, Q. (2019). Analysis and evaluation of the flavonoid content of rhizomes of 211 different Golden buckwheat accessions (*Fagopyrum cymosum* complex). *J. Guizhou Norm. Univ.* 37, 25–30. doi: 10.16614/j.gznuj.zrb.2019.04.005
- Xu, W., Dubos, C., and Lepiniec, L. (2015). Transcriptional control of flavonoid biosynthesis by MYB–bHLH–WDR complexes. *Trends Plant Sci.* 20, 176–185. doi: 10.1016/j.tplants.2014.12.001
- Yang, B., Xin, M. T., Lu, D. M., Qu, Q. L., and Huang, L. Q. (2007). Studies on HPLC-fingerprint of Rhizoma *Fagopyri Dibotryis*. *China J. Chin. Mater. medica* 32, 1171–1174. doi: 10.3321/j.issn:1001-5302.2007.12.014
- Yang, W., Su, Y., Dong, G., Qian, G., Shi, Y., Mi, Y., et al. (2020). Liquid chromatography–mass spectrometry-based metabolomics analysis of flavonoids and anthraquinones in *Fagopyrum tataricum* L. Gaertn. (tartary buckwheat) seeds to trace morphological variations. *Food Chem.* 331:127354. doi: 10.1016/j.foodchem.2020.127354
- Yao, P., Huang, Y., Dong, Q., Wan, M., Wang, A., Chen, Y., et al. (2020). FtMYB6, a light-induced SG7 R2R3-MYB transcription factor, promotes flavonol biosynthesis in tartary buckwheat (*Fagopyrum tataricum*). *J. Agric. Food Chem.* 68, 13685–13696. doi: 10.1021/acs.jafc.0c03037
- Zhang, D., Jiang, C., Huang, C., Wen, D., Lu, J., Chen, S., et al. (2019). The light-induced transcription factor FtMYB116 promotes accumulation of rutin in *Fagopyrum tataricum*. *Plant. Cell Environ.* 42, 1340–1351. doi: 10.1111/pce.13470
- Zhang, F., Ma, Z., Qiao, Y., Wang, Z., Chen, W., Zheng, S., et al. (2022). Transcriptome sequencing and metabolomics analyses provide insights into the flavonoid biosynthesis in *Torreya grandis* kernels. *Food Chem.* 374:131558. doi: 10.1016/j.foodchem.2021.131558
- Zhang, J., Kuang, Y., Liu, L., Yang, S. X., and Zhao, L. C. (2016). Chemical constituents from root tubers of *Fagopyrum dibotrys*. *Chin. Tradit. Herb. Drugs* 47, 722–725. doi: 10.7501/j.issn.0253-2670.2016.05.005
- Zhang, K., Logacheva, M. D., Meng, Y., Hu, J., Wan, D., Li, L., et al. (2018). Jasmonate-responsive MYB factors spatially repress rutin biosynthesis in *Fagopyrum tataricum*. *J. Exp. Bot.* 69, 1955–1966. doi: 10.1093/jxb/ery032
- Zhang, L., Li, X., Ma, B., Gao, Q., Du, H., Han, Y., et al. (2017). The tartary buckwheat genome provides insights into rutin biosynthesis and abiotic stress tolerance. *Mol. Plant* 10, 1224–1237. doi: 10.1016/j.molp.2017.08.013
- Zhao, C., Wang, F., Lian, Y., Xiao, H., and Zheng, J. (2020). Biosynthesis of citrus flavonoids and their health effects. *Crit. Rev. Food Sci. Nutr.* 60, 566–583. doi: 10.1080/10408398.2018.1544885
- Zhao, Y., Zhou, M., Xu, K., Li, J., Li, S., Zhang, S., et al. (2019). Integrated transcriptomics and metabolomics analyses provide insights into cold stress response in wheat. *Crop J.* 7, 857–866. doi: 10.1016/j.cj.2019.09.002
- Zhou, S., Kremling, K. A., Bandillo, N., Richter, A., Zhang, Y. K., Ahern, K. R., et al. (2019). Metabolome-scale genome-wide association studies reveal chemical diversity and genetic control of maize specialized metabolites. *Plant Cell* 31, 937–955. doi: 10.1105/tpc.18.00772

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Huang, Wang, Tang, Ren, Shi, Zhu, Deng, Liang, Wang and Chen. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.