INTERACTIONS BETWEEN BIOCHEMICAL PATHWAYS PRODUCING PLANT COLORS AND SCENTS

EDITED BY: Xiumin Fu, Renata Rivera-Madrid and Lourdes Gómez-Gómez

PUBLISHED IN: Frontiers in Plant Science

frontiers Research Topics

frontiers

ISSN 1664-8714
ISBN 978-2-88976-582-9
DOI 10.3389/978-2-88976-582-9

About Frontiers

Frontiers is more than just an open-access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

Frontiers Journal Series

The Frontiers Journal Series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing. All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the Frontiers Journal Series operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

Dedication to Quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews.
Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view. By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the Frontiers Journals Series: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area! Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers Editorial Office: frontiersin.org/about/contact

INTERACTIONS BETWEEN BIOCHEMICAL PATHWAYS PRODUCING PLANT COLORS AND SCENTS

Topic Editors:
Xiumin Fu, South China Botanical Garden, Chinese Academy of Sciences (CAS), China
Renata Rivera-Madrid, Scientific Research Center of Yucatán (CICY), Mexico
Lourdes Gómez-Gómez, University of Castilla-La Mancha, Spain

Citation: Fu, X., Rivera-Madrid, R., Gómez-Gómez, L., eds. (2022). Interactions Between Biochemical Pathways Producing Plant Colors and Scents. Lausanne: Frontiers Media SA. doi: 10.3389/978-2-88976-582-9

Table of Contents

frontiers | Frontiers in Plant Science

EDITORIAL
published: 22 June 2022
doi: 10.3389/fpls.2022.955431

Editorial: Interactions Between Biochemical Pathways Producing Plant Colors and Scents

Xiumin Fu[1]*, *Lourdes Gómez-Gómez*[2]* and *Renata Rivera-Madrid*[3]*

[1] Guangdong Provincial Key Laboratory of Applied Botany and Key Laboratory of South China Agricultural Plant Molecular Analysis and Genetic Improvement, South China Botanical Garden, Chinese Academy of Sciences, Guangzhou, China, [2] Departamento de Ciencia y Tecnología Agroforestal y Genética, Instituto Botánico, Facultad de Farmacia, Universidad de Castilla-La Mancha, Albacete, Spain, [3] Centro de Investigación Científica de Yucatán A.C. (CICY), Mérida, Mexico

Keywords: anthocyanins, carotenoids, pigments, secondary metabolism, tissues, volatiles

Editorial on the Research Topic

Interactions Between Biochemical Pathways Producing Plant Colors and Scents

INTRODUCTION

The pigments and volatiles metabolites in plants fulfill the function of helping with pollination and spreading seeds (Sosenski and Parra-Tabla, 2019). These compounds also confer economic value on some fruits, vegetables, and flowers by giving them the important quality factors of color and aroma. In addition, these compounds are particularly interesting from the point of view of human health due to their properties as antioxidants, anti-inflammatory, or anticancer activities (Ávila-Román et al., 2021). The pigments in plants mainly comprise carotenoids, anthocyanins, chlorophyll, and betaines (Tanaka et al., 2008; Sudhakar et al., 2016). Volatile compounds are always divided into categories of terpenoids, phenylpropanoids/benzenoids, and fatty acid derivatives by their synthesis pathway (Maffei, 2010).

It has been widely reported that the composition or content of pigments and volatile compounds changes during the ripening of fruits and the development of flowers. Previous studies mainly focused on single pathway. Now with the development of omics technologies, many studies began to study the correlation between the changes of a metabolite in particular with respect to those associated with global changes of the other metabolites as a whole by combining metabolome, transcriptome, and genome analysis. The articles in this Research Topic are grouped into two themes as follows:

a) Application of omics technologies in the study of biosynthesis and regulation of pigments.
b) The correlation between pigments, volatile or other compounds.

APPLICATION OF OMICS TECHNOLOGIES IN THE STUDY OF BIOSYNTHESIS AND REGULATION OF PIGMENTS

Carotenoids are biosynthesized mainly *via* the MEP pathway in plants. Most of the genes in the carotenoid biosynthetic pathway have been identified in plants (Ruiz-Sola and Rodríguez-Concepción, 2012). Carotenoid isomerase (CRTISO) is a key enzyme to catalyze pro-lycopene to lycopene in the carotenoid biosynthetic pathway (Pinheiro et al., 2019). Jiang et al. identified the BoaCRTISO function in Chinese Kale. They found this gene could affect the color of the leaf by regulating the carotenoid and chlorophyll biosynthetic gene transcripts. Apart from

OPEN ACCESS

Edited and reviewed by:
Liang Guo,
Huazhong Agricultural University, China

***Correspondence:**
Xiumin Fu
Fuxiumin@scbg.ac.cn
Lourdes Gómez-Gómez
MariaLourdes.Gomez@uclm.es
Renata Rivera-Madrid
renata@cicy.mx

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

Received: *28 May 2022*
Accepted: *09 June 2022*
Published: *22 June 2022*

Citation:
Fu X, Gómez-Gómez L and Rivera-Madrid R (2022) Editorial: Interactions Between Biochemical Pathways Producing Plant Colors and Scents. Front. Plant Sci. 13:955431. doi: 10.3389/fpls.2022.955431

the biosynthesis genes in the carotenoid pathway, genes related to the formation of carotenoid sequestration structures and the plastids biogenesis also play an important role in carotenoid metabolism in crops (Lu et al., 2006; Lopez et al., 2008). The lower proportions of esterified xanthophylls were caused by the fewer expression levels of xanthophyll esterase (XES), resulting in the pale-yellow flower of petunia (Kishimoto et al., 2019).

A candidate gene *BrWF3*, homologous to *AtPES2*, was found possibly participate in xanthophyll esterification in Chinese cabbage flower (Yang et al.). The SNP deletion of *BrWF3* in the white petals of Chinese cabbage caused the protein to non-function. Therefore, the authors deduced that the lost function of BrWF3 results in the failure combination of carotenoid and polar lipids in the white petals of Chinese cabbage.

The anthocyanin biosynthesis pathway has been reported since 1980 (Holton and Cornish, 1995), and the functions of many genes in the pathway have been identified in recent years. The metabolic pathways and transcriptional regulation of anthocyanins have been intensively studied in model plants. Numerous species in nature could accumulate anthocyanin and may show different levels or types of metabolites. Based on the recent application of omics technologies, we can better understand their metabolites and their regulatory mechanisms. Anthocyanins contribute to *Lycoris* flowers' color. Wang N. et al. identified a hub gene *LrDFR1* through comparative transcriptome analysis, which possibly plays an important role in the anthocyanin accumulation in the *Lycoris* flower. Their experiment results suggested that *LrDFR1* acts as a positive regulator in anthocyanin biosynthesis. Fu et al. carried out integrated transcriptomics and metabolomics analysis in five *C. japonica* cultivars with different color petals and they identified *CjANS* and *Cj4CL* as key contributors to the diversity of petal color of *C. japonica*. Except for structural genes, the transcription factors, such as MYB, bHLH, WD, and MADS-box, play important roles in the regulation of anthocyanin biosynthesis (Zhang et al., 2014). Zhang et al. identified genes responsible for anthocyanin accumulation in the wild blueberry of Wufanshu (*V. bracteatum* Thunb.) by analysis of transcriptomes, and they identified VbMYBA as the transcriptional activator responsible for anthocyanin accumulation. In addition, they found that VbMYBA could activate strong anthocyanin accumulation in tobacco leaves. Li et al. gave a new sight of anthocyanin biosynthesis in peanuts by interaction joint analysis of multi-omics and miRNA. They suggested that *HCT* and chalcone biosynthesis-related candidate genes (*Ah21440*, *AhCHS*, and *AhCHI*) were the target genes of *AhmiR2950*, *AhmiR398*, *AhmiR50*, and *AhmiR51*.

Xiao et al. found that pigment-related compounds could be used as biomarker metabolites for genetic breeding screening. They compared the metabolomes of flower petals of *Nicotiana tabacum* L. (red) and *Nicotiana rustica* L. (orange) species using UPLC-ESI-MS/MS. A batch of novel metabolic biomarkers including flavonoids (kaempferol-3-o-rut, quercetin-glu, and rutin) and carotenoids (lutein and β-carotene) for these species were identified.

THE CORRELATION BETWEEN PIGMENTS, VOLATILE OR OTHER COMPOUNDS

Carotenoids could be oxidized by enzymatic activities (CCD, LOX) or non-enzymatically to generate apocarotenoids in plants, including crocins and bixin, which contribute to pigment development, and β-ionone, α-ionone, β-cyclocitral, β-damascenone, 6-methyl-5-hepten-2-one, and safranal, among others, that contribute to fragrance. The review of Zheng et al. summarized the underlying mechanism of regulation, metabolism, and genetics of apocarotenoid diversity, particularly focusing on apocarotenoid pigments and volatiles. In addition, they proposed a strategy that combines chromatography-mass spectrometry apocarotenoid profiling with multi-omics (such as QTL mapping, GWAS, and RNA-seq) to find new QTLs for apocarotenoid metabolite traits and identify their underlying genes.

Bixin is the second most economically important natural apocarotenoid. Bixin has an orange-red color and is produced mainly from *Bixa orellana* seeds. The bixin metabolic pathway was proposed as early as 20 years ago by Bouvier et al. (2003), but it has not been deeply studied at the biochemical and molecular levels. Us-Camas et al. identified two new genes *BoCCD1* and *BoCCD4* that participate in the biosynthesis of bixin aldehyde, which is the first product of bixin biosynthesis of *Bixa orellana* by using an *in vivo* (*E. coli*) and *in vitro* approach. They analyzed the generated products by LC-ESI-QTOF-MS/MS.

Xi et al. compared the carotenoids and aroma volatile apocarotenoids profiles of fruits of three different colored apricot cultivars. They found the total carotenoid contents were negatively correlated with the transcripts of *CCD1* and *CCD4* genes, while the aroma volatile apocarotenoid contents (mainly β-ionone) were positively regulated. These results suggested that CCD1 and CCD4 may be the key regulatory points for the profiles of color and aroma in apricot fruits. Therefore, these two genes can potentially be used as the targets for molecular breeding.

β-ionone is also a significant contributor to the fragrance of sweet osmanthus flowers and is mainly determined by the CCD4 enzyme (Han et al., 2019). Yan et al. performed a comprehensive analysis of the MYB-related transcription factor superfamily in sweet osmanthus and provided novel insights into the roles of OfMYB-related genes in sweet osmanthus as regulators of volatile compounds. They found that OfMYB1R114 and OfMYB1R70 members accelerated β-ionone formation. While OfMYB1R201 was involved in decreasing the β-ionone content. The mechanism of how these candidate MYB-related genes involved in the regulation of aroma substances remains to be further studied.

The metabolism of carotenoids has been studied intensively for decades, while the interaction mechanism between carotenoids and norisoprenoids (β-ionone) under abiotic stress remained unclear in many plants. Liu et al. found that carotenoid metabolism in peach fruit was significantly influenced by expression levels of carotenoid pathway genes (*PpPSY*, *PpLCY-B*, *PpLCY-E*, P*pCHY-B,* and *PpCCD4*) under UV-B irradiation. They deduced that the increased β-carotene and the decreased volatiles β-ionone were partially caused by

the inhibition of *PpCCD4* expression level under the UV-B irradiation. These results suggested that some stress factors could affect the carotenoid metabolism and then influence or change the volatile patterns.

Anthocyanins and a variety of aroma compounds (benzaldehyde, phenylacetaldehyde, and methyl salicylate) are derived from phenylalanine *via* the shikimate pathway. However, studies on the underlying mechanism of the relationship between anthocyanins and phenylalanine-derived volatiles are few. Mei, Wan et al. reported a specific tea (*Camellia sinensis*) variety with purple flowers, which accumulate a high concentration of anthocyanins. Meanwhile, tea flowers contain special volatile benzenoid-phenylpropanoids (BPs), such as 1-phenylethanol (1-PE) and acetophenone (AP). What would happen to the volatile compounds when the flower color mutates naturally from white color to purple color? According to the results, they found the flux to the benzenoid-phenylpropanoids (BPs) was also enhanced along with the anthocyanins accumulation in the flower of the tea plant.

In addition to the volatiles generated from the degradation of certain metabolites in plants, such as carotenoids or phenylalanine-derived compounds, volatiles contributing to the aroma of plants are as well directly synthesized, as the case of monoterpenes (linalool, geraniol, and their derivatives), which are derived mainly through the MEP pathway. Huang et al. identified monoterpenes as the main scent components in *Phalaenopsis bellina* flower, and they found that TPS-b and TPS-e/f enzymes are involved in the monoterpene biosynthesis in the P. *bellina* floral scent. Interestingly, the P. *bellina* flower accumulates anthocyanins, but the main aroma released is not phenylalanine-derived volatiles; while this flower could release terpenoid aroma from the MEP pathway, but could not accumulate carotenoids. Therefore, the regulation of different derived volatile compounds is very complex in plants. Ke et al. identified HcMYB1 could activate the structural gene *HcBSMT2* involved in methyl benzoate biosynthesis, and HcMYB2 can also activate the structural gene *HcTPS5* involved in linalool biosynthesis in *H. coronarium* flowers. These findings have shed light on the regulation of volatile compounds in plants.

The changes in plant color could affect not only volatiles but also other metabolism compounds. Mei, Lin et al. compared metabolites in fresh and fermented tea (*Camellia sinensis*) leaves between "Yinghong 9" (green leaves) and "Huangyu" (mutant yellow leaves) cultivars by using targeted metabolomics. Apart from the significant difference in pigments between these two cultivars, they also found alterations in polyphenols and volatiles. Proanthocyanidins are the main pigment substances in brown cotton, and studies have shown that fiber color is negatively correlated with fiber yield and quality. But the underlying mechanism between proanthocyanidins biosynthesis and metabolism in cotton fiber is unclear. Wang Z. et al. investigated the key structure and regulatory genes in the proanthocyanidins biosynthesis of brown cotton by combing with transcriptome co-expression network and metabolome analysis, and thus established the transcriptional regulatory network of proanthocyanidins biosynthesis and flavonoid metabolism in cotton.

Taken together, the Frontiers Research Topic presented here documents recent advances in pigments and volatiles biology research. In the present volume, the authors use advanced omics technologies to elucidate the biosynthesis and regulation of pigments and volatile and address the interactions between pigments and their related secondary compounds in plants. These results would give insight into the interaction between pigments and volatile and inspire further advances in the study of plant metabolites interaction.

AUTHOR CONTRIBUTIONS

All authors contributed equally to the manuscript and approved it for publication.

FUNDING

This work was supported by the project from the Young Scientists Fund of the National Natural Science Foundation of China (Grant No. 31902074), the Youth Innovation Promotion Association of Chinese Academy of Sciences (2022351), and the CAS Scholarship. RR-M work was supported by the Consejo Nacional de Ciencia y Tecnología (CONACYT) (Fronteras de la Ciencia No. 2016-01-1716).

