



Integrated Metabolome and Transcriptome Analyses Reveal Dissimilarities in the Anthocyanin Synthesis Pathway Between Different Developmental Leaf Color Transitions in *Hopea hainanensis* (Dipterocarpaceae)

OPEN ACCESS

Edited by:

Saneyoshi Ueno,
Forestry and Forest Products
Research Institute, Japan

Reviewed by:

Xiaopeng Fu,
Huazhong Agricultural University,
China
Jinhuan Chen,
Beijing Forestry University, China

*Correspondence:

Xianbang Wang
wangxb@caf.ac.cn
Zhiqiang Wu
wuzhiqiang@caas.cn

† These authors have contributed
equally to this work

Specialty section:

This article was submitted to
Plant Bioinformatics,
a section of the journal
Frontiers in Plant Science

Received: 07 December 2021

Accepted: 07 February 2022

Published: 03 March 2022

Citation:

Huang G, Liao X, Han Q, Zhou Z,
Liang K, Li G, Yang G, Tembrock LR,
Wang X and Wu Z (2022) Integrated
Metabolome and Transcriptome
Analyses Reveal Dissimilarities
in the Anthocyanin Synthesis Pathway
Between Different Developmental
Leaf Color Transitions in *Hopea
hainanensis* (Dipterocarpaceae).
Front. Plant Sci. 13:830413.
doi: 10.3389/fpls.2022.830413

Guihua Huang^{1†}, Xuezhu Liao^{2†}, Qiang Han¹, Zaizhi Zhou¹, Kunnan Liang¹,
Guangyou Li¹, Guang Yang³, Luke R. Tembrock⁴, Xianbang Wang^{1*} and Zhiqiang Wu^{2*}

¹ State Key Laboratory of Tree Genetics and Breeding, Research Institute of Tropical Forestry, Chinese Academy of Forestry, Guangzhou, China, ² Shenzhen Branch, Guangdong Laboratory for Lingnan Modern Agriculture, Genome Analysis Laboratory of the Ministry of Agriculture and Rural Affairs, Agricultural Genomics Institute at Shenzhen, Chinese Academy of Agricultural Sciences, Shenzhen, China, ³ Guangdong Eco-Engineering Polytechnic, Guangzhou, China, ⁴ Department of Agricultural Biology, Colorado State University, Fort Collins, CO, United States

Changes in plant leaf color during development are directly related to the accumulation or degradation of certain phytochemicals such as anthocyanins. Since some anthocyanins can be beneficial to human health and provide insights into the biology of leaves, the underlying processes and timing by which plants produce these molecules has been the focus of numerous studies. The tree species *Hopea hainanensis* generally produces green leaves at all growth stages; however, a few explored individuals have been identified possessing red leaves on the top of the seedlings at a young stage. While the phenomenon of leaf color varying with age has been studied in several species, the underlying mechanisms are largely unknown in *H. hainanensis*. Using a metabolomics approach, the young red leaves in *H. hainanensis* were found to contain higher levels of anthocyanins and flavonoids than the young green-leaved individuals. Among anthocyanins, pelargonidin and cyanidin were the most likely candidates contributing to the red color of the young leaves. Transcriptome results indicated the genes related to the production of these anthocyanins were significantly upregulated, leading to greater accumulation of red pigments. Specifically, the expression of several *MYB* and *bHLH* genes in young red leaf lines was significantly higher than that in the young green leaf lines, especially *HhMYB66*, *HhMYB91*, *HhMYB6*, and *HhbHLH70*. As such these four transcription factors are probably the main regulatory genes resulting in young red leaves in *H. hainanensis*. From these results, comparative analyses with other species can be made to better understand the evolution of pigment biosynthesis and how anthocyanins function in plant metabolism and evolution/adaptation.

Keywords: plant pigments, leaf maturation, trees, anthocyanins, transcription factors

INTRODUCTION

Differences in leaf color have been observed by humans for millennia, especially in temperate climates where many trees reveal colorful pigments in autumn as chlorophyll is degraded and secondary pigments are made evident. Beyond the aesthetics of leaf coloration, pigment molecules play an essential role in leaf physiology and metabolism through harvesting light and providing protection from the damaging effects of UV radiation (Demmig-Adams and Iii, 1996; Chen, 2014; Rahnasto-Rilla et al., 2018). The protective effects of some plant pigment molecules, such as anthocyanins, are known to provide similar functions in human biochemistry, and thus a great deal of work has been conducted to better understand these molecules and how they can be used medicinally (Santos-Buelga et al., 2014; Demmig-Adams et al., 2020; Jideani et al., 2021). Therefore, improving our understanding of plant pigment biosynthetic pathways could have profound influences on both plant and human health. For example, the development of cultivars with specific leaf or fruit pigmentation could be improved by molecular breeding and/or genetic engineering of specific steps in pigment biosynthesis gene networks only after the pathways are properly mapped.

Hopea hainanensis Merr. et Chun (坡垒 *po lei*) is a tropical evergreen tree species in the Dipterocarpaceae family. The current range of this species is restricted to scattered occurrences in dense tropical forests on the island of Hainan and a small number of climactically similar locations in Northern Vietnam. This species is well-known due to the fine-grained durable hardwood it possesses, which has historically been used for a wide variety of applications such as in the production of railroad sleepers, mechanical appliances, fishing vessels, docks, bridges, and in constructing buildings. Since this species is now uncommon in the wild due to intensive logging and habitat loss (Ly et al., 2018), the uses of the wood are now mainly restricted to making decorative furniture. In addition to timber production, white aromatic resin from the trunk of *H. hainanensis* has been used in making medicine, perfume, and paint. The characterization of *H. hainanensis* for medicinal compounds has resulted in the discovery of several novel and bioactive compounds, including the acetylcholinesterase inhibitor hopeahainol (Ge et al., 2008, 2009). Moreover, *H. hainanensis* is a deep-rooted tree with tolerance to rocky and shallow soils making it highly valuable in ecosystem services related to soil and water preservation and as such it has been often used in landscaping and ecological conservation projects.

Most recent research on *H. hainanensis* has focused on conservation biology and ecology, seed germination, breeding, reintroduction, and silvicultural attributes (Xiaoying et al., 2017; Lu et al., 2020). Although some chemistry work has been conducted regarding the medical chemistry of *H. hainanensis* (Ge et al., 2008, 2009), relatively little work has been completed regarding the developmental biology of leaf pigment phytochemistry in this species.

In the course of previous studies on *H. hainanensis*, we found that some trees possessed green leaves through the developmental process, whereas others produced leaves that were initially red when first developing and eventually became green as the leaves

matured. Breeding experiments between the two leaf forms indicated that the trait was heritable and as such likely not the results of a plastic gene by environment response. Given the desirable nature of this trait in the development of ornamental cultivars and the potential underlying adaptive characteristics associated with pigmentation (such as adaption to different light levels), it was determined that further work should be conducted to map the genetic pathways of these two color forms. Young red leaf forms have been studied in several plant species (Han et al., 2019; Waterland et al., 2019; El-Nakhel et al., 2020); however, it has not been characterized in *H. hainanensis*. Given this gap in knowledge, we carried out metabolome and transcriptome analyses of different colored leaves at different stages of development for *H. hainanensis* to answer the following questions regarding red leaf pigmentation: (1) what phytochemicals are responsible for the red coloration in young red leaves, (2) which genes are upregulated in the biosynthesis of red pigments, and (3) how do these attributes differ between the two young leaf forms and to other known pigment pathways? The results from these investigations will elucidate the mechanisms responsible for young red leaves in *H. hainanensis* and provide a thoroughly described trait for testing hypotheses regarding the ecology and evolution of the two leaf forms and improve our understanding of how increased red pigmentation in young leaves could influence other biosynthetic pathways especially as pertains to medicinal compounds, environmental adaptation, and photosynthetic efficiency. Young red leaf lines are considered desirable for horticultural uses, and thus knowledge about the formation of this trait can improve cultivar selection and could lead to the development of new more desirable and resilient cultivars for use in landscaping and ecological restoration. Except for cultivar characterization, the inclusion of different phytochemical variants into the germplasm repositories (especially among rare trees like *H. hainanensis*) is essential to preserve the diversity of traits for breeding and introduction.