ACKNOWLEDGMENTS

We thank all the authors for their contributions to this Research Topic. We also thank reviewers for their time and valuable comments and the Frontiers Editorial Team for their constant support.

REFERENCES

Ávila-Román, J., García-Gil, S., Rodríguez-Luna, A., Motilva, V., and Talero, E. (2021). Anti-inflammatory and anticancer effects of microalgal carotenoids. *Mar. Drugs* 19:100531. doi: 10.3390/md19100531

Bouvier, F., Dogbo, O., and Camara, B. (2003). Biosynthesis of the food and cosmetic plant pigment bixin (Annatto). *Science* 300, 2089–2091. doi: 10.1126/science.1085162

Han, Y., Wang, H., Wang, X., Li, K., Dong, M., Li, Y., et al. (2019). Mechanism of floral scent production in *Osmanthus fragrans* and the production and regulation of its key floral constituents, β-ionone and linalool. *Horticult. Res.* 6:4. doi: 10.1038/s41438-019-0189-4

Holton, T. A., and Cornish, E. C. (1995). Genetics and biochemistry of anthocyanin biosynthesis. *Plant Cell* 7, 1071–1083. doi: 10.2307/3870058

Kishimoto, S., Oda-Yamamizo, C., and Ohmiya, A. (2019). Comparison of petunia and calibrachoa in carotenoid pigmentation of corollas. *Breedi. Sci.* 69, 117–126. doi: 10.1270/jsbbs.18130

Lopez, A. B., Van Eck, J., Conlin, B. J., Paolillo, D. J., O'Neill, J., and Li, L. (2008). Effect of the cauliflower *Or* transgene on carotenoid accumulation and chromoplast formation in transgenic potato tubers. *J. Exp. Bot.* 59, 213–223. doi: 10.1093/jxb/erm299

Lu, S., Van Eck, J., Zhou, X., Lopez, A. B., O'Halloran, D. M., Cosman, K. M., et al. (2006). The cauliflower *Or* gene encodes a DnaJ cysteine-rich domain-containing protein that mediates high levels of β-carotene accumulation. *Plant Cell* 18, 3594–3605. doi: 10.1105/tpc.106.046417

Maffei, M. E. (2010). Sites of synthesis, biochemistry and functional role of plant volatiles. *South Afri. J. Bot.* 76, 612–631. doi: 10.1016/j.sajb.2010.03.003

Pinheiro, T. T., Peres, L. E. P., Purgatto, E., Latado, R. R., Maniero, R. A., Martins, M. M., et al. (2019). Citrus carotenoid isomerase gene characterization by complementation of the "Micro-Tom" tangerine mutant. *Plant Cell Rep.* 38, 623–636. doi: 10.1007/s00299-019-02393-2

Ruiz-Sola, M. Á., and Rodríguez-Concepción, M. (2012). Carotenoid biosynthesis in Arabidopsis: a colorful pathway. *Arabidopsis Book* 10:e0158. doi: 10.1199/tab.0158

Sosenski, P., and Parra-Tabla, V. (2019). *Secondary Metabolites: Attracting Pollinators*. p. 1–9. doi: 10.1002/9780470015902.a0000909.pub2

Sudhakar, P., Latha, P., and Reddy, P. V. (2016). "Chapter 15 - plant pigments," in *Phenotyping Crop Plants for Physiological and Biochemical Traits*, eds P. Sudhakar, P. Latha and P. V. Reddy. (Cambridge, MA: Academic Press), 121–127. doi: 10.1016/B978-0-12-804073-7.00015-6

Tanaka, Y., Sasaki, N., and Ohmiya, A. (2008). Biosynthesis of plant pigments: anthocyanins, betalains and carotenoids. *Plant J.* 54, 733–749. doi: 10.1111/j.1365-313X.2008.03447.x

Zhang, Y., Butelli, E., and Martin, C. (2014). Engineering anthocyanin biosynthesis in plants. *Curr. Opin. Plant Biol.* 19, 81–90. doi: 10.1016/j.pbi.2014.05.011

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.

frontiers in Plant Science

ORIGINAL RESEARCH
published: 18 December 2020
doi: 10.3389/fpls.2020.607715

The Genes of *CYP*, *ZEP*, and *CCD1/4* Play an Important Role in Controlling Carotenoid and Aroma Volatile Apocarotenoid Accumulation of Apricot Fruit

*Wanpeng Xi[1,2], Lina Zhang[2], Shengyu Liu[2] and Guohua Zhao[1]**

[1] College of Food Science, Southwest University, Chongqing, China, [2] College of Horticulture and Landscape Architecture, Southwest University, Chongqing, China

OPEN ACCESS

Edited by:
Xiumin Fu,
South China Botanical Garden (CAS), China

Reviewed by:
Huping Zhang,
Nanjing Agricultural University, China
Tianhu Sun,
Cornell University, United States
Ai-Sheng Xiong,
Nanjing Agricultural University, China
Chunhua Zhou,
Yangzhou University, China

****Correspondence:***
Guohua Zhao
zhaogh@swu.edu.cn

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

Received: *18 September 2020*
Accepted: *02 December 2020*
Published: *18 December 2020*

Citation:
Xi W, Zhang L, Liu S and Zhao G (2020) The Genes of CYP, ZEP, and CCD1/4 Play an Important Role in Controlling Carotenoid and Aroma Volatile Apocarotenoid Accumulation of Apricot Fruit. Front. Plant Sci. 11:607715. doi: 10.3389/fpls.2020.607715

Carotenoids are important coloration molecules and indispensable component of the human diet. And these compounds confer most of the apricot fruit yellow or orange color. In China, fruit of some apricot cultivar present light-yellow color but strong flowery flavor, however, the chemical mechanism remains unknown. Here, carotenoids and aroma volatile apocarotenoids (AVAs) in three skin types of apricot cultivars (orange, yellow, and light-yellow skinned) were determined by HPLC and GC-MS, respectively. And the transcript levels of carotenogenic genes were analyzed by qRT-PCR. The orange-skinned cultivars "Hongyu" and "Danxing" fruit presented the most abundant total carotenoid, β-carotene and specific α-carotene contents, and β-carotene (52–77%) increased to become the dominant carotenoid during fruit ripening. The transcript levels of lycopene β-cyclase (*LCYb*) and β-carotene hydroxylase (*CHYb*) sharply increased during ripening. The yellow-skinned cultivars "Sulian No. 2" and "Akeyaleke" fruit contained lower levels of total carotenoids and β-carotene but were rich in phytoene. The light-yellow coloration of "Baixing" and "Luntaixiaobaixing" fruit was attributed to low amounts of total carotenoids, lutein, and neoxanthin and an absence of β-cryptoxanthin, but high level of aroma volatile apocarotenoids (AVAs) such as β-ionone were detected in these cultivars fruit, accompanied by low transcript levels of carotene hydroxylase (*CYP*) and zeaxanthin epoxidase (*ZEP*) but high levels of carotenoid cleavage dioxygenase 1 (*CCD1*) and *CCD4*. Correlation analysis showed that the expression level of *CCD1* negatively correlated with carotenoid accumulation but positively with AVAs production. These collected results suggest that both carotenoid biosynthesis and degradation are important for apricot coloration and aroma formation. CYP, ZEP, CCD1, and CCD4 may be the key regulation points for carotenoid and AVAs accumulation in apricot fruit, which provide important targets for quality-oriented molecular breeding.

Keywords: apricot, carotenoid, apocarotenoid, color, aroma, ripening

INTRODUCTION

Carotenoids are pigments that play a major role in the protection of plants against photooxidative processes. They are efficient antioxidants that scavenge singlet molecular oxygen and peroxyl radicals. In humans, carotenoids are part of the antioxidant defense system (Spirt et al., 2010) and include the vitamin A precursor β-carotene, which can enhance immunity and reduce the incidence of various diseases, such as cancer and cardiovascular disease (Fraser and Bramley, 2004). In addition to protecting plants and improving human health, carotenoids confer an attractive visual appearance of plants. The accumulation of carotenoids results in red, yellow, and orange color in plants (Tanaka et al., 2008). Therefore, carotenoids represent a double index of plant appearance and an indispensable component of the human diet (Saini et al., 2015). Based the point, increasing attention is being paid to the biosynthesis and metabolism of carotenoids in horticultural crops (Enfissi et al., 2017).

Apricot (*Prunus armeniaca* L.) is a very important fruit tree within Rosaceae, and its fruits are appreciated by consumers for their characteristic color, flavor, and juiciness, which are strongly related to the variety and ripening process (Botondi et al., 2003); these fruits are also important sources of fiber, provitamin A, vitamins, and phenolic compounds (Erdogan-Orhan and Kartal, 2011; Melgarejo et al., 2014). As one of the world's sources of apricot, China has rich apricot germplasm resources, and many cultivars with diverse quality traits exist (Zhang et al., 2019). These apricots present colorful appearance that ranges from light-yellow to yellow or orange, and the fruits of some cultivars exhibit a red blush against the yellow or orange background. These different resource types are ideal natural materials for studying the mechanisms of coloration and colorant metabolism in fruit.

In recent years, carotenoid metabolism has been investigated in many horticultural crops (Yuan et al., 2015), and the biosynthesis pathways have been well established. Carotenoids are biosynthesized in plastids and stored in plant cells. The biosynthesis process starts with the condensation of two geranylgeranyl pyrophosphate (GGPP) molecules by phytoene synthase (PSY) to form colorless phytoene as a 15-*cis* isomer, after which phytoene is converted into red lycopene via a series of desaturation and isomerization steps. Next, all-*trans*-lycopene is cyclized to form orange carotenoid molecules and ultimately hydroxylated to yield yellow xanthophyll (Nisar et al., 2015). Carotenoids can be oxidized to apocarotenoids through symmetrical 9–10 and 9′–10′ cleavage by carotenoid cleavage dioxygenases (CCDs) (McQuinn et al., 2015). The carotenoids existing in many fruits have been characterized, and the profile depends on many factors, such as environmental conditions, the cultivar and the fruit ripening process, and many studies have suggested that the dynamic profile of carotenoids during ripening serves as a basis for understanding their molecular mechanism (Luan et al., 2020).

Previous studies have shown that the yellow or orange background skin coloration of apricots is determined by carotenoids (Le Bourvellec et al., 2018) and that carotenoid accumulation in apricots is regulated by ethylene (Marty et al., 2005). However, the molecular basis of carotenoid accumulation during apricot fruit development has yet to be reported, and the differences between various cultivars remains elusive. In this study, we compared the profile of carotenoid and aroma volatile apocarotenoids in three skin types of apricot cultivar peels and flesh and investigated the pattern of 11 genes involved in carotenogenesis during fruit development and ripening. Our results reveal new insight into apricot coloration and aroma formation, and provide important information for understanding carotenoid metabolism and regulation in apricot fruit.

MATERIALS AND METHODS

Fruit Materials

Apricots were collected from the orchard of the National Fruit Tree Germplasm Repository, Academy of Xinjiang Agricultural Sciences, Luntai, Xinjiang, China (45″191 N, 86″031 E). All experimental trees were planted at a 3–4 m spacing in 1998 in rows in a north-south orientation. The same tree shape (open-center shape), fertilization management, and pest control were used for all experimental trees. Experimental design was a singletree plot complete randomized design with 10 individual trees as replications for each cultivar. From April to July 2017, the fruits of six cultivars ("Hongyu," HY, with a red peel and orange flesh; "Danxing," DX, with a red peel and orange flesh; "Sulian No. 2," SL, with an orange-red peel and yellow flesh; "Akeyaleke," AK, with an orange-red peel and yellow flesh; "Baixing," BX, with a light-yellow peel and flesh; "Luntaixiaobaixing," LT, with a light-yellow peel and flesh) were picked at the fruitlet (F, 21 days after blossoming, DAB), enlargement (E, 32 DAB for "HY" and "DX," 27 DAB for "SL," "AK," "BX," and "LT"), turning (T, 61 DAB for "HY," "SL" and "BX," and 56 DAB, 58 DAB and 57 DAB for "DX," "AK," and "LT," respectively), commercial maturation (CM, 74 DAB for "HY," "DX," "SL" and "AK," and 88 DAB for "BX," 65 DAB for "LT"), and fully ripe (FR, 82 DAB for "HY" and "DX," 85 DAB for "SL" and "AK," 91 DAB for "BX," 74 DAB for "LT") stages. After being picked, the fruits were immediately transported to the laboratory, and fruits without mechanical damage or pests were used. Each replicate consisted of 50 fruits, among which 20 fruits were used to determine the basic fruit quality index, and the others were divided into peel and flesh tissues, which were cut into small cubes, immediately frozen in liquid nitrogen and stored at −80°C for later analysis. Three replicates were used for each sample.

Equations Determination of the Basic Fruit Quality Index

The firmness of the flesh was measured with a hardness tester (Model: HL-300, Xianlin Non Detection Device Co., Ltd., Nanjing, China) with an 8 mm probe. A handheld digital refractometer (B32T Brix Meter, Guangzhou Ruiqi Trading Co., Ltd., Guangdong, China) was used to determine the total soluble solid (TSS) content of the fruit. Titratable acids (TA) were determined by the NaOH titration method. The titratable acidity was calculated with the following formula: TA

[mmol/100(mL)] = [(c × V_1)/V_0] × [100/V] × 100, where c is the molar concentration of the NaOH standard solution, which was 0.1 mol/L; V_1 is the volume of NaOH solution consumed by the titration; V_0 is the volume of the sample used for titration, which was 30 mL; V is the volume of the juice, which was 10 mL; and 100 is the volume after the juice is diluted. The titratable acidity of fruit juice is often expressed as the acid percentage, calculated according to the following formula: TA (%) = [(c × V_1 × k)/V_0] × [100/V] × 100. The main organic acid in apricot is malic acid, and the conversion factor, k, of malic acid is 0.067. Fruit color at different developmental stages was measured with a Hunter Mini Scanning Colorimeter (Hunter Associates Laboratory, Inc., Reston, VA, United States), and the color index CCI (citrus color index) was calculated with the following formula: CCI = 1000 × a*/(L* × b*).

Carotenoid Extraction and Quantification

The extraction and quantification of carotenoids was carried out according to our previous study (Zheng et al., 2016). Twenty grams of flesh or eight grams of peel was dissolved in 50 mL of extraction solvent (hexane/acetone/ethanol, 50:25:25). After standing for 30 min, the mixtures were centrifuged for 5 min at 6500 rpm. The top colored layer of hexane was recovered and transferred to a volumetric flask. The hexane extract was blown dry with nitrogen and dissolved in 2 mL of methyl tert-butyl ether (MTBE). Then, the solution was transferred to 2 mL of 10% methanol/potassium hydroxide. The mixture was allowed to stand and separate in a separatory funnel, then rinsed twice with water and once with 0.1% butylhydroxytoluene (BHT)/MTBE. The rinse solution was transferred to a brown bottle and dried with nitrogen. After drying, 2 mL methanol/acetone (2:1) was added to the brown bottle for dissolution and filtering with a 0.22 m filter membrane, and the resultant sample was prepared for carotenoid determination.

Carotenoids were separated and detected by HPLC (Waters, Milford, MA, United States) with C30 chromatography columns (250 mm × 4.6 mm, 5 μm, YMC, Wilmington, NC, United States) and photodiode array detector. The mobile phase flow rate was 1 mL/min, the column temperature was 25°C, the detection wavelength was 288 nm, and the injection volume was 20 μL. The mobile phase consisted of methanol, MTBE, and water. Carotenoids were identified by comparison to standard retention times and UV-visible spectral peaks. The quantification of carotenoids was performed using the standard curve method, and the concentration of carotenoids was expressed as μg/g fresh weight (FW). All carotenoid standards including β-carotene, α-carotene, phytoene, β-cryptoxanthin, neoxanthin, lutein, and violaxanthin were purchased from Sigma-Aldrich. Three replicates were used for each sample.

Aroma Volatile Apocarotenoid Extraction and Quantification

The AVA contents in flesh were analyzed by gas chromatography mass spectrometry (GC-MS) as previously described (Zheng et al., 2016). A solid-phase microextraction (SPME) needle with a 1-cm long fiber coated with 65 μm of polydimethylsiloxane, and divinybenzene (Supelco Inc., Bellefonte, PA, United States) was used for volatile extraction. The identification and quantification of volatiles was performed on an Agilent 6890N GC equipped with a flame ionization detector (FID) detector and a DB-WAX column (0.32 mm, 30 m, 0.25 μm, J&W Scientific, Folsom, CA, United States). All volatiles were quantified according to standard curves of authentic compounds. β-Ionone, dihydro-β-ionone, β-damascenone, and 6-methyl-5-hepten-2-one were obtained from Sigma (St. Louis, MO, United States). Extracts from three triplicate tissue samples were analyzed.

RNA Extraction and cDNA Synthesis

Total RNA was isolated and extracted from fruit flesh using a Tiangen reagent kit as described previously (Shan et al., 2008). RNA quantity and quality were determined in a NanoDrop 2000 spectrophotometer and through denaturing agarose gel electrophoresis. The first strand of cDNA was synthesized from 4 μg of total RNA as a template according to the protocol of the SuperScript^{TM} III RT-PCR First-Strand Synthesis System (Invitrogen, CA, United States) and then used as a template to conduct real-time quantitative PCR.

Real-Time Quantitative PCR Analysis

The specific primers for the eleven genes involved in carotenoid biosynthesis (*PSY*, *PDS*, *ZDS*, *CRTISO*, *LCYb*, *CHYb*, *CYP*, *ZEP*, *NCED*, *CCD1*, and *CCD4*) and two reference genes (26 S ribosome and Actin) (**Supplementary Table 1**) were designed on the basis of our RNA-Seq data (NCBI Sequence Read Archive: PRJNA506502). Gene transcript levels were detected using an iQ5 instrument (Bio-Rad Laboratories, Inc., United States) with the SYBR Premix Ex Taq II Kit [TaKaRa Biotechnology (Dalian) Co., Ltd., China]. The amplification procedure was as follows: 95°C for 1 min, followed by 40 cycles at 95°C for 20 s, 58°C for 20 s, and 72°C for 30 s. The Ct values of the reactions were recorded, and gene transcript quantification was performed based on the relative expression of the selected genes against that of Actin using the $2^{-\Delta\Delta CT}$ method (Zhang et al., 2017). Three biological replicates were used for each analysis.

Statistical Analysis

Data were expressed as the mean of three biological replicates ± standard deviation (SD) using Microsoft Excel 2010. Statistical differences for each treatment point were tested by Fisher's protected least squares difference (LSD) test at a 0.01 probability. The correlation analysis was conducted by RStudio (Version 1.1.463). All graphs were drawn with Origin Pro 2018 (OriginLab Corporation, Northampton, MA, United States).

RESULTS

Evaluation of Fruit Ripening According to Basic Fruit Quality Parameters

The changes in basic fruit quality parameters for the six apricot cultivars of three skin types are shown in **Figure 1**. During fruit development and ripening, the fruit weight of all cultivars

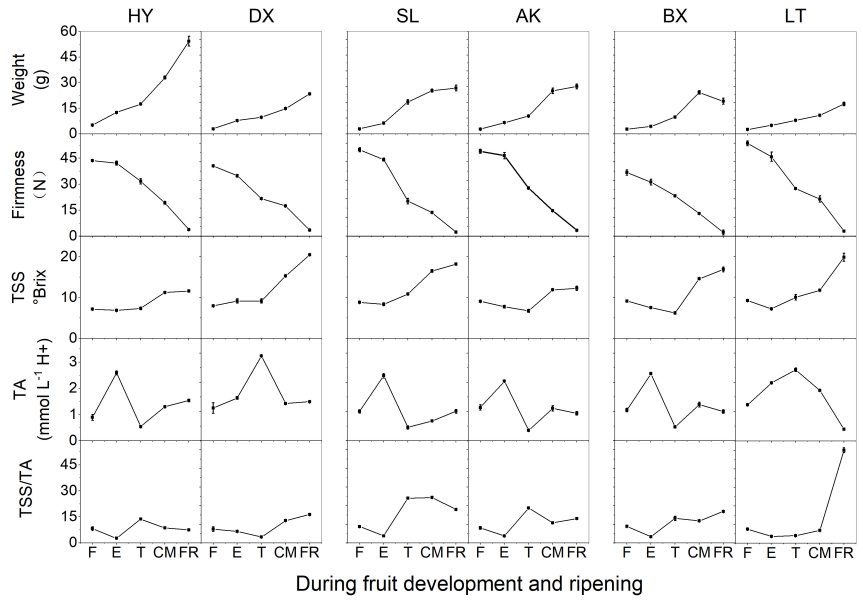


FIGURE 1 | Basic quality parameters of apricots during development and ripening. TA, titratable acid; TSS, total soluble solid; TSS/TA, the ratio of TSS and TA. F, E, T, CM, and FR represent fruitlet, enlargement, turning, commercial maturation, and fully ripe developmental stages of fruit, respectively. All data are expressed as the means ± standard deviations of three biological replicates.

increased significantly ($p < 0.01$). Although the TSS content remained stable from the F to E stages in all cultivars, it increased rapidly from the T to FR stages, and similar changes were observed for TSS/TA. Firmness decreased markedly throughout the development and ripening period ($p < 0.01$). The TA contents of the six cultivars increased to the maximum at the E or T stage and then decreased in the later stage of fruit development. The changes in fruit quality parameters suggested that the early stages, F and E, correspond to development, while the late stages, T, CM, and FR, correspond to ripening.

Color Parameters of Different Cultivars During Fruit Development and Ripening

The peel color of the fruits changed dramatically throughout the fruit development period ($p < 0.01$). According to the peel color at the fully ripe (FR) stage, the six apricot cultivars could be clearly divided into three skin color types: orange, yellow, and light-yellow (**Figure 2**). With the development of fruit, the values of L*, a*, b*, and c* continuously increased (**Supplementary Table 2**), indicating that the fruit peel color gradually changed from green to yellow. Notably, the hue values (h) of the six apricot cultivars at the F stage were basically the same, ranging from 114.91 ± 0.59 to 125.31 ± 6.38 (**Supplementary Table 2**). At the FR stage, the *h* values of the two orange cultivars, "HY" and "DX," decreased to 72.57 ± 0.40 and 67.36 ± 2.90, respectively, which were significantly lower than the *h* values of the two yellow cultivars, "SL" and "AK," and much lower than those of the two light-yellow cultivars, "BX" and "LT," respectively (**Supplementary Table 2**), which is completely consistent with the definition of the *h* value as representing reddish-purple at 0°, yellow at 90°, and bluish-green at 180°. In addition, the color differences between the cultivars were well described by the CCI values. At the FR stage, the six cultivars were divided into three types based on their CCI values, which was consistent with the peel color of each cultivar at the fully ripe stage (**Figure 2**).

Carotenoid Profile of Each Cultivar During Fruit Development and Ripening

In total, seven carotenoids were identified from the tested apricots, including β-carotene, α-carotene, phytoene, β-cryptoxanthin, neoxanthin, lutein, and violaxanthin. Among these carotenoids, β-carotene was predominant in the fruits of all apricot cultivars, followed by lutein. The total carotenoid content in the peel was markedly higher than that in the flesh despite fruit development (**Figure 3** and **Supplementary Figures 1, 2**).

The total carotenoid content continuously increased in the orange cultivars "HY" and "DX" throughout the development period, while it decreased during fruit ripening in the yellow ("SL" and "AK") and light-yellow ("BX" and "LT") cultivars (**Figure 3**). The contents of five of the detected carotenoids (β-carotene, α-carotene, phytoene, β-cryptoxanthin, and neoxanthin) generally increased in all six cultivars throughout the fruit development period, whereas those of the other two carotenoids (lutein and violaxanthin) continued to increase until the E or T stage and then decreased until the FR stage in the peel and continued to decrease throughout the remaining development stages in the flesh.

The highest total carotenoid content was observed in the orange cultivars, while the total carotenoid contents detected in the yellow and light-yellow cultivars were only 27–63% and 19–20%, respectively, of those in the orange cultivars. It was notable that the composition and content of carotenoids differed in the

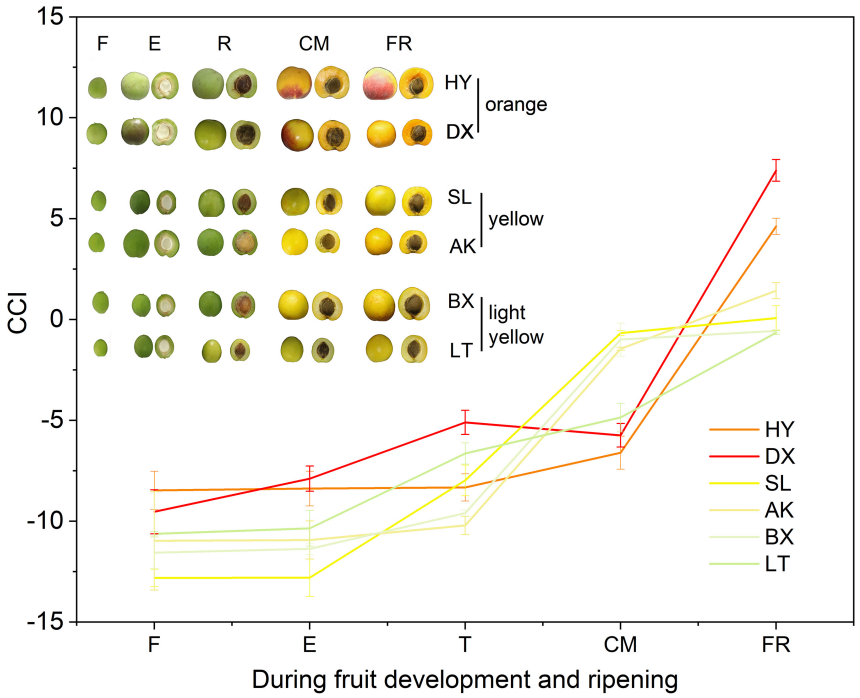


FIGURE 2 | Citrus color index (CCI) values of three types of cultivars during fruit development and ripening. F, E, T, CM, and FR represent fruitlet, enlargement, turning, commercial maturation, and fully ripe developmental stages of fruit. All data are expressed as the means ± standard deviations of three biological replicates.

cultivars. In the orange cultivars, only phytoene was not detected in "HY," while "DX" showed enrichment of all carotenoids identified. The contents of β-carotene, β-cryptoxanthin, and specific α-carotene in "HY" and "DX" were markedly higher than those in the yellow and light-yellow cultivars ($p < 0.01$). Conversely, although the content of phytoene increased in all cultivars throughout the development period, the lowest content was observed in the orange cultivars. In the yellow cultivars, α-carotene was not detected, and higher contents of neoxanthin and violaxanthin were detected than in the orange and light-yellow cultivars. The content of β-carotene was significantly lower than that in the orange cultivars ($p < 0.01$). In the light-yellow cultivars, α-carotene, and β-cryptoxanthin were not detected, and the content of β-carotene was lower than that in any of the other cultivars included in this study.

The proportions of the carotenoids varied significantly between fruits of different stages in each cultivar (**Figure 4**). Lutein was the predominant carotenoid in all cultivars in the early development period (from F to T), accounting for 52.7–89.7% of the total carotenoids in peels and 34.4–83% of the total carotenoids in flesh. In contrast, the proportion of lutein decreased dramatically during ripening (from T to FR), accounting for 8.3–44.3% of the total carotenoids in peels and 0–30.4% of the total carotenoids in flesh, and this carotenoid was replaced by increasing amounts of β-carotene and β-cryptoxanthin. During the ripening period, the proportion of β-carotene in flesh increased to 15.6–76.6% in the orange cultivars, 33–47.3% in the yellow cultivars, and 45.9–49.8% in the light-yellow cultivar "BX," but this carotenoid did not accumulate in "LT," and the proportion of β-cryptoxanthin in flesh increased to 19.4–53.9% in the orange cultivars and no detected (ND) to 38.8% in the yellow cultivars, but this carotenoid did not accumulate in either light-yellow cultivar. Moreover, at the FR stage, the carotenoids in the flesh of "HY" and "DX" were dominated by β-carotene, which is a natural pigment conferring an attractive orange color, accounting for 76.6% and 52% of the total carotenoids, respectively. The flesh of "SL" and "AK" was dominated by β-cryptoxanthin and β-carotene, respectively, accounting for 38.8% and 47.3% of the total carotenoids. The flesh of "BX" and "LT" was dominated by β-carotene and lutein, respectively, accounting for 45.9% and 82.2% of the total carotenoids.