RESULTS

Metabolites in *Hopea hainanensis*

Hopea hainanensis is an economically and ecologically important species in Dipterocarpaceae, which was found to possess young red-leaved individuals in about 10% of the individuals sampled (Figure 1).

To investigate what compounds contribute to the red-colored young leaves in *H. hainanensis*, the total metabolites of RU (red young leaves of red lines), RL (green mature leaves of red lines), GU (green young leaves green lines), and GL (green mature leaves of green lines) were detected by UPLC-MS/MS followed by hierarchical cluster analysis (HCA) to analyze the differences in accumulated metabolites of the different samples. The metabolites possessing similar accumulation patterns were clustered together and indicated that there are more differences in young leaves than mature leaves between the comparison of RU vs. GU and RL vs. GL. However, even in the mature leaves of the two lines, several metabolites were differentially accumulated. The HCA results also showed that some of the metabolites of RU



FIGURE 1 | Different leaf color phenotypes in *Hopea hainanensis*. Young red-leaved example on the left and standard green-leaved example on the right.

were significantly different from those of the other three sample types. Nearly half of all detected metabolites were significantly accumulated in RU, some of which are likely contribute to red young leaves (**Figure 2A**).

In addition, the results of principal component analysis (PCA) with all samples showed that principal component 1 (PC1) and principal component 2 (PC2) accounted for 59.77 and 21.57% of the total variation, respectively. In the PCA, the biological replicates were tightly clustered while the four sample types were broadly separated on the graph. The RU samples were clearly differentiated from the other three samples, yet the mature leaves from the different lines were also plotted separately (and non-overlapping) in the PCA plot. These results not only indicate the repeatability and reliability of the data, but also indicate metabolic differences by stage and type across all samples (**Figure 2B**). To resolve the key metabolites responsible for the young red leaves in *H. hainanensis*, 719 metabolites were detected when all RU, RL, GU, and GL were considered. The most abundant compounds were phenolic acids (114), flavonoids (109), lipids (105), amino acids and derivatives (77), and organic acids (74), among which the flavonoids are considered the most important in this study because of their crucial role in plant pigmentation (**Figure 2C**). Of the 719 metabolites, 636 metabolites were shared across all four sample types, with only four metabolites (including flavonoids related to red color) unique to RU (**Figure 2D**), and

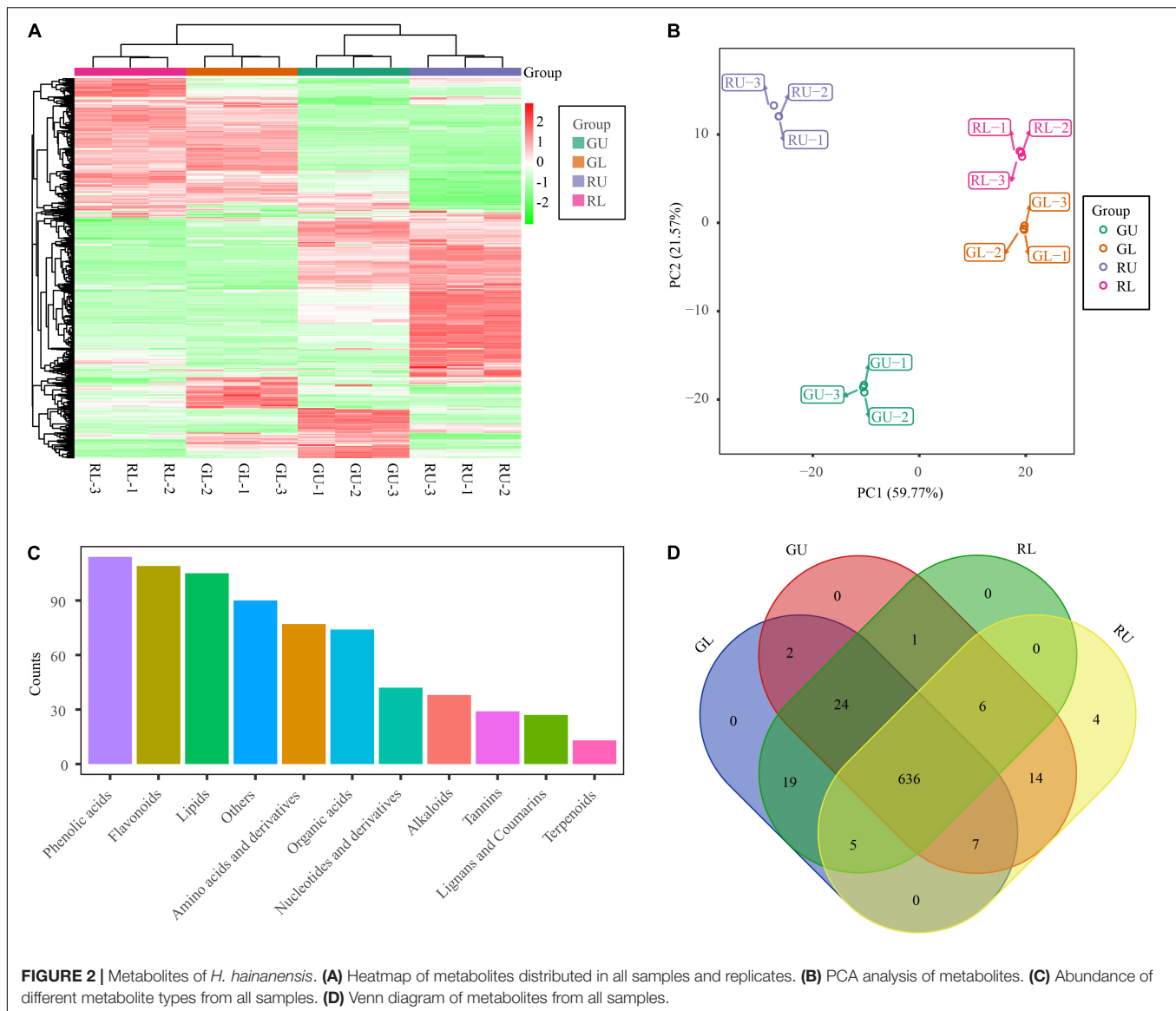
no specific metabolites were uniquely accumulated in the other three samples. Given that flavonoids as well as several unique flavonoids are found in higher abundance in RU, it is reasonable to suggest that compounds of this type are involved with red coloration among these samples (**Figures 2A,C,D**).

Differential Accumulation of Flavonoid Metabolites Between RU and GU

To further confirm that flavonoid metabolites were responsible for the formation of red young leaves in *H. hainanensis*, we compared the metabolites between groups (GL vs. GU, RU vs. GU, RL vs. GL, and RL vs. RU), selected different metabolites in each group based on the screening condition that the variable importance in the projection (VIP) ≥ 1 with $|\text{Log}_2\text{FC}| \geq 1$, and successfully obtained 424 differentially accumulated metabolites (DAMs) in RL vs. RU and 236 DAMs in RU vs. GU, among which 189 metabolites were shared between these two comparisons (**Figure 3A**). We found that RU had 119 metabolites with decreased accumulation and 117 with increased accumulation compared with GU. The RL samples had 224 metabolites with decreased accumulation and 200 metabolites with increased accumulation compared with RU. The RL samples had 49 metabolites with decreased accumulation and 46 metabolites with increased accumulation compared with GL. The GL samples had 183 metabolites with decreased accumulation and 151 metabolites with increased accumulation compared with GU (**Figure 3A**). We further performed k-means clustering on the DAMs in the comparison group of all the samples and divided these differentially accumulated metabolites into 9 subclasses. Among the 9 subclasses 3, 8, and 9 had unique patterns of metabolites in the RU samples. The DAMs in subclasses 3 and 9 were mainly phenolic and organic acids, while subclass 8 was mainly made up of flavonoids, indicating a high accumulation of these compounds in RU relative to GU, GL, and RL (**Figures 3B,C**). This strongly indicates that flavonoid metabolites are the primary compounds responsible for young leaves in *H. hainanensis*.