Aroma Volatile Apocarotenoid Profile of Each Cultivar During Fruit Development and Ripening

Four aroma volatile apocarotenoids were identified from the fruit of cultivars tested, including β-ionone, dihydro-β-ionone, 6-methyl-hepten-2-one, and β-damascenone (**Figure 5**). β-ionone was the predominant component and its content increased rapidly from the T stage in all yellow and light-yellow cultivars, but it remained very low level in orange cultivar throughout ripening process. Similar changes dihydro-β-ionone, 6-methyl-hepten-2-one, and β-damascenone were observed these varieties during ripening. Even increased four aroma volatile apocarotenoids were also detected in yellow cultivars "SL" and "AK"; however, their abundance of these compounds in

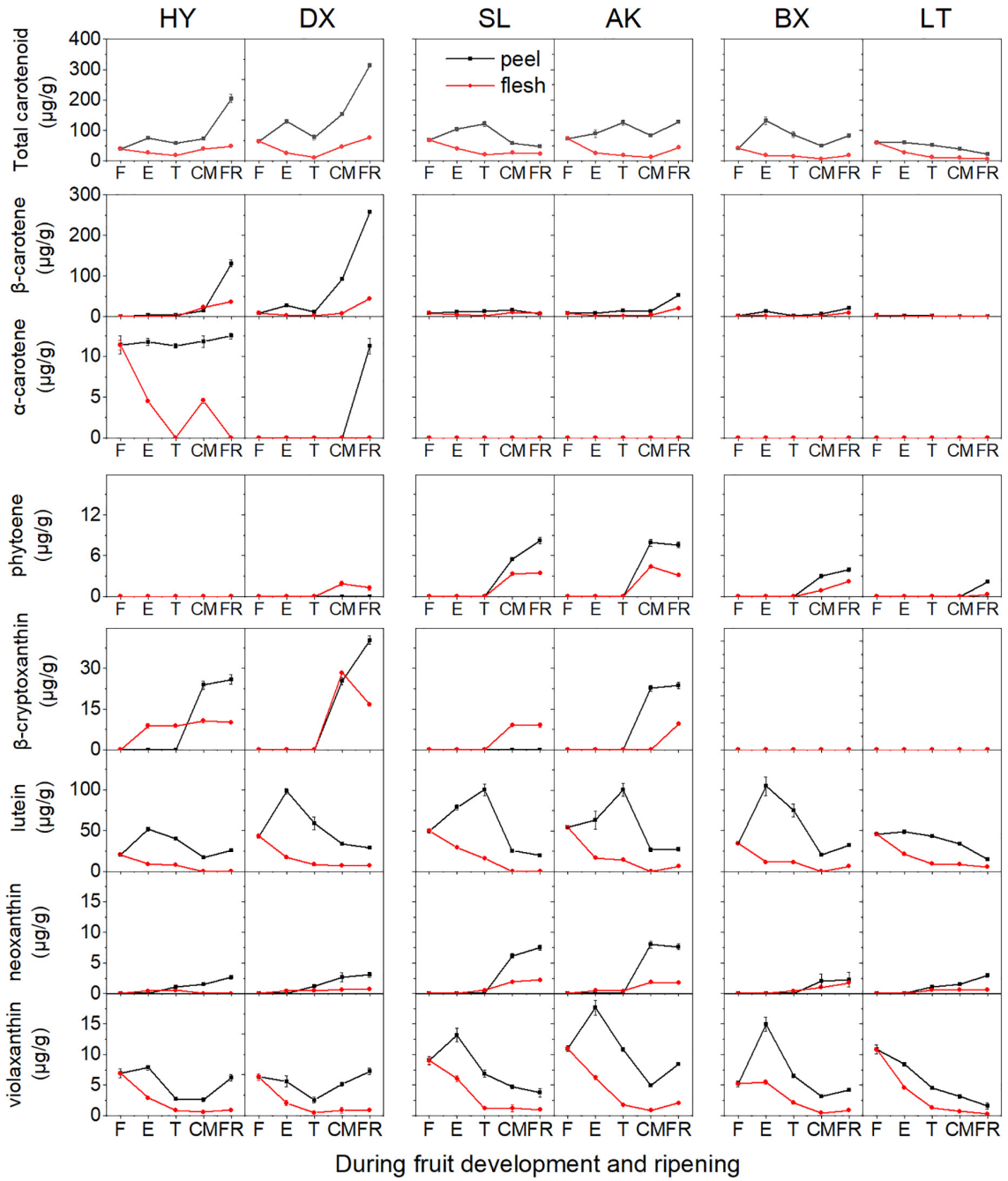


FIGURE 3 | Carotenoid content (μg/g fresh weight) of apricot fruit during development and ripening. F, E, T, CM, and FR represent fruitlet, enlargement, turning, commercial maturation, and fully ripe developmental stages of fruit. All data are expressed as the means ± standard deviations of three biological replicates.

light-yellow cultivars "BX" and "LT" were considerably higher than those in yellow cultivars during ripening.

Differential Transcription Levels of Genes Involved in Carotenoid Biosynthesis

The expression of 11 genes involved in carotenoid biosynthesis was analyzed during the entire developmental period (**Figure 6**). Among these genes, the transcript levels of six genes (*PSY*, *PDS*, *ZDS*, *CRTISO*, *LCYb*, and *CHYb*) continued to increase as the fruits matured and generally showed the highest transcript levels in the orange cultivars, which corresponded to the highest levels of phytoene, β-carotene, α-carotene, and β-cryptoxanthin found in orange flesh. It is worth noting that the expression of *CHYb* in the yellow cultivars was significantly increased during ripening ($p < 0.01$), which might have resulted in the simultaneous

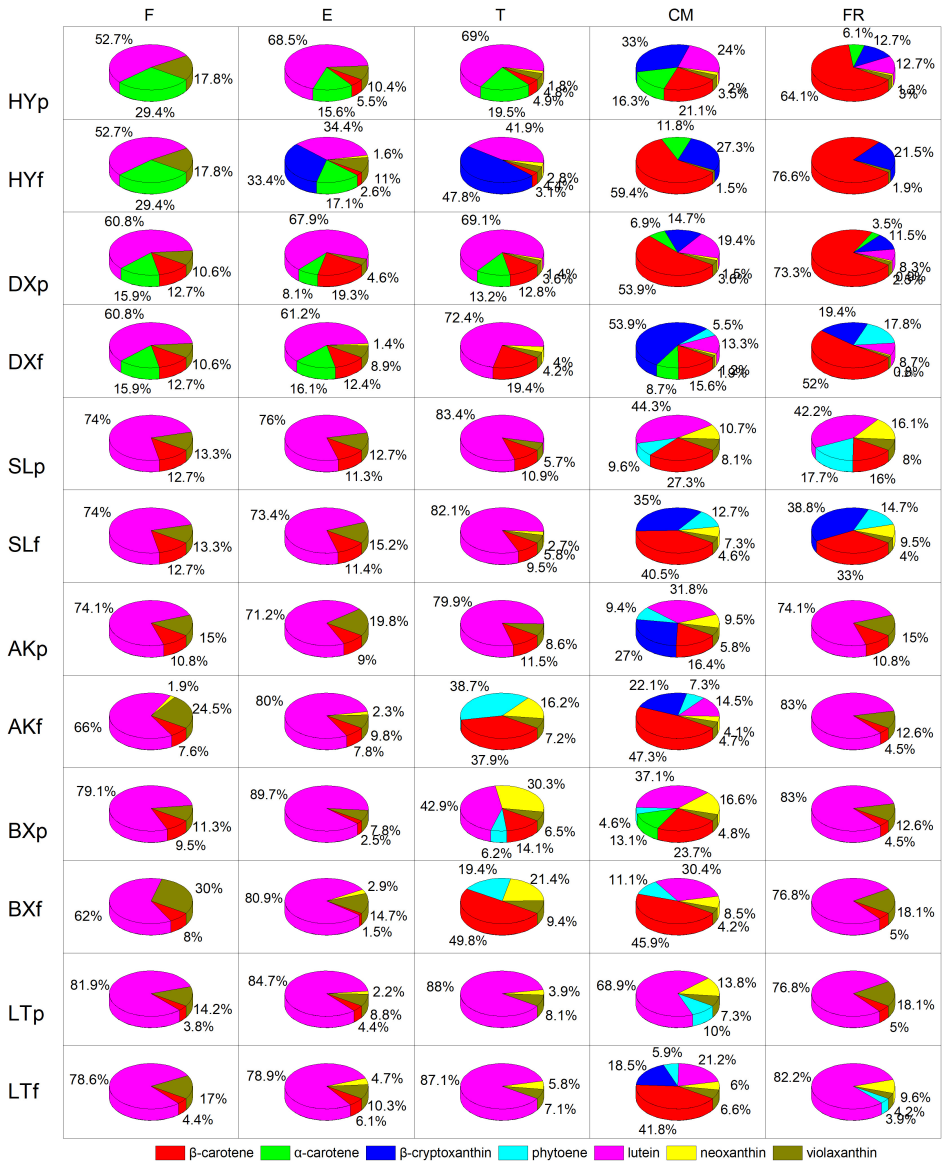


FIGURE 4 | Proportion of carotenoids during fruit development and ripening. p and f represent the peel and flesh, respectively. βca, β-carotene; αca, α-carotene; βcr, β-cryptoxanthin; phy, phytoene; lut, lutein; neo, neoxanthin; vio, violaxanthin. F, E, T, CM, and FR represent five developmental stages of fruit. All data are expressed as the means ± standard deviations of three biological replicate.

increase in the proportion of β-cryptoxanthin from ND to 38.8%. In addition, the expression of *CYP* continued to increase during the early development stages (from F to E) then decreased until the FR stage in all six cultivars, corresponding to the decline in lutein content. The transcript level of *ZEP* in the yellow and light-yellow cultivars was similar to that of *CYP*, which was consistent with the decrease in the violaxanthin content. Although the expression of *ZEP* increased in the orange cultivars, the content of violaxanthin decreased unexpectedly as the fruit matured, which might have been due to the increased transcript level of *NCED*, whose product converts violaxanthin into abscisic acid. In addition, the transcript levels of *NCED*, *CCD1*, and *CCD4* continued to increase throughout the fruit development period were highest in the light-yellow cultivars, which could result in more carotenoids in "BX" and "LT" being converted into apocarotenoids, leading to lower carotenoid levels.

Correlation Between Carotenoid Contents, Aroma Volatile Apocarotenoid Contents and Expression Levels of Carotenogenic Genes

Correlation analysis was used to investigate the relationship between carotenoid contents, aroma volatile apocarotenoid contents and expression levels of genes in carotenoid biosynthesis pathway (**Figure 7**). A high positive correlation between the

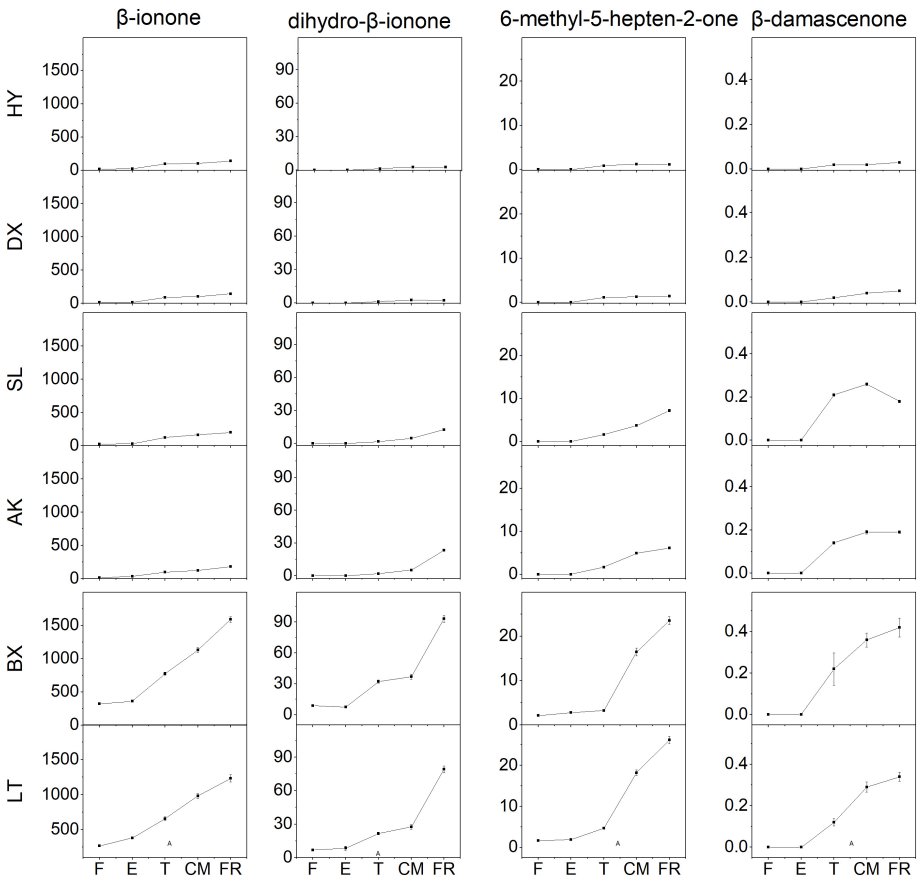


FIGURE 5 | Aroma volatile apocarotenoid content of apricot fruit development and ripening. F, E, T, CM, and FR represent fruitlet, enlargement, turning, commercial maturation, and fully ripe developmental stages of fruit. All data are expressed as the means ± standard deviations of three biological replicates.

total carotenoid, β-carotene, β-cryptoxanthin contents, and expression level of *ZEP* in all cultivars studied was observed ($p < 0.01$, $r > 0.71$), high correlation was also found between β-carotene content and expression level of *CHYb*. Significant negative correlation was found between carotenoids accumulation and expression levels of *CCD1* or *CCD4*. Significant positive correlation was observed between four AVAs (β-ionone, dihydro-β-ionone, 6-methyl-hepten-2-one, and β-damascenone) contents and the expressions of *CCD1* and *CCD4* ($p < 0.01$, r ranged 0.72–0.95), while no significant correlation was found between the expressions of *NCED* and AVAs content ($p < 0.01$, $r < 0.67$).

DISCUSSION

In flowering plants, the composition and accumulation of carotenoids cause many fruits to exhibit different colors, such as yellow, orange or red (Sagawa et al., 2016). An increasing number of studies have shown that the color differences in the fruits of different cultivars are closely related to the composition and content of carotenoids and the transcript levels of key genes in the carotenoid metabolism pathway (Qi et al., 2019). The specific reason for the color difference in apricot cultivars is unknown. In the present study, we first explored the different mechanisms of carotenoid accumulation in three skin types of apricot cultivars during fruit development and ripening.