Transcriptome Analyses

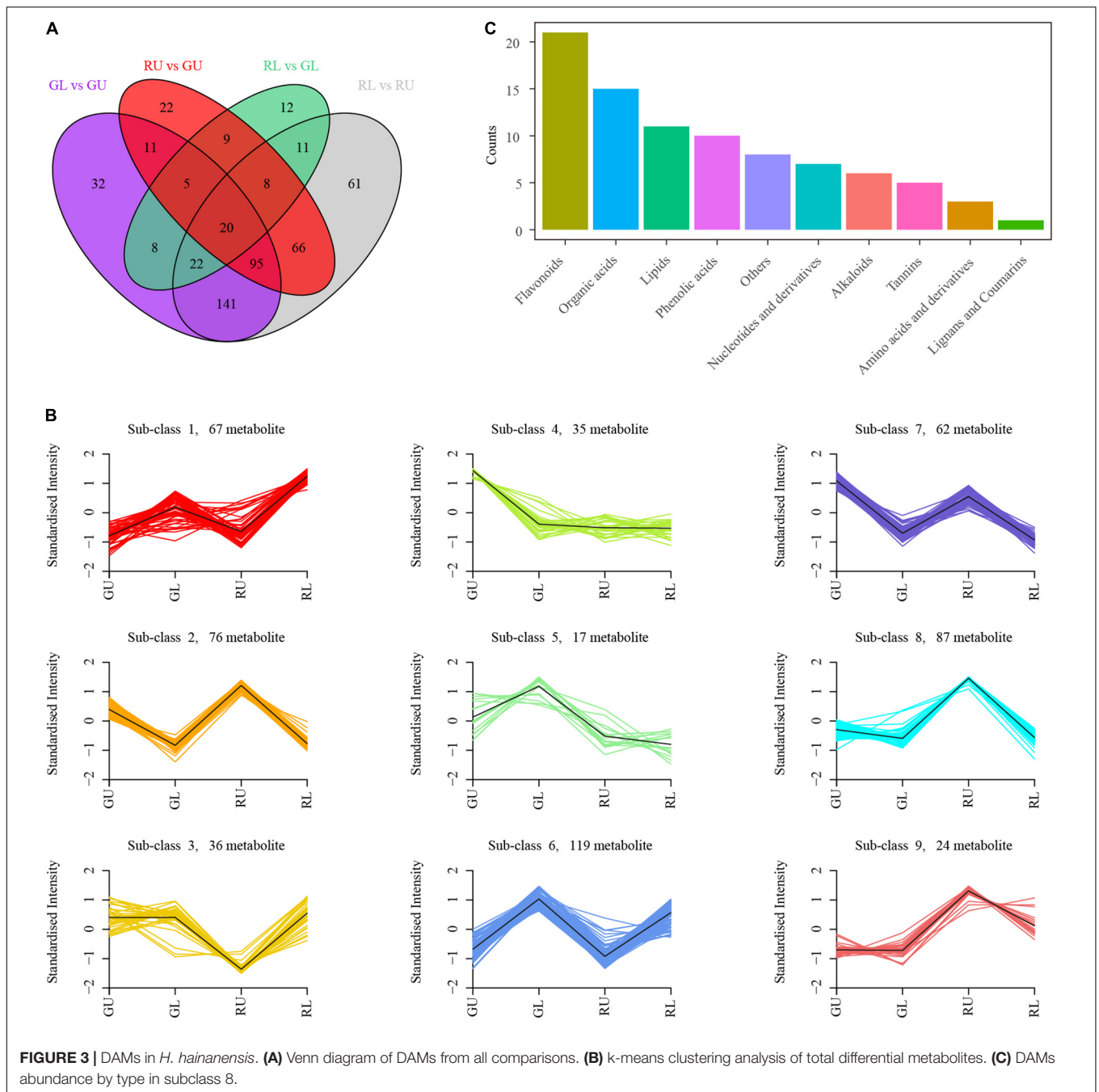
To further confirm metabolite differences between samples RU, RL, GU, and GL, RNA-seq was used to assess the differences in transcript abundance and type. A total of 86,454 transcripts were obtained by Trinity. The longest cluster sequence obtained by Corset was defined as a unigene for subsequent analysis with 83,078 such unigenes generated. Comparisons of GL vs. GU, RL vs. GL, RL vs. RU, and RU vs. GU were used to select differentially expressed genes (DEGs) with the threshold “ $|\log_2\text{Fold Change}| \geq 1$ and $\text{FDR} < 0.05$.” From this 14,195 DEGs (8,203 upregulated and 5,992 downregulated) from GL vs. GU; 7,718 (4,150 upregulated and 3,568 downregulated) from RL vs. GL; 19,488 (10,839 up-regulated and 8,649 downregulated) from RL vs. RU; and 11,584 (6,289 upregulated and 5,295 downregulated) from GL vs. GU were identified (**Figure 4A**). The RL vs. RU comparison contained the greatest number of DEGs and was further scrutinized for pigment-related pathways. The KEGG enrichment results of RU vs. GU showed that DEGs



were significantly enriched in pathways related to flavonoid, stilbenoid, diarylheptanoid, and gingerol biosynthesis as well as photosynthesis (**Figure 4B**). Among the DEGs, the genes related to the flavonoid pathway were the most highly enriched in RU, providing additional evidence that these compounds were responsible for red leaf coloration in *H. hainanensis*.

Gene expression in some instances can be controlled by transcription factors (TFs). The TFs known as trans-acting factors, often regulate plant growth and development as well as the biosynthesis of secondary metabolites by activating or inhibiting gene expression (Latchman, 1993; Schwechheimer and Bevan, 1998). Among them, MYB and bHLH play an important role in the regulation of plant pigment (flavonoid-anthocyanin) biosynthesis (Zhuang et al., 2019). In our study, a total of 1,151 TFs were identified, with the most abundant including AP2/ERF-ERF (88), MYB-related (80), C2H2 (76), WRKY (72), bHLH (55), MYB (52), GRAS (52), NAC (51),

and bZIP (44). Within the top 20 TF types, the transcription factors related to flower color in plants such as MYB accounted for 6.11% and bHLH 6.46% of all TFs (**Figure 4C**). Given that MYB and bHLH are known to be involved with anthocyanin accumulation, it was not surprising to find that some of these TFs were found in significantly higher abundance in RU. Specifically, three MYB TFs *HhMYB66* (Cluster-15967.19622), *HhMYB91* (Cluster-15967.12496), and *HhMYB6* (Cluster-15967.32152) and one bHLH TF *HhbHLH70* (Cluster-15967.18674) were found to be significantly upregulated in RU. The other MYB and bHLH TFs were also upregulated in RU (**Figure 4D**). Additionally, two bHLH TFs, Cluster-15967.47825 and Cluster-27423.0, were highly expressed in RL and GL, respectively, suggesting a putative role in leaf development (**Figure 4D**). Thus, increases in red leaf pigments among RU samples appears to be driven in part by the increased abundance of anthocyanin-related TFs.



Association Analysis of Genes and Metabolites Related to Red Young Leaf Formation in *Hopea hainanensis*

As flavonoids are the main DAMs in the red leaves of *H. hainanensis*, we selectively compared DAMs of phenylpropanoid biosynthesis (ko00940), flavonoid biosynthesis (ko00941), and anthocyanin biosynthesis (ko00942) to assess RU vs. GU and RL vs. RU. From these comparisons, 10 metabolites were found to be significantly accumulated in RU, including the anthocyanin pelargonidin-3-O-glucoside, the

dihydroflavonol naringenin-7-O-glucoside, and the flavonoid myricetin as well as others in each of these three classes (Figure 5). Metabolite pme3392 was annotated as anthocyanin, pelargonidin-3-O-glucoside, which is a pigment known to produce red coloration and is found in the highest relative abundance in RUs (Figure 5B). Therefore, we suggest that this may be the most important metabolite contributing to the coloration of red young leaves.

Based on the results of metabolomic analysis, anthocyanins were the main metabolites found to differ in the red young leaves

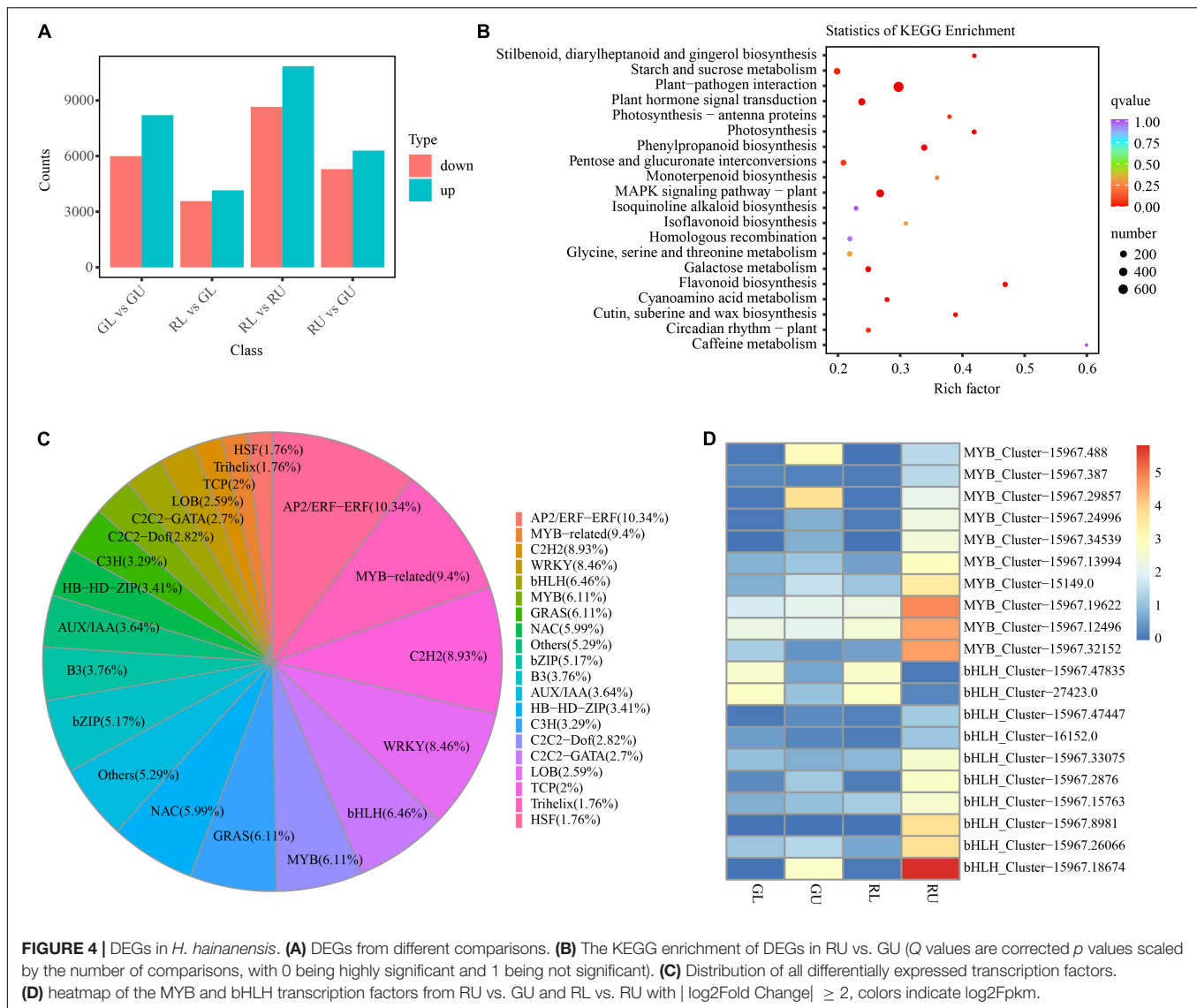


FIGURE 4 | DEGs in *H. hainanensis*. **(A)** DEGs from different comparisons. **(B)** The KEGG enrichment of DEGs in RU vs. GU (Q values are corrected *p* values scaled by the number of comparisons, with 0 being highly significant and 1 being not significant). **(C)** Distribution of all differentially expressed transcription factors. **(D)** heatmap of the MYB and bHLH transcription factors from RU vs. GU and RL vs. RU with $|\log_2\text{Fold Change}| \geq 2$, colors indicate $\log_2\text{Fpkm}$.

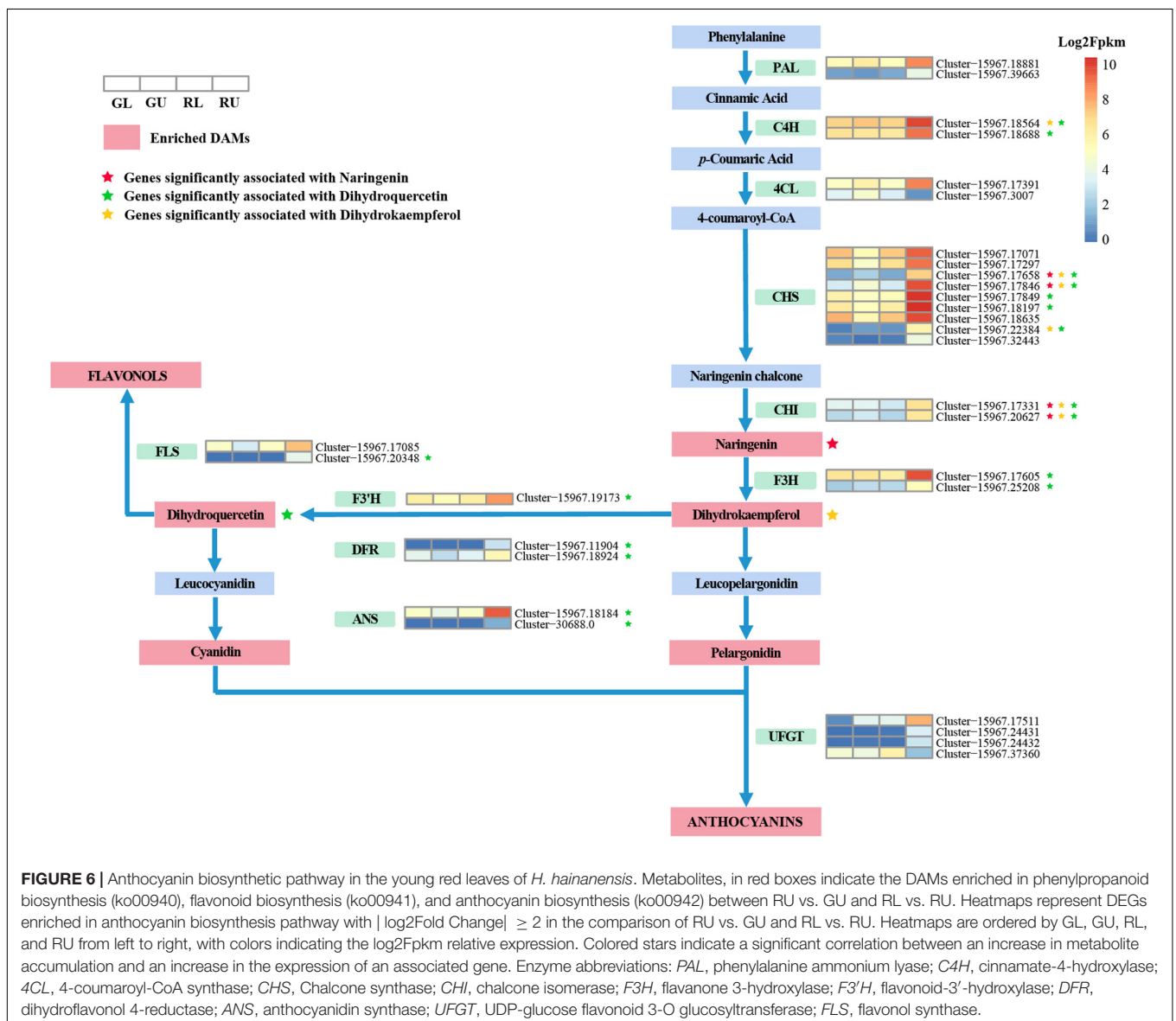
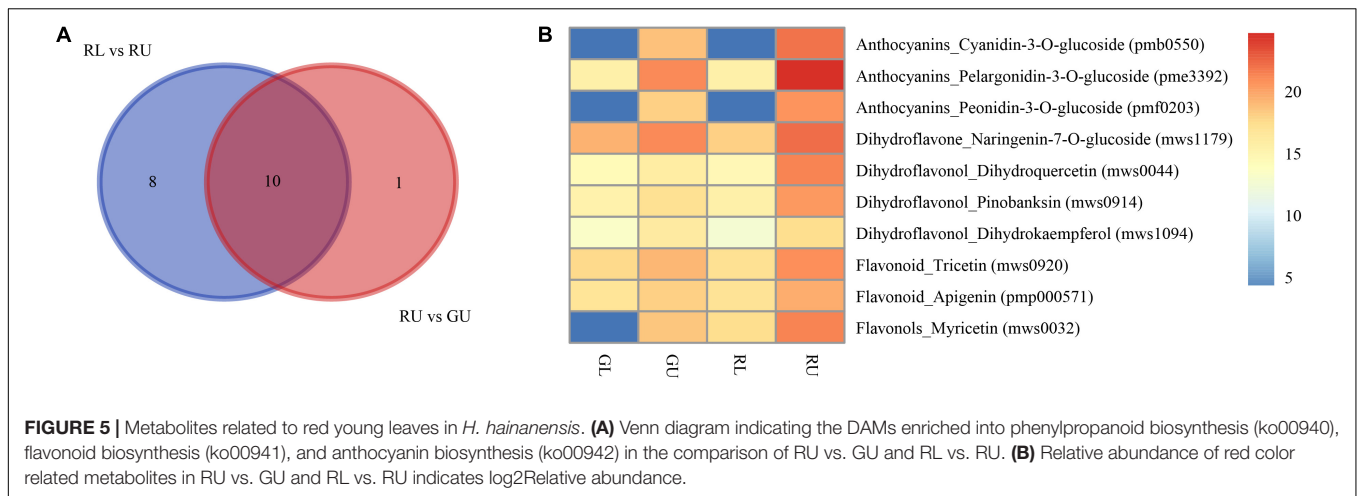
of *H. hainanensis*. Thus, we focused on anthocyanin biosynthesis. DEGs associated with anthocyanin biosynthesis pathway with $|\log_2\text{Fold Change}| \geq 2$ in comparisons of RU vs. GU and RL vs. RU were also selected for further analysis. From this, a total of 30 DEGs (Figure 6) were found, of which 2 *PAL* genes, 2 *C4H* genes, and 1 *4CL* gene, which are involved in precursor biosynthesis of flavonoids, were significantly upregulated in RU. Similarly, the anthocyanidin-associated genes, *CHS*, *CHI*, *F3H*, *F3'H*, *DFR*, and *ANS*, were found to have higher expression levels in RU. In addition, two *FLS* genes which are the key genes in the conversion of dihydroflavonols to flavonols were also significantly upregulated in RU. Interestingly, some downstream *UGT* genes involved in the transport of anthocyanidins were downregulated in RU. When the DEGs and DAMs associated with the anthocyanin biosynthesis pathway in *H. hainanensis* were analyzed separately from the total data set, we found that among the previously screened 30 DEGs, there were 17 genes that had a significant correlation to 3 metabolites, namely