Three Skin Types of Apricots Present Different Carotenoid and AVA Profile

Previous studies have shown that the color change of fruit is closely associated with the contents and proportions of carotenoids during development and ripening (Yuan et al., 2015). Here, the sharply contrasting peel colors of the three skin types of apricot cultivars indicated that there was a significant difference in the accumulation of carotenoids. The total carotenoid contents of the orange cultivars were much higher than those of the yellow and light-yellow cultivars, suggesting that the degree of fruit coloration may depend on the content of total carotenoids. The contents and proportions of β-carotene and β-cryptoxanthin increased significantly, and these compounds became the main pigments during the ripening of the orange cultivars, which explains why these cultivars show a vivid orange color when their fruits are fully ripe; these results are consistent with findings for the same color type of tomato and carrot (Rodriguez-Concepcion and Stange, 2013), but differ from findings for

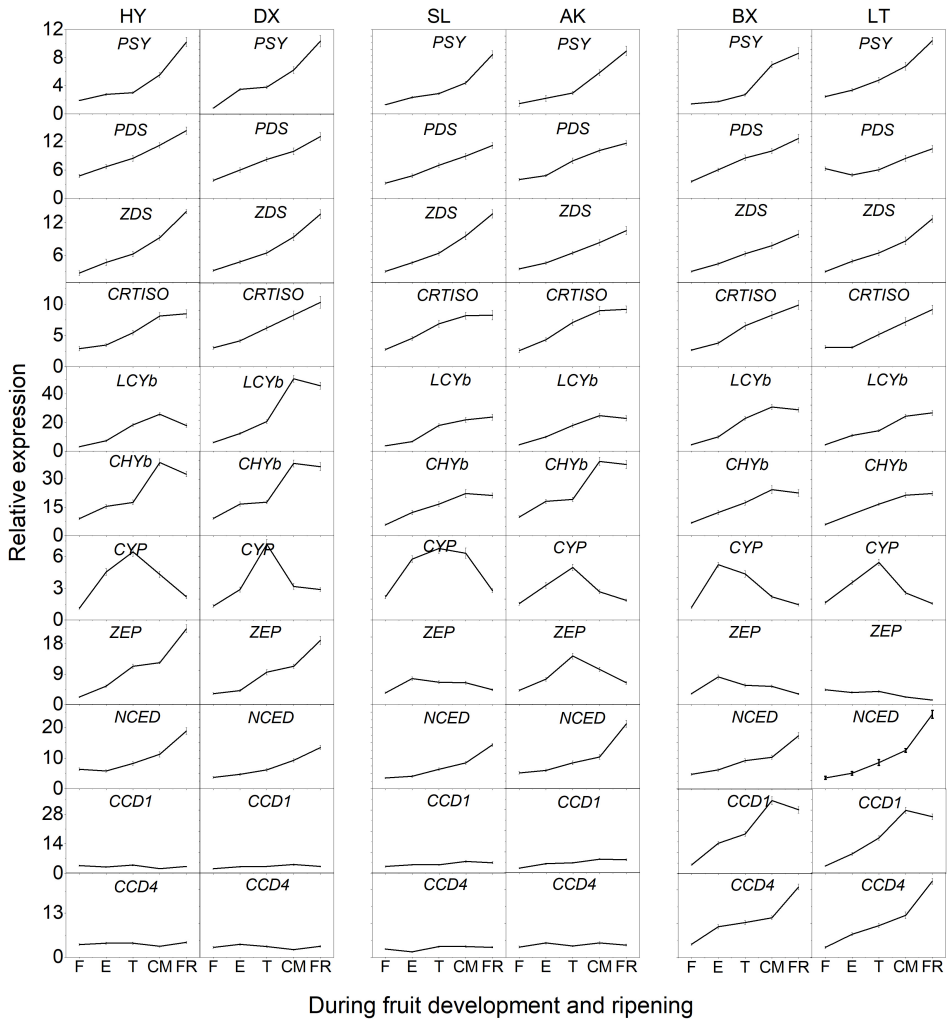


FIGURE 6 | Expression pattern of genes involved in carotenoid biosynthesis during fruit development and ripening. F, E, T, CM, and FR represent fruitlet, enlargement, turning, commercial maturation, and fully ripe developmental stages of fruit. PSY, phytoene synthase; PDS, phytoene desaturase; Z-ISO, ζ-carotene isomerase; ZDS, ζ-carotene desaturase; CRTISO, carotenoid isomerase; LCYe, lycopene ε-cyclase; LCYb, lycopene β-cyclase; CHYb, β-carotene hydroxylase; CYP, cytochrome P450-type monooxygenase 97C; ZEP, zeaxanthin epoxidase; VDE, violaxanthin de-epoxidase; NXS, neoxanthin synthase; CCD, carotenoid cleavage dioxygenase; NCED, 9-*cis*-epoxycarotenoid dioxygenase. All data are expressed as the means ± standard deviations of three biological replicates.

pepper (Ha et al., 2007) and papaya (Schweiggert et al., 2011). The increase in the content of β-cryptoxanthin was accompanied by a decrease in the content of lutein, indicating that the β-branch was strengthened in the ripening orange and yellow cultivars. However, β-cryptoxanthin was not detected in the light-yellow cultivars, suggesting that the carotenoid biosynthetic pathway in the light-yellow cultivars only involved the α-branch. The content of lutein was greatly reduced, making lutein the major pigment in the light-yellow cultivars. In contrast, lutein is the main pigment of most yellow cultivars, including pepper and *Gentiana lutea* (Zhu et al., 2014). On the whole, orange and yellow skin cultivars accumulated rich both the carotenoid composition and the total carotenoid content, these cultivars belong to the "carotenoid accumulation" cultivars during fruit development and ripening, but light-yellow skin cultivars decreased the carotenoid throughout the period, they can be as "carotenoid loss" cultivars. On the contrary, "carotenoid loss" cultivars produced rich AVAs, but "carotenoid accumulation" cultivars not. The fluctuation of carotenoids and AVAs in different skin types apricot cultivars suggested that there is a putative link between carotenoid accumulation and AVA produce in different skin types of apricot cultivars.

Carotenoid Accumulation Highly Correlated With the Transcript Level of Carotenoid Biosynthesis Related Genes

Our mRNA transcript level analysis showed that the transcript dynamics of genes involved in carotenoid biosynthesis could explain the differences in carotenoid accumulation patterns in cultivars with different skin types. For example, a low carotenoid content of loquat fruit is correlated with lower mRNA transcript

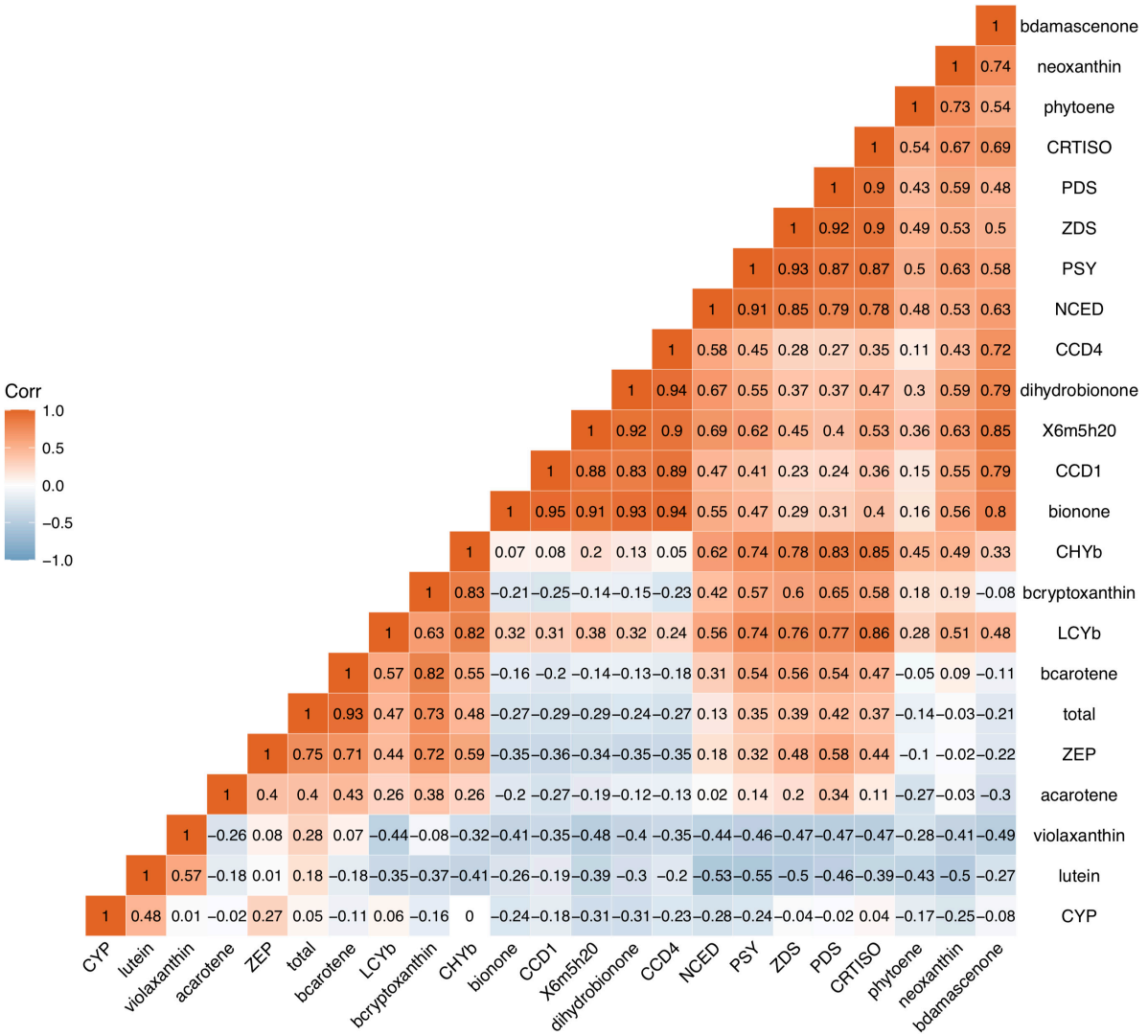


FIGURE 7 | Correlation analysis between carotenoid contents, aroma volatile apocarotenoid contents and expression levels of carotenogenic genes. b -ionone, dihydroionone, X6m5h20, bdamascenone, total, bcarotene, acarotene, and bcryptoxanthin, represent β-ionone, dihydro-β-ionone, 6-methyl-hepten-2-one, β-damascenone, total, bcarotene, acarotene, and bcryptoxanthin total carotenoid, β-carotene, α-carotene, and β-cryptoxanthin, respectively. These genes are same as them in **Figure 6**.

levels of *PSY1*, *CYCB*, and *BCH* (Fu et al., 2012; Hadjipieri et al., 2017). Phytoene synthase (PSY) is the core determinant of the total amount of carotenoid accumulation in plant tissues (Burkhardt et al., 2010), this enzyme produces phytoene, providing the synthetic precursor for other carotenoids. Previous studies have shown that upregulation of *PSY* in watermelon promotes carotenoid accumulation (Lv et al., 2015). In the present study, we found that the transcript patterns of *PSY* were not significantly different among the three cultivars, but the relative transcription level of the *PSY* gene increased significantly, accompanied by carotenoid accumulation during apricot fruit development and ripening, suggesting that *PSY* plays an important regulatory role in carotenoid biosynthesis and accumulation. Phytoene desaturase (PDS) and ζ-carotene desaturase (ZDS) are responsible for the synthesis of lycopene from phytoene, and mutations in the *PDS* and *ZDS* genes result in the accumulation of phytoene and ζ-carotene in maize and *Arabidopsis*, respectively (Dong et al., 2007). Here, we found that phytoene was not predominant during ripening in the three skin types of apricot cultivars, suggesting no significant correlation between carotenoid accumulation and the expression of the *PDS* and *ZDS* genes, and these two genes play relatively minor roles in apricot carotenoid accumulation during ripening. *Arabidopsis*, tomato and melon mutants of carotenoid isomerase (CRTISO) accumulate *cis*-carotenes in the etioplasts of seedlings or chromoplasts of the fruit (Galpaz et al., 2013). We found that the transcript level of *CRTISO* increased significantly and promoted the accumulation of β-carotene

during the process of fruit development and ripening. Lycopene β-cyclase (LCYb) allows lycopene to synthesize β-carotene (Alquézar et al., 2009), and a high transcript level of *LCYb* contributes to the accumulation of β-carotene. The accumulation pattern of β-carotene in different kiwifruit cultivars is related to the expression of the *LCYb* gene (Ampomah-Dwamena et al., 2009). The content of β-carotene in the fruit was found to be increased 3.8 times after the *LCYb* gene was introduced into the tomato (Rosati et al., 2010). In this work, higher expression of *LYCb* was accompanied by β-carotene accumulation in the three cultivars. β-carotene hydroxylase (CHYb) catalyzes the conversion of β-carotene to β-cryptoxanthin and zeaxanthin, and the expression of *CHYb* results in higher accumulation of β-cryptoxanthin in the orange and yellow cultivars than in the light-yellow cultivars. The transcript level of the ε-carotene hydroxylase gene *CYP97* (*CYP*) is positively correlated with lutein production in the carotenoid biosynthesis pathway (Yuan et al., 2015), and it has been shown that the *Arabidopsis* lut5 mutant, which is defective in CYP97A3 hydroxylase, accumulates higher levels of α-carotene (Kim and DellaPenna, 2006). The expression level of *CYP* and the content of lutein first increased and then decreased during development and ripening in the three examined skin types of apricot cultivars. Zeaxanthin epoxidase (ZEP) is an enzyme in the xanthophyll cycle, and ZEP catalyzes the conversion of zeaxanthin to violaxanthin. Loss of *ZEP* function in the aba1 mutant of *Arabidopsis* and the aba2 mutant of tobacco causes the accumulation of high zeaxanthin levels in leaves and lower ABA levels (Marin et al., 1996). In this study, the change in the content of violaxanthin was consistent with the change in the *ZEP* gene transcript level, indicating that the *ZEP* gene positively regulates the accumulation of violaxanthin in the carotenoid biosynthesis pathway.

Two Carotenoid Cleavage Dioxygenase Genes Play an Important Role in Coloration and Aroma Formation of Apricots

Apocarotenoids are formed by the enzymatic cleavage of carotenoids with the assistance of carotenoid cleavage dioxygenase (CCD) family proteins (**Figure 1**; Walter and Strack, 2011), which are non-heme iron (II) dependent enzymes (Harrison and Bugg, 2014). In plants, CCDs generally include 9-*cis*-epoxycarotenoid dioxygenase (NCED), CCD7, CCD8, CCD4, and CCD1 (Walter et al., 2010). NCED enzymes specifically cleave double bond at the 11,12 position of 9-*cis*-epoxycarotenoids, resulting in the production of *cis*-xanthoxin, an abscisic acid (ABA) precursor (Schwartz et al., 2001). AtCCD7 catalyzes the asymmetric cleavage of β-carotene at the 9′,10′ position, producing 10′-apo-β-caroten-10′-al and β-ionone, while 10′-apo-β-caroten-10′-al is subsequently cleaved by CCD8, yielding strigolactones (SLs) (Seto and Yamaguchi, 2014). AtCCD4, a negative regulator of the carotenoid accumulation in Arab. thaliana seeds, cleaves C9–C10 double bond of carotenoids (Bruno et al., 2016). The CCD4 protein cleaves β-cryptoxanthin and zeaxanthin at the 7,8 or 7′,8′ position to form β-citraurin, thereby influencing color formation in citrus (Ma et al., 2013). Previous studies found that the emission of β-ionone from *Petunia hybrida* flowers correlated strongly with the expression levels of *PhCCD1* in corollas (Simkin et al., 2004). In strawberry, the expression of *CCD1* reduces the content of lutein (García-Limones et al., 2008). Studies have shown that *CCD4* can control the pigmentation of peach flesh, and high *CCD4* transcript abundance in white flesh is related to the release of carotenoid-derived volatiles (Falchi et al., 2013). During apricot fruit development and ripening, we found that the CCDs gene family members *CCD1* and *CCD4* showed lower transcript levels in the orange and yellow cultivars but higher transcript levels in the light-yellow cultivars. The light-yellow cultivars exhibit the lowest total carotenoid and β-carotene contents but the highest transcript levels of the corresponding genes among the three skin types of cultivars. By contrast, the orange cultivar shows the lowest transcript levels and the highest total carotenoid and β-carotene contents. Simultaneously, the richest contents of aroma volatile apocarotenoids were detected in light-yellow fruit, their contents were almost undetected in orange cultivars, and the contents positively correlated significantly with the expression levels of *CCD1* and *CCD4*, but correlated with negatively total carotenoid accumulation. Furthermore, we compared the DNA sequences of *CCD1* and *CCD4* from the six experimental cultivars found that these sequences are the same, showing that expression levels of *CCD1* and *CCD4* is important for carotenoid accumulation in these cultivar fruit. These findings suggest that the total carotenoid and β-carotene contents in apricot fruit are negatively regulated by *CCD1* and *CCD4*.