dihydroquercetin (mws0044), dihydrokaempferol (mws1094), and naringenin-7-o-glucoside (mws1179), in the comparison of RU vs. GU (Pearson correlation coefficient (PCC) ≥ 0.8).

DISCUSSION

Hopea hainanensis is a valuable tree that produces high-quality wood and has numerous ecological benefits, but has become increasingly rare in recent decades, which makes the study and preservation of this species even more important. In the process of investigation and research, we found that most of the mature leaves and young leaves of *H. hainanensis* in natural forests are green, while a small minority were found to have green mature leaves and young red leaves. From breeding work, it was found that this trait was fixed and not the result of phenotypic plasticity.

From previous studies, it is well known that the red pigments in most flowers, fruits, and seeds are flavonoid compounds



(Falcone Ferreyra et al., 2012). In this study, the metabolome of *H. hainanensis* leaves were comprehensively analyzed, and over 700 compounds were identified. The metabolome analyses showed that pigments, such as flavonoids were significantly accumulated in the leaf forms differently than those in young green-leaved forms. This indicated that the difference in leaf color mainly resulted from the differences in the accumulation of pigments. Leaf pigments are essential compounds in the metabolic functioning of leaves. Better understanding of the pathways by which these compounds are produced helps researchers to elucidate the evolution of plant pigments and infer their function as well as provide a blueprint for making alterations to a given pathway, resulting in selective and/or increased production of desired compounds. Chalcone synthase (*CHS*) is the first step in controlling flavonoid biosynthesis (Winkel-Shirley, 2001; Chaudhary et al., 2016), and it is the first enzyme identified in the flavonoid biosynthetic pathway. *CHS* is an important regulatory gene located upstream in the flavonoid biosynthesis pathway, and its overexpression may positively affect the expression of downstream *CHI* genes that affect the production of flavonoids (Kreuzaler and Hahlbrock, 1972; Awasthi et al., 2016). Here we saw the upregulation of *CHS*, which is accompanied by the upregulation of *CHI* and its downstream counterpart naringenin. In the anthocyanin biosynthetic pathway, dihydrokaempferol is a key intermediate product that can be converted to dihydroquercetin by the enzyme *F3'H*, and/or converted to dihydromyricetin by *F3'5'H* (Falcone Ferreyra et al., 2012). Here, we found that two *F3H* genes were upregulated, and dihydrokaempferol increased. An *F3'H* gene was found to be highly expressed along with increases in the accumulation of the direct downstream product in the red young leaves. This result is similar to a study that determines the pathway by which red longan (*Dimocarpus longan*) fruits were produced. Yi et al. (2021) revealed that genes related to enzymes leading up to dihydrokaempferol were significantly upregulated in red pericarp longan fruits. While the pathways to produce red pigments might be similar between these two distantly related species, the timing for activating them is quite different as the pathway is initiated in young leaves of *H. hainanensis* and in maturing fruits of longan. Apart from these intermediate products mentioned above, their downstream genes and products were also upregulated in *H. hainanensis*. On the whole, we found *C4H*, *CHS*, *CHI*, *F3H*, *F3'H*, *DFR*, *ANS*, and *FLS* were positively correlated with differentially accumulated metabolic intermediates in the comprehensive analysis of the metabolome and transcriptome, and these genes possessed significantly higher expression levels, which is consistent with results in the purple-leaved Jujube (Li et al., 2021). These results suggest that the high expression of *C4H*, *CHS*, *CHI*, *F3H*, and *F3'H* could promote the accumulation of dihydroquercetin (mws0044) and dihydrokaempferol (mws1094). The accumulation of these two metabolites will recruit and promote the expression of downstream genes, such as *DFR*, *ANS*, and *FLS*, to further generate downstream products. In addition, we found that in the anthocyanin biosynthesis of RU vs. RL, some *UFGT* genes (cluster-15967.49721) were negatively correlated with pelargonidin-3-o-glucoside (−0.831) as the log2Fold Change

of this gene in RU vs. GU was −1.83. Downregulation of some *UFGT* genes was also found in red pericarp longan fruits (Yi et al., 2021) and thought to be involved with increased anthocyanin accumulation. Furthermore, the high expression of *FLS* and the accumulation of the flavonoid apigenin (pmp000571) also revealed that flavonoids, in addition to anthocyanins, pelargonidin, and cyanidin, are also responsible for the red color of young red-leaved lines.

Flavonoid biosynthesis genes are often regulated by the interaction of transcription factors in different gene families. In monocotyledons and dicotyledons, genes involved in the anthocyanin biosynthesis pathway are differentially regulated by R2R3-MYB transcription factors, bHLH, and WD40 proteins (Grotewold, 2005; Petroni and Tonelli, 2011; Xu et al., 2015; Zhuang et al., 2019). Therefore, the combination of R2R3-MYB, bHLH, and WD40 transcription factors and their interactions (MYB-bHLH-WD40 complex) determine the activation and spatio-temporal expression of anthocyanin synthesis structural genes (Petroni and Tonelli, 2011). Unsurprisingly, our results also showed similarities to other studies in apple and grape (Walker et al., 2007; An et al., 2020), wherein the TFs, MYB and bHLH, are the key regulators of anthocyanin biosynthesis and accumulation. The MYB and bHLH transcription factors are present in all eukaryotes and are the two largest families of plant transcription factors (Feller et al., 2011). More than 20 years ago, the first transcription factor encoding proteins in the MYB domain required for anthocyanin synthesis was discovered in plants (Paz-Ares et al., 1987). Since then more studies have found that MYB transcription factors, such as some R2R3-MYBs, including AtMYB75/PAP1, AtMYB90/PAP2, AtMYB113, and AtMYB114 that control anthocyanin biosynthesis in vegetative tissues play an important role in flavonoid biosynthesis (Gonzalez et al., 2008). As for the bHLH transcription factors, the abnormal expression of *bHLH3* in mulberry fruits was shown to disrupt the balance of the flavonoid metabolic network, leading to changes in the content and proportion of anthocyanins, flavonoids, and flavonols in the different colored mulberry fruits (Li et al., 2020). In this process, MYB transcription factors that are mainly involved in the regulation of flavonoid biosynthetic genes, for example, MYB transcription factors SbY1 in *Sorghum bicolor*, may have an impact on the expression of *CHS*, *CHI*, and *DFR*, resulting in regulation of 3-deoxyflavonoid biosynthesis (Du et al., 2009). In *Solanum tuberosum*, StD is thought to regulate the *F3H*, *DFR*, and *F3'5'H* genes, and thus play a role in the regulation of anthocyanin biosynthesis (Jung et al., 2009). In our study, we also found that three MYB TFs *HhMYB66* (Cluster-15967.19622), *HhMYB91* (Cluster-15967.12496), and *HhMYB6* (Cluster-15967.32152), and one bHLH TF *HhbHLH70* (Cluster-15967.18674) were significantly in higher abundance in RU. Moreover, *CHS*, *CHI*, *F3H*, *F3'H*, and *DFR* were also upregulated, implying that these genes may be the key genes regulating flavonoid/anthocyanin synthesis in *H. hainanensis* red leaves.

By combining the results of metabolome and transcriptome data we suggest that there are two main anthocyanin biosynthesis pathways after branching from dihydrokaempferol, and the significant upregulation of genes in these pathways leads to an increased accumulation of anthocyanin and flavonoid

derivatives, which together are responsible for the young red leaves of *H. hainanensis*. Now, the red pigments responsible for young red leaves have been identified, while the follow-up work examining the simultaneously upregulated genes related to stilbenoid production and photosynthesis (found in this study) needs to be conducted to uncover potentially synergistic roles of these phytochemicals in leaf metabolism.

MATERIALS AND METHODS

Plant Materials

During the investigation and cultivation of *H. hainanensis*, we found that most *H. hainanensis* have green mature and young leaves, whereas a small number of individuals possess mature green leaves and young red leaves. After we collected seeds and raised seedlings individually (the seedlings from one tree's seeds are regarded as a line), it was found that their offspring had the same leaf-color, and this pattern was maintained after reproduction, which indicated that genetic differences controlled the leaf color trait. To investigate the metabolomic and transcriptomic differences between the red and green leaf lines, we chose one red leaf individual (referred to as Line 5 in our lab breeding nomenclature) and one green leaf pedigree (referred to as Line 10) for further analyses. The red-leaved individuals were found during the course of forest surveys. Both red and green individuals have been grown from seeds for 2 years in the outdoor common garden at the Research Institute of Tropical Forestry (Chinese Academy of Forestry).

Next, we selected three independent samples from each pedigree to represent the biological duplications (referred to as Red1, Red2, and Red3; Green1, Green2, and Green3). To optimize variation among red and green leaves, we sampled the leaves along the axis of the stem from the 2nd expanded leaf (young leaves) to the 6th unfolded leaf (mature leaves) with a total of 12 samples. Duplicated samples from young red-leaved lines (RU), green mature leaves in red lines (RL), young green leaves in green lines (GU), and green mature leaves in green lines (GL) were harvested and stored at -80°C after snap-freezing with liquid nitrogen.

UPLC-MS/MS Instrumentation and Analyses

The freeze-dried leaves were crushed using a mixer mill (MM 400, Retsch, Haan, German) with a zirconia bead for 1.5 min at 30 Hz. From this 100 mg powder was weighed and extracted overnight at 4°C with 1.2 mL 70% aqueous methanol. Following centrifugation at 12,000 rpm for 10 min, the supernatant was removed and syringe-filtered (SCAA-104, $0.22\mu\text{m}$ pore size; ANPEL, Shanghai, China)¹ before UPLC-MS/MS analysis.

The sample extracts were analyzed using a UPLC-ESI-MS/MS system (UPLC, SHIMADZU Nexera X2²; MS, Applied Biosystems 4500 Q TRAP³). The analytical conditions were

as follows for UPLC: column, Agilent SB-C18 ($1.8\mu\text{m}$, $2.1\text{mm} \times 100\text{mm}$); the mobile phase consisted of solvent A, pure water with 0.1% formic acid, and solvent B, acetonitrile with 0.1% formic acid. Sample measurements were performed with a gradient program that employed the starting conditions of 95% A, 5% B. Within 9 min, a linear gradient to 5% A, 95% B was programmed, and a composition of 5% A, 95% B was kept for 1 min. Subsequently, a composition of 95% A, 5% B was adjusted within 1.10 min and kept for 2.9 min. The column oven was set to 40°C and the injection volume was $4\mu\text{L}$. The effluent was injected into an ESI-triple quadrupole-linear ion trap (QTRAP)-MS.

The mass spectrometry analysis followed the method of Chen et al. (2013). Linear ion trap (LIT) and triple quadrupole (QQQ) scans were acquired on a triple quadrupole-linear ion trap mass spectrometer (Q TRAP), AB4500 Q TRAP UPLC/MS/MS system, equipped with an ESI Turbo Ion-Spray interface and operating in positive and negative ion modes and controlled by Analyst 1.6.3 software (AB Sciex). The ESI source operation parameters were as follows: ion source turbo spray; source temperature 550°C ; ion spray voltage (IS) 5,500 V (positive ion mode)/-4,500 V (negative ion mode); ion source gas I (GSI), gas II (GSII), and curtain gas (CUR) were set at 50, 60, and 25.0 psi, respectively; the collision gas (CAD) was set to high. Instrument tuning and mass calibration were performed with 10 and $100\mu\text{mol/L}$ polypropylene glycol solutions in QQQ and LIT modes, respectively. The QQQ scans were acquired as MRM experiments with collision gas (nitrogen) set to medium. DP and CE for individual MRM transitions were done further with DP and CE optimization. A specific set of MRM transitions were monitored for each period according to the metabolites eluted within this period (Falcone Ferreyra et al., 2012).

Unsupervised PCA (principal component analysis) (Chen et al., 2009) was performed by statistics function `prcomp` within R⁴. The data was normalized before unsupervised PCA. The hierarchical cluster analysis (HCA) results of samples and metabolites were presented as heatmaps with dendrograms, while the Pearson correlation coefficients (PCC) between samples were calculated by the `cor` function in R and presented as heatmaps. Both HCA and PCC were carried out by R package `pheatmap`⁵. For HCA, normalized signal intensities of metabolites (unit variance scaling) are visualized as a color spectrum. Significantly accumulated metabolites between groups were determined by $\text{VIP} \geq 1$ and absolute Log_2FC (fold change) ≥ 1 . VIP values were extracted from orthogonal partial least squares discriminant analysis (OPLS-DA) results, which also contain score plots and permutation plots, and was generated using R package `MetaboAnalystR` (Chong and Xia, 2018). Data were log-transformed (log_2) and mean centered before OPLS-DA. To avoid overfitting, a permutation test (200 permutations) was performed. Identified metabolites were annotated using the KEGG compound database⁶ (Kanehisa and Goto, 2000) and

¹<http://www.anpel.com.cn>

²www.shimadzu.com.cn

³www.appliedbiosystems.com.cn

⁴www.r-project.org

⁵<https://cran.r-project.org/web/packages/pheatmap/index.html>

⁶<http://www.kegg.jp/kegg/compound>

the annotated metabolites were then mapped to the KEGG pathway database⁷. Pathways with significantly accumulated metabolites were then fed into metabolite sets enrichment analysis (MSEA), with significance determined by *p*-values from hypergeometric tests.