CONCLUSION

The total carotenoid content continuously increased in the orange cultivars during ripening, while it decreased in the light-yellow cultivars, and the orange cultivars presented the highest levels of total carotenoids, followed by the yellow cultivars. The composition and contents of carotenoids differed in the cultivars. High levels of β-carotene and specific α-carotene contributed greatly to the coloration of the orange cultivars, but the yellow cultivars were characterized by high levels of oxidated carotenoids and phytoene, low levels of β-carotene, an absence of β-carotene and β-cryptoxanthin and decreased levels of oxidated carotenoids, resulting in light-yellow coloration. Based on the accumulation patterns of carotenoids in different apricot cultivars during development and ripening, the apricot cultivars can be divided into two types: "carotenoid accumulation" and "carotenoid loss" types of cultivars. In contrast, "carotenoid loss" types of cultivars fruit were characterized by rich AVAs such as β-ionone. The differences in carotenoid and AVAs accumulation between the cultivars were coordinately determined by the differential expression of carotenoid biosynthetic and cleaving genes, especially *PSY*, *PDS*, *ZDS*, *LCYb*, *ZEP*, *NCED*, *CCD1*, and *CCD4*. However, it is necessary to further investigate how these enzymes regulates carotenoids and AVAs accumulation through substrate specific identification *in vivo* and *in vitro* and transgenic means to verify the role played by their enzyme.

This study unveils the chemical mechanism of color and aroma of different skin types of apricot cultivars, provides useful targets for controlling fruit quality in future work.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**, further inquiries can be directed to the corresponding author.

AUTHOR CONTRIBUTIONS

GZ designed the experiments. WX performed the experiments and wrote the manuscript. LZ and SL analyzed the data. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by the National Natural Science Foundation of China (No. 31872046).

ACKNOWLEDGMENTS

We thank Junhui Zhou (Department of Cell Biology and Molecular Genetics, University of Maryland) for comments on an earlier version of this manuscript.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2020.607715/full#supplementary-material

Supplementary Figure 1 | HPLC chromatogram of carotenoids from the peel of "Danxing" at the fully ripe stage.

Supplementary Figure 2 | HPLC chromatogram of carotenoids from the peel of "Baixing" during development and ripening. The black line, blue line, green line, indigo line, and pink line represent the fruitlet (F), enlargement (E), turning (T), commercial maturation (CM), and fully ripe (FR) stages.

Supplementary Table 1 | Primer sequences for qRT-PCR.

Supplementary Table 2 | Changes in chromatism parameters of apricot fruit during development and ripening.

REFERENCES

Alquézar, B., Zacarías, L., and Rodrigo, M. J. (2009). Molecular and functional characterization of a novel chromoplast-specific lycopene β-cyclase from *Citrus* and its relation to lycopene accumulation. *J. Exp. Bot.* 60:1783. doi: 10.1093/jxb/erp048

Ampomah-Dwamena, C., Mcghie, T., Wibisono, R., Montefiori, M., Hellens, R. P., and Allan, A. C. (2009). The kiwifruit lycopene beta-cyclase plays a significant role in carotenoid accumulation in fruit. *J. Exp. Bot.* 60:3765. doi: 10.1093/jxb/erp218

Botondi, R., DeSantis, D., Bellincontro, A., Vizovitis, K., and Mencarelli, F. (2003). Influence of ethylene inhibition by 1-methylcyclopropene on apricot quality, volatile production, and glycosidase activity of low- and high-aroma varieties of apricots. *J. Agric. Food Chem.* 51, 1189–1200. doi: 10.1021/jf025893o

Bruno, M., Koschmieder, J., Wuest, F., Schaub, P., Fehling-Kaschek, M., Timmer, J., et al. (2016). Enzymatic study on AtCCD4 and AtCCD7 and their potential to form acyclic regulatory metabolites. *J. Exp. Bot.* 67, 5993–6005. doi: 10.1093/jxb/erx347

Burkhardt, P. K., Beyer, P., Wünn, J., Klöti, A., Armstrong, G. A., Schledz, M., et al. (2010). Transgenic rice (*Oryza sativa*) endosperm expressing daffodil (*Narcissus pseudonarcissus*) phytoene synthase accumulates phytoene, a key intermediate of provitamin a biosynthesis. *Plant J.* 11, 1071–1078. doi: 10.1046/j.1365-313X.1997.11051071.x

Dong, H., Deng, Y., Mu, J. Y., Lu, Q. T., Wang, Y. Q., Xu, Y. Y., et al. (2007). The *Arabidopsis* spontaneous cell death1 gene, encoding a ζ-carotene desaturase essential for carotenoid biosynthesis, is involved in chloroplast development, photoprotection and retrograde signaling. *Cell Res.* 17:458. doi: 10.1038/cr.2007.51

Enfissi, E. M. A., Nogueira, M., Bramley, P. M., and Fraser, P. D. (2017). The regulation of carotenoid formation in tomato fruit. *Plant J.* 89, 774–788. doi: 10.1111/tpj.13428

Erdogan-Orhan, I., and Kartal, M. (2011). Insights into research on phytochemistry and biological activities of *Prunus armeniaca L.* (apricot). *Food Res. Int.* 44, 1238–1243. doi: 10.1016/j.foodres.2010.11.014

Falchi, R., Vendramin, E., Zanon, L., Scalabrin, S., Cipriani, G., Verde, I., et al. (2013). Three distinct mutational mechanisms acting on a single gene underpin the origin of yellow flesh in peach. *Plant J.* 76, 175–187. doi: 10.1111/tpj.12283

Fraser, P. D., and Bramley, P. M. (2004). The biosynthesis and nutritional uses of carotenoids. *Prog. Lipid Res.* 43, 228–265. doi: 10.1016/j.plipres.2003.10.002

Fu, X. M., Kong, W. B., Peng, G., Zhou, J. Y., Azam, M., Xu, C. J., et al. (2012). Plastid structure and carotenogenic gene expression in red- and white-fleshed loquat (*Eriobotrya japonica*) fruits. *J. Exp. Bot.* 63, 341–354. doi: 10.1093/jxb/err284

Galpaz, N., Burger, Y., Lavee, T., Tzuri, G., Sherman, A., Melamed, T., et al. (2013). Genetic and chemical characterization of an EMS induced mutation in *Cucumis melo* CRTISO gene. *Arch. Biochem. Biophys.* 539, 117–125. doi: 10.1016/j.abb.2013.08.006

García-Limones, C., Schnäbele, K., Blanco-Portales, R., Bellido, M. L., Caballero, J. L., Schwab, W., et al. (2008). Functional characterization of FaCCD1: a carotenoid cleavage dioxygenase from strawberry involved in lutein degradation during fruit ripening. *J. Agric. Food Chem.* 56, 9277–9285. doi: 10.1021/jf801096t

Ha, S. H., Kim, J. B., Park, J. S., Lee, S. W., and Cho, K. J. (2007). A comparison of the carotenoid accumulation in *Capsicum* varieties that show different ripening colours: deletion of the capsanthin-capsorubin synthase gene is not a prerequisite for the formation of a yellow pepper. *J. Exp. Bot.* 17, 3135–3144. doi: 10.1007/s00894-010-0927-x

Hadjipieri, M., Georgiadou, E. C., Marin, A., Diaz-Mula, H. M., Goulas, V., Fotopoulos, V., et al. (2017). Metabolic and transcriptional elucidation of the carotenoid biosynthesis pathway in peel and flesh tissue of loquat fruit during on-tree development. *BMC Plant Biol.* 17:102. doi: 10.1186/s12870-017-1041-3

Harrison, P. J., and Bugg, T. D. (2014). Enzymology of the carotenoid cleavage dioxygenases: reaction mechanisms, inhibition and biochemical roles. *Arch. Biochem. Biophys.* 544, 105–111. doi: 10.1016/j.abb.2013.10.005

Kim, J., and DellaPenna, D. (2006). Defining the primary route for lutein synthesis in plants: the role of *Arabidopsis* carotenoid β-ring hydroxylase CYP97A3. *Proc. Natl. Acad. Sci. U.S.A.* 103, 3474–3479. doi: 10.1073/pnas.0511207103

Le Bourvellec, C., Gouble, B., Bureau, S., Reling, P., Bott, R., Ribas-Agusti, A., et al. (2018). Impact of canning and storage on apricot carotenoids and polyphenols. *Food Chem.* 240, 615–625. doi: 10.1016/j.foodchem.2017.07.147

Luan, Y., Fu, X., Lu, P., Grierson, D., and Xu, C. (2020). Molecular mechanisms determining the differential accumulation of carotenoids in plant species and varieties. *Crit. Rev. Plant Sci.* 39, 125–139. doi: 10.1080/07352689.2020.1768350

Lv, P., Li, N., Lui, H., Gu, H. H., and Zhao, W. E. (2015). Changes in carotenoid profiles and in the expression pattern of the genes in carotenoid metabolisms during fruit development and ripening in four watermelon cultivars. *Food Chem.* 174, 52–59. doi: 10.1016/j.foodchem.2014.11.022

Ma, G., Zhang, L., Matsuta, A., Matsutani, K., Yamawaki, K., Yahata, M., et al. (2013). Enzymatic formation of β-citraurin from β-cryptoxanthin and zeaxanthin by carotenoid cleavage dioxygenase4 in the flavedo of *Citrus* fruit. *Plant Physiol.* 163, 682–695. doi: 10.1104/pp.113.223297

Marin, E., Nussaume, L., Quesada, A., Gonneau, M., Sotta, B., Hugueney, P., et al. (1996). Molecular identification of zeaxanthin epoxidase of *Nicotiana plumbaginifolia*, a gene involved in abscisic acid biosynthesis and corresponding to the ABA locus of *Arabidopsis thaliana*. *Embo J.* 15, 2331–2342. doi: 10.1002/j.1460-2075.1996.tb00589.x

Marty, I., Bureau, S., Sarkissian, G., Gouble, B., Audergon, J. M., and Albagnac, G. (2005). Ethylene regulation of carotenoid accumulation and carotenogenic gene expression in colour-contrasted apricot varieties (*Prunus armeniaca*). *J. Exp. Bot.* 56, 1877–1886. doi: 10.1093/jxb/eri177

McQuinn, R. P., Giovannoni, J. J., and Pogson, B. J. (2015). More than meets the eye: from carotenoid biosynthesis, to new insights into apocarotenoid signaling. *Curr. Opin. Plant Biol.* 27, 172–179. doi: 10.1016/j.pbi.2015.06.020

Melgarejo, P., Calín-Sánchez, A., Carbonell-Barrachina, A. A., Martínez-Nicolás, J. J., Legua, P., Martínez, R., et al. (2014). Antioxidant activity, volatile composition and sensory profile of four new very-early apricots (*Prunus armeniaca L.*). *J. Sci. Food Agric.* 94, 85–94. doi: 10.1002/jsfa.6201

Nisar, N., Li, L., Lu, S., Khin, N. C., and Pogson, B. J. (2015). Carotenoid metabolism in plants. *Mol. Plant* 8, 68–82. doi: 10.1016/j.molp.2014.12.007

Qi, Y. W., Liu, X. S., Zhang, Q., Wu, H. X., Yan, D., Liu, Y. F., et al. (2019). Carotenoid accumulation and gene expression in fruit skins of three differently colored persimmon cultivars during fruit growth and ripening. *Sci. Hortic.* 248, 282–290. doi: 10.1016/j.scienta.2018.12.042

Rodriguez-Concepcion, M., and Stange, C. (2013). Biosynthesis of carotenoids in carrot: an underground story comes to light. *Arch. Biochem. Biophys.* 539, 110–116. doi: 10.1016/j.abb.2013.07.009

Rosati, C., Aquilani, R., Dharmapuri, S., Pallara, P., Marusic, C., Tavazza, R., et al. (2010). Metabolic engineering of beta-carotene and lycopene content in tomato fruit. *Plant J.* 24, 413–420. doi: 10.1046/j.1365-313x.2000.00880.x

Sagawa, J. M., Stanley, L. E., Lafountain, A. M., Frank, H. A., Liu, C., and Yuan, Y. W. (2016). An R2R3—MYB transcription factor regulates carotenoid pigmentation in *Mimulus lewisii* flowers. *New Phytol.* 209, 1049–1057. doi: 10.1111/nph.13647

Saini, R. K., Nile, S. H., and Park, S. W. (2015). Carotenoids from fruits and vegetables: chemistry, analysis, occurrence, bioavailability and biological activities. *Food Res. Int.* 76, 735–750. doi: 10.1016/j.foodres.2015.07.047

Schwartz, S. H., Qin, X., and Zeevaart, J. A. (2001). Characterization of a novel carotenoid cleavage dioxygenase from plants. *J. Biol. Chem.* 276, 25208–25211. doi: 10.1074/jbc.M102146200

Schweiggert, R. M., Steingass, C. B., Heller, A., Esquivel, P., and Carle, R. (2011). Characterization of chromoplasts and carotenoids of red- and yellow-fleshed papaya (*Carica papaya L.*). *Planta* 234, 1031–1044. doi: 10.1007/s00425-011-1457-1

Seto, Y., and Yamaguchi, S. (2014). Strigolactone biosynthesis and perception. *Curr. Opin. Plant Biol.* 21, 1–6. doi: 10.1016/j.pbi.2014.06.001

Shan, L., Li, X., Wang, P., Cai, C., Zhang, B., Sun, C. D., et al. (2008). Characterization of cDNAs associated with lignification and their expression profiles in loquat fruit with different lignin accumulation. *Planta* 227, 1243–1254. doi: 10.1007/s00425-008-0696-2

Simkin, A. J., Schwartz, S. H., Auldridge, M., Taylor, M. G., and Klee, H. J. (2004). The tomato carotenoid cleavage dioxygenase 1 genes contribute to the formation of the flavor volatiles β-ionone, pseudoionone, and geranylacetone. *Plant J.* 40, 882–892. doi: 10.1111/j.1365-313x.2004.02263.x