RNA Sequencing and Differentially Expressed Genes Analysis

Total RNA was extracted from the young and mature leaves of *H. hainanensis* red and green lines. RNA-Seq was performed by Biomarker Technologies Co., Ltd. (Beijing, China). Sequencing libraries were generated using the NEBNext[®]UltraTM RNA Library Prep Kit for Illumina[®] (New England Biolabs, Ipswich, MA, United States) following the manufacturer's recommendations. Sample-specific indices were added during library preparation such that sequencing could be multiplexed. The library preparations were sequenced on an Illumina HiSeq X platform (Illumina, Inc., San Diego, CA, United States).

After filtering and quality control of the raw sequence data, clean reads were *de novo* assembled by Trinity v2.6.6 (Grabherr et al., 2011; Haas et al., 2013), and Corset v 1.07 (Davidson and Oshlack, 2014) was used to perform hierarchical clustering of the transcripts by comparing the number of reads and expression patterns of the transcripts. The transcript sequences obtained by Trinity were used as the reference sequence, and the longest Cluster sequence obtained by Corset hierarchical clustering was used as a unigene for subsequent analysis. To conduct the gene function annotation, the unigene sequences were compared with KEGG, NR, Swiss-Prot, GO, COG/KOG, and Trembl databases using the BLAST software (Altschul et al., 1990). Then the amino acid sequences predicted from unigenes were compared with Pfam database by HMMER software (Mistry et al., 2013) to obtain unigene annotation information. Plant-transcription factor prediction was performed using iTAK software (Zheng et al., 2016), which integrates PLNTFDB and PLANTTFDB with TFs identified by hmmscan.

The transcriptome assembled by Trinity was used as a reference sequence, and the clean reads of each sample were mapped on the reference to calculate the mapping rate of each sample with bowtie2 in RSEM. DESeq2 v1.22.2 (Varet et al., 2016) was used to obtain the DEGs between any two biological conditions. After that the Benjamini-Hochberg method was used to conduct multiple hypothesis testing correction on the hypothesis test probability (*p*-value) to obtain the false discovery rate (FDR). DEGs were selected under the conditions of $|\log_2\text{Fold Change}| \geq 1$ and $\text{FDR} < 0.05$. Finally, we analyzed the KEGG pathway enrichment in all DEGs and drew the corresponding network regulation pathway map.

CONCLUSION

Integrated metabolome and transcriptome analyses of young and mature leaves of red and green lines of *H. hainanensis* revealed that the accumulation of anthocyanins and flavonoids

in RU was significantly higher in red samples than in the three green samples. The main color-related products in red leaves were anthocyanins and flavonoids, among which pelargonidin and cyanidin were the most important metabolites. Transcriptome results showed that the key genes in the anthocyanin pathway in RU were expressed at significantly higher levels compared to green-leaved samples, leading to the greater accumulation of red-colored metabolites. In addition to the structural genes in the anthocyanin/flavonoid pathway, the TFs MYB and bHLH were highly expressed in RU over the three green-leaved samples. Three MYBs *HhMYB66* (Cluster-15967.19622), *HhMYB91* (Cluster-15967.12496), and *HhMYB6* (Cluster-15967.32152), and one bHLH *HhbHLH70* (Cluster-15967.18674) were expressed at significantly higher levels in RU. As such these four TFs are likely essential to initiate the pathway associated with increased anthocyanin and flavonoids found in RU. While the anthocyanin pathway in *H. hainanensis* is similar to others described from distantly related plant lineages, it is also different in several aspects, such as the absence of delphinidin derivatives. The anthocyanin/flavonoid pathway described here provides a valuable resource for the study of leaf pigment evolution as well as in the development of red-leaved cultivars for use in horticultural and forestry applications.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found below: NCBI; PRJNA795198.

AUTHOR CONTRIBUTIONS

GH designed and supervised implementation of the studies. GL and GY were responsible for sampling. XL and QH supervised the statistical analyses, constructed the tables, and wrote the manuscript. ZZ and KL cultivated the seedlings of *H. hainanensis*. XW and ZW carried out all technical aspects and crafted the final version. LT provided suggestions on various aspects of the study and helped edit the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by the "Fundamental Research Funds for the Central Non-profit Research Institution of CAF" to GH (Grant No. CAFYBB2020SZ005) and the Chinese Academy of Agricultural Sciences Elite Youth Program to ZW.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpls.2022.830413/full#supplementary-material>