Spirt, S. D., Lutter, K., and Stahl, W. (2010). Carotenoids in photooxidative stress. *Curr. Nutr. Food Sci.* 6, 36–43. doi: 10.2174/157340110790909572

Tanaka, Y., Sasaki, N., and Ohmiya, A. (2008). Biosynthesis of plant pigments: anthocyanins, betalains and carotenoids. *Plant J.* 54, 733–749. doi: 10.1111/j.1365-313X.2008.03447.x

Walter, M. H., Floss, D. S., and Strack, D. (2010). Apocarotenoids: hormones, mycorrhizal metabolites and aroma volatiles. *Planta* 232, 1–17. doi: 10.1007/s00425-010-1156-3

Walter, M. H., and Strack, D. (2011). Carotenoids and their cleavage products: biosynthesis and functions. *Nat. Prod. Rep.* 28, 663–692. doi: 10.1039/c0np00036a

Yuan, H., Zhang, J. X., Nageswaran, D., and Li, L. (2015). Carotenoid metabolism and regulation in horticultural crops. *Hortic. Res.* 2:15036. doi: 10.1038/hortres.2015.36

Zhang, Q. F., Liu, M. Y., and Ruan, J. Y. (2017). Metabolomics analysis reveals the metabolic and functional roles of flavonoids in light-sensitive tea leaves. *BMC Plant Biol.* 17:64. doi: 10.1186/s12870-017-1012-8

Zhang, Q. Y., Feng, C., Li, W. H., Qu, Z. H., Zeng, M., and Xi, W. P. (2019). Transcriptional regulatory networks controlling taste and aroma quality of apricot (*Prunus armeniaca L.*) fruit during ripening. *BMC Genomics* 20:45. doi: 10.1186/s12864-019-5424-8

Zheng, H. W., Zhang, Q. Y., Quan, J. P., Zheng, Q., and Xi, W. P. (2016). Determination of sugars, organic acids, aroma components, and carotenoids in grapefruit fleshs. *Food Chem.* 205, 112–121. doi: 10.1016/j.foodchem.2016.03.007

Zhu, C. F., Yang, Q. J., Ni, X. Z., Bai, C., Sheng, Y. M., Shi, L. X., et al. (2014). Cloning and functional analysis of the promoters that upregulate carotenogenic gene expression during flower development in *Gentiana lutea*. *Physiol. Plant.* 150, 493–504. doi: 10.1111/ppl.12129

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.

frontiers in Plant Science

ORIGINAL RESEARCH
published: 07 May 2021
doi: 10.3389/fpls.2021.646222

Fine Mapping and Candidate Gene Identification of a White Flower Gene *BrWF3* in Chinese Cabbage (*Brassica rapa* L. ssp. *pekinensis*)

Shuangjuan Yang[1,2†], Xinxin Tian[1,2†], Zhiyong Wang[1†], Xiaochun Wei[1], Yanyan Zhao[1], Henan Su[1], Xiaobin Zhao[1], Baoming Tian[2*], Yuxiang Yuan[1,2*] and Xiao-Wei Zhang[1,2*]*

[1] Institute of Horticulture, Henan Academy of Agricultural Sciences, Postgraduate T&R Base of Zhengzhou University, Zhengzhou, China, [2] School of Agricultural Sciences, Zhengzhou University, Zhengzhou, China

Flower color is an important trait in plants. However, genes responsible for the white flower trait in Chinese cabbage are rarely reported. In this study, we constructed an F_2 population derived from the Y640-288 (white flower) and Y641-87 (yellow flower) lines for the fine mapping of the white flower gene *BrWF3* in Chinese cabbage. Genetic analysis indicated that *BrWF3* was controlled by a single recessive gene. Using BSA-seq and KASP assays, *BrWF3* was fine-mapped to an interval of 105.6 kb. Functional annotation, expression profiling, and sequence variation analyses confirmed that the *AtPES2* homolog, *Bra032957,* was the most likely candidate gene for *BrWF3*. Carotenoid profiles and transmission electron microscopy analysis suggested that *BrWF3* might participate in the production of xanthophyll esters (particularly violaxanthin esters), which in turn disrupt chromoplast development and the formation of plastoglobules (PGs). A SNP deletion in the third exon of *BrWF3* caused the loss of protein function, and interfered with the normal assembly of PGs, which was associated with reduced expression levels of genes involved in carotenoid metabolism. Furthermore, we developed and validated the functional marker TXBH83 for *BrWF3.* Our results provide insight into the molecular mechanism underlying flower color pigmentation and reveal valuable information for marker-assisted selection (MAS) breeding in Chinese cabbage.

Keywords: ***Brassica rapa*, white flower, gene cloning, carotenoid, plastoglobule**

OPEN ACCESS

Edited by:
Lourdes Gómez-Gómez, University of Castilla-La Mancha, Spain

Reviewed by:
Maria Jesus Rodrigo, Institute of Agrochemistry and Food Technology (IATA), Spain
Dámaso Hornero-Mendez, Consejo Superior de Investigaciones Científicas (CSIC), Spain
Melaku Gedil, International Institute of Tropical Agriculture (IITA), Nigeria

***Correspondence:**
Shuangjuan Yang
sjyang_0614@163.com
Baoming Tian
tianbm@zzu.edu.cn
Yuxiang Yuan
yuxiangyuan126@126.com
Xiao-Wei Zhang
xiaowei5737@163.com

† These authors have contributed equally to this work and share first authorship

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

Received: *14 January 2021*
Accepted: *13 April 2021*
Published: *07 May 2021*

Citation:
Yang S, Tian X, Wang Z, Wei X, Zhao Y, Su H, Zhao X, Tian B, Yuan Y and Zhang X-W (2021) Fine Mapping and Candidate Gene Identification of a White Flower Gene BrWF3 in Chinese Cabbage (Brassica rapa L. ssp. pekinensis). Front. Plant Sci. 12:646222. doi: 10.3389/fpls.2021.646222

INTRODUCTION

Flower color is one of the most important traits in plants, which provides a visual signal to attract insects for pollination (Kevan and Baker, 1983; Ariizumi et al., 2014). It also protects plants against disease and UV radiation and helps to maintain the normal physiological function of floral organs (Koes et al., 1994). In breeding, flower color can be used for identifying true/false

Abbreviations: BSA-seq, bulked segregant analysis coupled with whole-genome sequencing; CDS, coding sequence; DEGs, differentially expressed genes; DH, doubled haploid; FKPM, the fragments per kilobase of transcript per million; InDel, insertion-deletion; KASP, Kompetitive allele-specific PCR; KEGG, Kyoto encyclopedia of genes and genomes; MAS, molecular assisted selection; PGs, plastoglobules; qRT-PCR, Quantitative real time PCR; SNP, single nucleotide polymorphism; TEM, transmission electron microscopy; LC-MS/MS, liquid chromatography-tandem mass spectrometry.

hybrids and for evaluating seed purity in hybrid production (Zhang et al., 2018b). Carotenoids, flavonoids and betalains are three main classes of natural pigments contributing to different flower colors, among which the accumulation of carotenoids can cause yellow, orange and red colorations (DellaPenna and Pogson, 2006; Grotewold, 2006). In nature, greater than 800 structurally different carotenoid compounds have been identified (Nisar et al., 2015; Ding et al., 2019), which are further divided into two main groups, carotenes and xanthophylls (Nisar et al., 2015; Ding et al., 2019). In many cases, xanthophylls (i.e., letein, zeaxanthin, and violaxanthin) are the most prevalent carotenoids in yellow flowers (Ohmiya, 2011). The amount of carotenoids is a net result of biosynthesis, degradation and stable storage (Deruere et al., 1994; Li and Yuan, 2013; Ariizumi et al., 2014; Nisar et al., 2015; Yuan et al., 2015). Thus, factors that are associated with these processes act together to regulate the final carotenoid levels. Almost all of the genes and enzymes that catalyze the core reactions of carotenoid biosynthesis and degradation have been identified in plants (Yuan et al., 2015), whereas only a few genes have been reported to be involved in carotenoid sequestration and storage.

Carotenoids accumulate at high levels in chromoplasts, which possess a superior storage capacity to deposit carotenoids in carotenoid-lipoprotein sequestering structures, such as plastoglobules (PGs) (Li and Yuan, 2013; Yuan et al., 2015; Li et al., 2016). These structures contain carotenoids, polar lipids and carotenoid-associated proteins (van Wijk and Kessler, 2017). Carotenoids occupying the inner core interact with the acyl residues of polar lipids, which subsequently interact with carotenoid-associated proteins via polar head groups (Deruere et al., 1994; Vishnevetsky et al., 1999). Genes participating in the biogenesis of chromoplasts and the formation of carotenoid sequestration structures exert a strong effect on carotenoid metabolism in crops. For example, the *Or* gene, which encodes a DnaJ cysteine-rich domain-containing protein, triggers the differentiation of non-colored plastids into chromoplasts with an increased capacity to accumulate β-carotene in cauliflower and potato (Lu et al., 2006; Lopez et al., 2008). Either the *fibrillin* gene in pepper or the *CHRC* gene in cucumber, which encode carotenoid-associated proteins, is positively associated with carotenoid accumulation (Vishnevetsky et al., 1996; Pozueta-Romero et al., 1997). Overexpression of the pepper *fibrillin* gene in tomato increases the levels of carotenoids in fruits (Simkin et al., 2007). The *PYP1* gene in tomato, which is homologous to *PES2* (*PHYTYL ESTER SYNTHASE 2*) in *Arabidopsis*, plays a vital role in the production of xanthophyll esters in tomato anthers and petals (Ariizumi et al., 2014). Functional disruption of PYP1 converts flower color from yellow to pale yellow (Ariizumi et al., 2014). In pale-yellow-flowered petunia, the lower expression of *xanthophyll esterase* (*XES*) and lower proportions of esterified xanthophylls are the main reasons for low carotenoid accumulation (Kishimoto et al., 2019). Overexpression of *XES* from petals of *Ipomoea obscura*, tomato (*PYP1* gene), and marigold (*Tagetes erecta*) in this pale-yellow-flowered petunia all increases the esterified xanthophylls and causes deeper yellow coloration in flowers (Kishimoto et al., 2020).

In *Brassica* species, several genes controlling flower color have been reported. In *B.napus* and *B.oleracea*, the white flower trait is controlled by a single dominant gene, *carotenoid cleavage dioxygenase 4* (*CCD4*). A CACTA-like transposable element insertion in *CCD4* results in a petal color transition from white to yellow (Zhang B. et al., 2015; Han et al., 2019). In *B.juncea*, two recessive genes (*Bjpc1* and *Bjpc2*) control the white flower gene (Zhang et al., 2018a,b). These two genes, which are located on chromosomes A02 and B04, respectively, are homologous to *AtPES2* and participate in xanthophyll esterification (Zhang et al., 2018a,b). In *B.rapa*, the *carotenoid isomerase* (*BrCRTISO*) gene controls orange flower color as well as the orange coloration of the inner leaves of Chinese cabbage (Su et al., 2014; Zhang J. X. et al., 2015). Recently, Zhang N. et al. (2020) reported that the white flower trait in Chinese cabbage is controlled by two recessive loci, *Brwf1* and *Brwf2*. These two genes are located on chromosomes A01 and A09 and encode a plastid-lipid associated protein (PAP) and a carotenoid isomerase enzyme, respectively. Another study revealed that the white flower trait in *B.rapa* is controlled by a single recessive gene (Rahman, 2001). However, the gene underlying this white flower trait has not been reported thus far.

In this study, we conducted positional cloning of the white flower gene (*BrWF3*) in Chinese cabbage by using F_2 populations derived from the white-flowered DH line 'Y640-288' and the yellow-flowered DH line 'Y641-87'. The *BrWF3* gene was mapped to chromosome A02, and *Bra032957*, which is homologous to *AtPES2*, was identified as the candidate gene. Based on carotenoids profile analysis and transmission electron microscopy (TEM) analysis, as well as transcriptome analysis, the *BrWF3* gene was predicted to participate in carotenoids esterification and the biogenesis of PGs. A functional marker of *BrWF3* was also developed and validated in our study. This work will promote molecular marker-assisted selection (MAS) breeding and the exploration of molecular mechanisms that regulate flower color variation in Chinese cabbage.

MATERIALS AND METHODS

Plant Materials

Two Chinese cabbage DH lines, white-flowered Y640-288 and yellow-flowered Y641-87, were used as parental lines to generate F_1 and F_2 populations for inheritance analysis and gene mapping. Additionally, three F_2 populations, (Y640-288 × SD369)-F_2, (Y640-288 × Chiifu)-F_2, (Y66-83 × R16-11)-F_2, were generated for marker validation by crossing the white-flowered DH lines Y640-288 and Y66-83 with the yellow-flowered DH lines SD369, Chiifu and R16-11. Furthermore, ten white-flowered and ten yellow-flowered DH lines (**Supplementary Table 1**) were also used to analyze mutations in the candidate gene. All the materials used in this study were provided by Institute of Horticulture, Henan Academy of Agricultural Sciences.

Transmission Electron Microscopy (TEM) Analysis

Petals from Y640-288 and Y641-87 flowers at the flowering stage were used for transmission electron microscopy (TEM) analysis, which was performed according to Ariizumi et al. (2004).

Carotenoid Identification and Quantification

Carotenoid composition was measured by MetWare (Wuhan, China). Petals from 10 white-flowered F_2 plants were combined to form one replicate W-bulk and petals from 10 yellow-flowered F_2 plants were included in the Y-bulk. In total, three replicates were assessed. Fresh petals were frozen in liquid nitrogen and stored at −80°C until needed. The direct extraction steps were performed according to Zhou et al. (2020). The saponified extraction steps were performed according to Inbaraj et al. (2008) with some modification. The direct and saponified extracts were then analyzed using an LC-APCI-MS/MS system (UHPLC, ExionLCs AD; MS, Applied Biosystems 6500 Triple Quadrupole). The chromatographic conditions and parameters for API 6500 Q TRAP LC-MS/MS System were performed as previously reported (Liu et al., 2020; Wang et al., 2020; Zhou et al., 2020). Specific procedures for extraction, identification and quantification of carotenoids were supported in **Supplementary Material 1**.

Carotenoids were identified by comparing their retention times and ion pair information (**Supplementary Table 2**). In saponified extracts, the integrated peak area was substituted into the linear equations of standard (Sigma, St. Louis, MO, United States) curves for content calculation (**Supplementary Table 3**); finally, the absolute content data for the carotenoids in the actual samples were obtained. Carotenoid content (μg/g) = B*C/1000/D, where B is the concentration (μg/mL) obtained by substituting the integrated peak area of a carotenoid in the sample into the corresponding standard curve, C is the resuspension volume (μL), and D is the mass of the weighed sample (g) (Liu et al., 2020; Wang et al., 2020; Zhou et al., 2020).