⁷<http://www.kegg.jp/kegg/pathway.html>

REFERENCES

- Altschul, S. F., Gish, W., Miller, W., Myers, E. W., and Lipman, D. J. (1990). Basic local alignment search tool. *J. Mol. Biol.* 215, 403–410. doi: 10.1016/S0022-2836(05)80360-2
- An, J. P., Wang, X. F., Zhang, X. W., Xu, H. F., Bi, S. Q., You, C. X., et al. (2020). An apple MYB transcription factor regulates cold tolerance and anthocyanin accumulation and undergoes MIEL1-mediated degradation. *Plant Biotechnol. J.* 18, 337–353. doi: 10.1111/pbi.13201
- Awasthi, P., Mahajan, V., Jamwal, V. L., Kapoor, N., Rasool, S., Bedi, Y. S., et al. (2016). Cloning and expression analysis of chalcone synthase gene from *Coleus forskohlii*. *J. Genet.* 95, 647–657. doi: 10.1007/s12041-016-0680-8
- Chaudhary, P. R., Bang, H., Jayaprakasha, G. K., and Patil, B. S. (2016). Variation in Key Flavonoid Biosynthetic Enzymes and Phytochemicals in 'Rio Red' Grapefruit (*Citrus paradisi* Macf.) during Fruit Development. *J. Agric. Food Chem.* 64, 9022–9032. doi: 10.1021/acs.jafc.6b02975
- Chen, M. (2014). Chlorophyll Modifications and Their Spectral Extension in Oxygenic Photosynthesis. *Annu. Rev. Biochem.* 83, 317–340. doi: 10.1146/annurev-biochem-072711-162943
- 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
- Chen, Y. H., Zhang, R. P., Song, Y. M., He, J. M., Sun, J. H., Bai, J. F., et al. (2009). RRLC-MS/MS-based metabolomics combined with in-depth analysis of metabolic correlation network: finding potential biomarkers for breast cancer. *Analyst* 134, 2003–2011. doi: 10.1039/b907243h
- 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
- Davidson, N. M., and Oshlack, A. (2014). Corset: enabling differential gene expression analysis for de novo assembled transcriptomes. *Genom. Biol.* 15:410. doi: 10.1186/s13059-014-0410-6
- Demmig-Adams, B., and Iii, W. W. A. (1996). The role of xanthophyll cycle carotenoids in the protection of photosynthesis. *Trends Plant Sci.* 1, 21–26. doi: 10.1016/s1360-1385(96)80019-7
- Demmig-Adams, B., Lopez-Pozo, M., Stewart, J. J., and Adams, W. W. III (2020). Zeaxanthin and Lutein: photoprotectors, Anti-Inflammatories, and Brain Food. *Molecules* 25:3607. doi: 10.3390/molecules25163607
- Du, H., Zhang, L., Liu, L., Tang, X. F., Yang, W. J., Wu, Y. M., et al. (2009). Biochemical and molecular characterization of plant MYB transcription factor family. *Biochemistry* 74, 1–11. doi: 10.1134/s0006297909010015
- El-Nakhel, C., Pannico, A., Graziani, G., Kyriacou, M. C., Giordano, M., Ritieni, A., et al. (2020). Variation in macronutrient content, phytochemical constitution and *in vitro* antioxidant capacity of green and red butterhead lettuce dictated by different developmental stages of harvest maturity. *Antioxidants* 9:300. doi: 10.3390/antiox9040300
- Falcone Ferreyra, M. L., Rius, S. P., and Casati, P. (2012). Flavonoids: biosynthesis, biological functions, and biotechnological applications. *Front. Plant Sci.* 3:222. doi: 10.3389/fpls.2012.00222
- Feller, A., Machemer, K., Braun, E. L., and Grotewold, E. (2011). Evolutionary and comparative analysis of MYB and bHLH plant transcription factors. *Plant J.* 66, 94–116. doi: 10.1111/j.1365-313X.2010.04459.x
- Ge, H. M., Yang, W. H., Zhang, J., and Tan, R. X. (2009). Antioxidant oligostilbenoids from the stem wood of *Hopea hainanensis*. *J. Agric. Food Chem.* 57, 5756–5761. doi: 10.1021/jf900756d
- Ge, H. M., Zhu, C. H., Shi da, H., Zhang, L. D., Xie, D. Q., Yang, J., et al. (2008). Hopeahainol A: an acetylcholinesterase inhibitor from *Hopea hainanensis*. *Chemistry* 14, 376–381. doi: 10.1002/chem.200700960
- Gonzalez, A., Zhao, M., Leavitt, J. M., and Lloyd, A. M. (2008). Regulation of the anthocyanin biosynthetic pathway by the TTG1/bHLH/Myb transcriptional complex in *Arabidopsis* seedlings. *Plant J.* 53, 814–827. doi: 10.1111/j.1365-313X.2007.03373.x
- Grabherr, M. G., Haas, B. J., Yassour, M., Levin, J. Z., Thompson, D. A., Amit, I., et al. (2011). Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nat. Biotechnol.* 29, 644–U130. doi: 10.1038/nbt.1883
- Grotewold, E. (2005). Plant metabolic diversity: a regulatory perspective. *Trends Plant Sci.* 10, 57–62. doi: 10.1016/j.tplants.2004.12.009
- Haas, B. J., Papanicolaou, A., Yassour, M., Grabherr, M., Blood, P. D., Bowden, J., et al. (2013). De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. *Nat. Protoc.* 8, 1494–1512. doi: 10.1038/nprot.2013.084
- Han, T., Wang, J., Ren, H., Yi, H., Zhang, Q., and Guo, Q. (2019). Changes in defense traits of young leaves in subtropical forests succession. *Plant Ecol.* 220, 305–320. doi: 10.1007/s11258-019-00916-1
- Jideani, A., Silungwe, H., Takalani, T., Omolola, A., Udeh, H., and Anyasi, T. (2021). Antioxidant-rich natural fruit and vegetable products and human health. *Int. J. Food Prop.* 24, 41–67.
- Jung, C. S., Griffiths, H. M., De Jong, D. M., Cheng, S., Bodis, M., Kim, T. S., et al. (2009). The potato developer (D) locus encodes an R2R3 MYB transcription factor that regulates expression of multiple anthocyanin structural genes in tuber skin. *Theor. Appl. Genet.* 120, 45–57. doi: 10.1007/s00122-009-1158-3
- Kanehisa, M., and Goto, S. (2000). KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Res.* 28, 27–30. doi: 10.1093/nar/28.1.27
- Kreuzaler, F., and Hahlbrock, K. (1972). Enzymatic synthesis of aromatic compounds in higher plants: formation of naringenin (5,7,4'-trihydroxyflavanone) from p-coumaroyl coenzyme A and malonyl coenzyme A. *FEBS Lett.* 28, 69–72. doi: 10.1016/0014-5793(72)80679-3
- Latchman, D. S. (1993). Transcription factors: an overview. *Int. J. Exp. Pathol.* 74, 417–422.
- Li, H., Yang, Z., Zeng, Q. W., Wang, S. B., Luo, Y. W., Huang, Y., et al. (2020). Abnormal expression of bHLH3 disrupts a flavonoid homeostasis network, causing differences in pigment composition among mulberry fruits. *Hort. Res.* 7:83. doi: 10.1038/s41438-020-0302-8
- Li, S., Deng, B., Tian, S., Guo, M., Liu, H., and Zhao, X. (2021). Metabolic and transcriptomic analyses reveal different metabolite biosynthesis profiles between leaf buds and mature leaves in *Ziziphus jujuba* mill. *Food Chem.* 347:129005. doi: 10.1016/j.foodchem.2021.129005
- Lu, X., Zang, R., Ding, Y., Huang, J., and Xu, Y. (2020). Habitat characteristics and its effects on seedling abundance of *Hopea hainanensis*, a Wild Plant with Extremely Small Populations. *Biodivers. Sci.* 28, 289–295.
- Ly, V., Nanthavong, K., Pooma, R., Hoang, V. S., Khou, E., and Newman, M. F. (2018). *Hopea hainanensis*. *IUCN Red List Threat. Species* 2018:e.T32357A2816074.
- Mistry, J., Finn, R. D., Eddy, S. R., Bateman, A., and Punta, M. (2013). Challenges in homology search: HMMER3 and convergent evolution of coiled-coil regions. *Nucleic Acids Res.* 41, e121. doi: 10.1093/nar/gkt263
- Paz-Ares, J., Ghosal, D., Wienand, U., Peterson, P. A., and Saedler, H. (1987). The regulatory c1 locus of *Zea mays* encodes a protein with homology to myb proto-oncogene products and with structural similarities to transcriptional activators. *EMBO J.* 6, 3553–3558.
- Petroni, K., and Tonelli, C. (2011). Recent advances on the regulation of anthocyanin synthesis in reproductive organs. *Plant Sci.* 181, 219–229. doi: 10.1016/j.plantsci.2011.05.009
- Rahnasto-Rilla, M., Tyni, J., Huovinen, M., Jarho, E., Kulikowicz, T., Ravichandran, S., et al. (2018). Natural polyphenols as sirtuin 6 modulators. *Sci. Rep.* 8:4163. doi: 10.1038/s41598-018-22388-5
- Santos-Buelga, C., Mateus, N., and De Freitas, V. (2014). Anthocyanins. Plant pigments and beyond. *J. Agric. Food Chem.* 62, 6879–6884. doi: 10.1021/jf501950s
- Schwechheimer, C., and Bevan, M. (1998). The regulation of transcription factor activity in plants. *Trends Plant Sci.* 3, 378–383. doi: 10.1016/S1360-1385(98)01302-8
- Varet, H., Brillet-Gueguen, L., Coppee, J. Y., and Dillies, M. A. (2016). SARTools: a DESeq2- and EdgeR-Based R Pipeline for Comprehensive Differential Analysis of RNA-Seq Data. *PLoS One* 11:e0157022. doi: 10.1371/journal.pone.0157022
- Walker, A. R., Lee, E., Bogs, J., McDavid, D. A. J., Thomas, M. R., and Robinson, S. P. (2007). White grapes arose through the mutation of two similar and adjacent regulatory genes. *Plant J.* 49, 772–785. doi: 10.1111/j.1365-313X.2006.02997.x
- Waterland, N., Moon, Y., Tou, J., Kopsell, D., Kim, M. J., and Park, S. (2019). Differences in Leaf Color and Stage of Development at Harvest Influenced Phytochemical Content in Three Cultivars of Kale (*Brassica oleracea* L. and *B. napus*). *J. Agric. Sci.* 11:14.

- Winkel-Shirley, B. (2001). Flavonoid biosynthesis. A colorful model for genetics, biochemistry, cell biology, and biotechnology. *Plant Physiol.* 126, 485–493. doi: 10.1104/pp.126.2.485
- Xiaoying, W., Fangfang, H., Xianhua, G., Weiqiang, Z., Yuhui, H., Xiuyu, X., et al. (2017). Introduction Performance of *Hopea hainanensis* and *Vatica mangachapoi* in Guangdong Tree Park. *Guangdong For. Sci. Technol.* 33, 52–56.
- 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
- Yi, D., Zhang, H., Lai, B., Liu, L., Pan, X., Ma, Z., et al. (2021). Integrative Analysis of the Coloring Mechanism of Red Longan Pericarp through Metabolome and Transcriptome Analyses. *J. Agric. Food Chem.* 69, 1806–1815. doi: 10.1021/acs.jafc.0c05023
- Zheng, Y., Jiao, C., Sun, H., Rosli, H. G., Pombo, M. A., Zhang, P., et al. (2016). iTAK: a Program for Genome-wide Prediction and Classification of Plant Transcription Factors, Transcriptional Regulators, and Protein Kinases. *Mol. Plant* 9, 1667–1670. doi: 10.1016/j.molp.2016.09.014
- Zhuang, H., Lou, Q., Liu, H., Han, H., Wang, Q., Tang, Z., et al. (2019). Differential Regulation of Anthocyanins in Green and Purple Turnips Revealed by Combined De Novo Transcriptome and Metabolome Analysis. *Int. J. Mol. Sci.* 20:4387. doi: 10.3390/ijms20184387
- 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, Liao, Han, Zhou, Liang, Li, Yang, Tembrock, Wang and Wu. 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.