Bulked Segregant Analysis (BSA) by Resequencing

Using the modified cetyltrimethylammonium bromide (CTAB) method, total genomic DNA was isolated from young leaves of the parents and F_2 (Y640-288 × Y641-87) plants (Liu et al., 2003). For BSA-seq, two DNA pools were constructed by mixing equal amounts of DNA from 50 white-flowered F_2 individuals (W-pool) and 50 yellow-flowered F_2 individuals (Y-pool). The Illumina HiSeq X Ten platform was used to generate 150-base paired-end reads for Y640-288, Y641-87, the W-pool and the Y-pool by Anoroad Biotech Co., Ltd. (Beijing, China). The raw data were deposited in the Sequence Read Archive (SRA) in NCBI as PRJNA682710.

The clean reads of Y640-288, Y641-87, the W-pool and the Y-pool were aligned to the *B.rapa* Chiifu reference genome version 1.5 using BWA software (Li and Durbin, 2010). SAMtools software package (version 1.3.1) (Li et al., 2009) was then used to call single-nucleotide polymorphism (SNP) and insertion/deletion (InDel) variants based on alignment files. To identify candidate regions associated with the white flower trait, the SNP-index and △(SNP-index) were calculated for all genomic positions in the W-pool and Y-pool. The SNP-index was estimated from the proportion of reads harboring SNPs among the entire number of reads compared to the reference genome sequence (Abe et al., 2012). Then △(SNP-index) was calculated by subtracting the SNP-index of the Y-pool from that of the W-pool (Takagi et al., 2013). We filtered out SNPs with SNP-index <0.3 or >0.8 simultaneously in the two bulked pools. Furthermore, SNPs with heterozygous genotypes in the parental lines were also excluded. A 1-Mb sliding window with a 50-kb increment was applied to slide across the genome, and Δ(SNP-index) graphs were plotted using the average Δ(SNP-index) against the positions of each sliding window. We calculated the statistical confidence intervals of Δ(SNP-index) among all SNP positions with given read depths under the null hypothesis of no major genes, and the 95% and 99% confidence intervals of the Δ(SNP-index) were then generated for each read depth according to Takagi et al. (2013).

Kompetitive Allele-Specific PCR (KASP) Marker and Linkage Map Development

To map the *BrWF3* gene, we extracted the 70-bp upstream and downstream sequences of the selected SNP for KASP marker development. For each selected SNP, two allele-specific forward primers and one common reverse primer were designed using the Primer Premier 5.0 program (Singh et al., 1998) according to the standard KASP guidelines. The two allele-specific primers were added with the standard FAM (5′-GAAGGTGACCAAGTTCATGCT-3′) and HEX (5′-GAAGGTCGGAGTCAACGGATT-3′) tails at the 5′ end. The developed KASP markers were first validated in the two parental lines and F_1 plants for polymorphism screening. Then, the polymorphic KASP markers (**Supplementary Table 4**) were employed to genotype the F_2 population using 135 individuals. KASP assays were performed as described by Yang et al. (2020). The genetic linkage map was constructed using JoinMap 4.0[1] software. Recombination values were converted into genetic map distances (cM) following the Kosambi mapping function (Kosambi, 1944).

Cloning and Sequence Analysis of the Candidate Genes

To clone the DNA and cDNA sequences of the putative candidate genes, primers (**Supplementary Table 5**) were designed according to the *B.rapa* reference genome. PCR amplification was performed in a total volume of 25 μL according to the manual supplied with Phanta Max Master Mix (Vazyme Biotech Co., Ltd., Nanjing, China). The PCR products were sequenced by Sunya Biotech Co., Ltd. (Zhengzhou, China), and sequence alignments were performed using DNAMAN software. The complete coding sequences of candidate gene from Y641-87 and Y640-288 were submitted to GenBank under the accession numbers: MW362118 (Y641-87) and MW362119 (Y640-288), respectively.

[1]https://www.kyazma.nl/index.php/JoinMap/

Quantitative Real-Time PCR (qRT-PCR)

Total RNA was extracted using the RNAprep Pure Plant Kit (Tiangen, Beijing, China) according to the manufacturer's instructions. First-strand cDNAs were synthesized in a 20 μL volume containing approximately 7 μg RNA and oligo (dT) primers using the TransScript One-Step gDNA Removal and cDNA Synthesis Kit (Trans, Beijing, China). qRT-PCR was performed with 2 × TB Green Premix Ex Taq II (TaKaRa, Japan) on a Roche LightCycler 480-II System (Roche Applied Sciences, Beijing, China). The *Bractin* gene was used as an internal control (Fujimoto et al., 2006; Takada et al., 2019). Each qRT-PCR experiment was performed in triplicate and the resultant mean value was used for qRT-PCR analysis (Zhang et al., 2013). Relative expression levels were calculated using the $2^{-\Delta\Delta Ct}$ method (Livak and Schmittgen, 2001).

Transcriptome Analysis

The W-bulk and Y-bulk each with three replicates used for carotenoid analysis were also used for transcriptome analysis. Six cDNA libraries were constructed and sequenced on the Illumina HiSeq X Ten platform at Metware Biotech Co., Ltd. (Wuhan, China). Raw reads were filtered by removing low-quality reads and reads containing the adapter or ploy-N using in-house Perl scripts available from Metware Biotech Co., Ltd. (Wuhan, China). The clean reads were aligned to the *B.rapa* V1.5 reference genome using HISAT2 software (Kim et al., 2015). Differentially expressed genes (DEGs) were identified using the DESeq2 package (v1.30.0) (Love et al., 2014). The *P*-value of the DEGs between samples was adjusted using the Benjamini & Hochberg method (Benjamini and Hochberg, 2000). Genes with an adjusted *P*-value $\leq$0.05 and $|\log2\text{ (fold change)}| \geq 1$ were recognized as DEGs. To determine the biological significance of the DEGs, Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analysis was implemented using KOBAS software (Wu et al., 2006).

RESULTS

Phenotypic Characterization and Genetic Analysis of White Flowers in Chinese Cabbage

The phenotypic analyses showed significant differences in flower color between the two parental lines. In Y641-87, the flower organs, particularly the petal tissue, showed stable yellow coloration at the flowering stage, whereas those of Y640-288 showed white coloration (**Figures 1A,B**). The flower color in Y640-288 was traditionally called white color. It was not like white papers as in *B.napus* and *B.oleracea* (Zhang B. et al., 2015; Han et al., 2019), but similar to that in *B.juncea* (Zhang et al., 2018a,b), the flowers of which still had pale yellow pigments. TEM analysis showed that the petals of Y641-87 had normal chromoplasts with numerous fully developed PGs, whereas the petals of Y640-288 showed abnormal chromoplasts with only a few irregular and small PGs (**Figures 1C,D**).

To investigate the genetic inheritance of white flowers in Chinese cabbage, we performed reciprocal crosses between Y641-87 and Y640-288. The resulting F_1 plants all displayed a yellow flower phenotype. The flowers of F_2 plants exhibited two types of colorations, corresponding to the coloration of either Y641-87 or Y640-288. Among 200 F_2 individuals, 142 exhibited yellow flowers, and 58 showed white flowers, corresponding to a segregation ratio of 3:1 by the Chi-square test (**Table 1**). In a larger F_2 population, the segregation ratio was also 3:1 (1775 yellow:596 white, $\chi^2 = 0.02$). These results demonstrated that the inheritance of the white flower trait in Y640-288 follows a monogenic recessive pattern. We named this white flower gene *BrWF3*.

Carotenoid Analysis in Yellow and White Flowers

To investigate whether the lower pigmentation in white flowers was due to decreased carotenoid accumulation, we analyzed the carotenoid profiles of a white petal pool (W-bulk) and a yellow petal pool (Y-bulk). We detected 20 carotenoid peaks in the Y-bulk in the direct extracts (**Figure 2** and **Supplementary Table 6**). Among these peaks, 9 peaks represented esterified carotenoids with retention times ranging from 5.5-7.5 min, as these peaks were not detected in the saponified sample (**Figure 2**). The esterified carotenoids are mostly derived from lutein and violaxanthin (**Supplementary Table 6**). In contrast, the composition and content of carotenoid esters in the W-bulk were much less than those in the Y-bulk (**Figure 2**). Analysis with saponification identified 10 carotenoids in both the W-bulk and Y-bulk (**Figure 2** and **Table 2**). Two major xanthophylls, violaxanthin and lutein, accounted for approximately 83 and 91% of the total carotenoids in the Y-bulk and the W-bulk, respectively. The total average content of violaxanthin and all carotenoids in Y-bulk was about 2.76 and 1.70 times higher than that in the W-bulk, whereas lutein content did not significantly differ between Y-bulk and W-bulk (**Table 2**). Taken together, these results indicated that white petals accumulate less xanthophylls esters (likely violaxanthin esters) than yellow petals, resulting in lower carotenoid accumulation and color pigmentation.

The *BrWF3* Gene Is Located on Chromosome A02

To map the *BrWF3* gene, we conducted BSA-seq using two pooled samples, which comprised 50 white-flowered (W-pool) and 50 yellow-flowered (Y-pool) F_2 plants, and two parental lines, Y640-288 and Y641-87. In total, 204, 209, 120, and 86 million raw data were generated for the W-pool, Y-pool, Y640-288 and Y641-87 (**Supplementary Table 7**), representing approximately 63-, 64-, 37- and 26-fold genome coverage, respectively, based on the estimated genome size of 485 Mb (Wang et al., 2011). The clean reads of each sample were mapped to the reference genome of the Chiifu cultivar. After filtering, a total of 358,141 SNPs and 54,500 InDels, which were distributed merely on ten chromosomes, were identified between the W-pool and the Y-pool. The Δ(SNP-index) of each position was calculated for

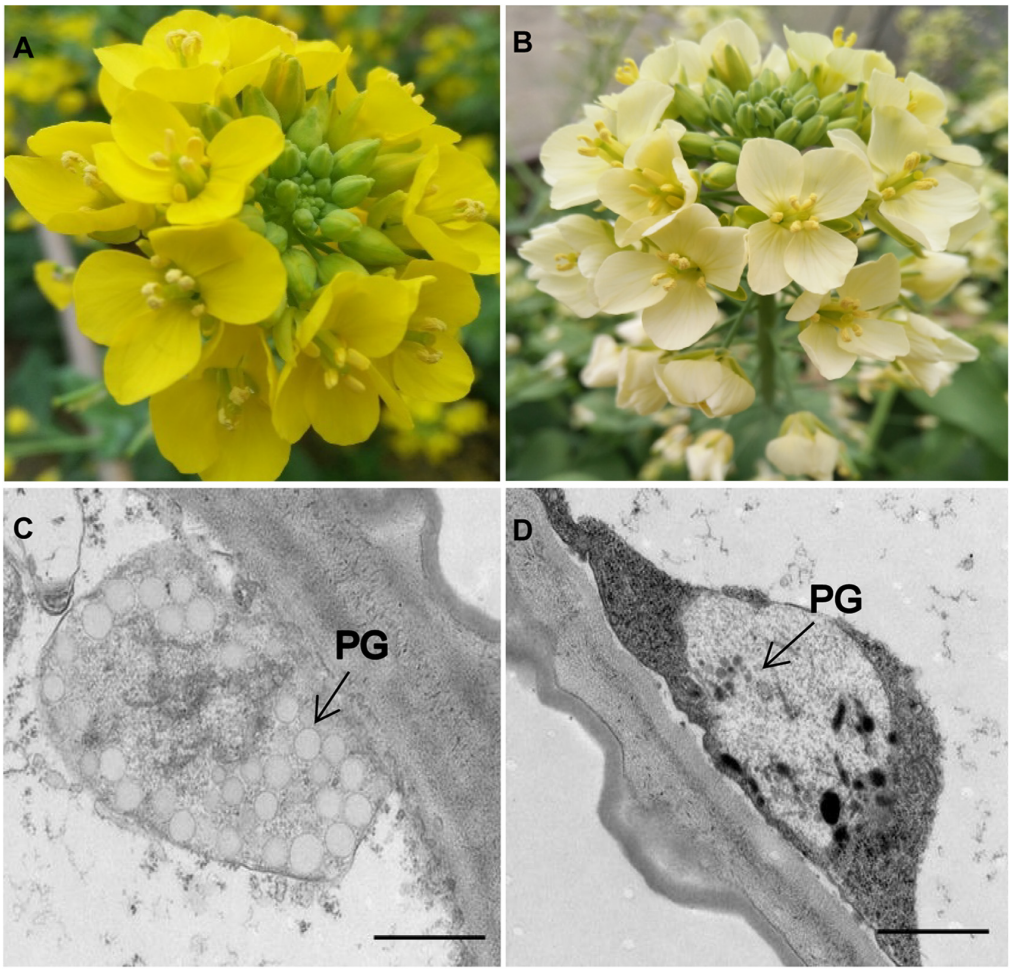


FIGURE 1 | Flower coloration and chromoplast ultrastructure in the parental lines. **(A)** Yellow-flowered Y641-87. **(B)** White-flowered Y640-288. **(C,D)** Plastid ultrastructure in Y641-87 and Y640-288. PG, plastoglobule; Scale bar, 1 μm in **(C,D)**.

TABLE 1 | Genetic analysis of the white flower trait in crosses between Y641-87 and Y640-288.

Population	Total	Yellow	White	Expected ratio	χ^2	$\chi^2_{0.05}$
P_1(Y641-87)	10	10	0	–	–	–
P_2(Y640-288)	10	0	10	–	–	–
F_1 (Y640-288 × Y641-87)	15	15	–	–	–	–
F_1 (Y641-87 × Y640-288)	15	15	–	–	–	–
F_2	200	142	58	3:1	1.71	3.84

sliding window analysis. According to the null hypothesis, a total of five adjacent regions on chromosome A02 (**Supplementary Figure 1** and **Supplementary Table 8**) exhibiting significant linkage disequilibrium were identified as the candidate region for white flower trait at a 95% significant level. These results were not consistent with the assumption that the white flower trait is controlled by a single recessive nuclear genetic locus. However, most of the genomic regions on other chromosomes exhibited a Δ(SNP-index) = 0. In theory, the SNP-index of the W- and Y-pools should be the same in the genomic regions that are not related to the phenotypic difference (SNP-index = 0.5), and Δ(SNP-index) should equal to 0 (Islam et al., 2016; Wang et al., 2017). Therefore, the most likely chromosome where *BrWF3* located was A02.

Fine Mapping of the *BrWF3* Gene

Based on BSA-seq analysis, 46 KASP markers previously available in our group (Yang et al., 2020) and 65 newly developed KASP markers, which were uniformly distributed across chromosome A02, were used to identify polymorphisms between the two parental lines (Y640-288 and Y641-87). The results showed that 29 KASP markers (**Supplementary Table 4**) exhibited polymorphism. These polymorphic markers were further genotyped in 135 F_2 plants for linkage analysis (**Supplementary Table 9**). The 135 F_2 plants is a subset from the 'whole' population containing 200 plants in **Table 1**. The results revealed no recombinant individuals between *BrWF3* and markers TXBH57, TXBH58, TXBH62 and TXBH30 and 3 recombinant individuals between

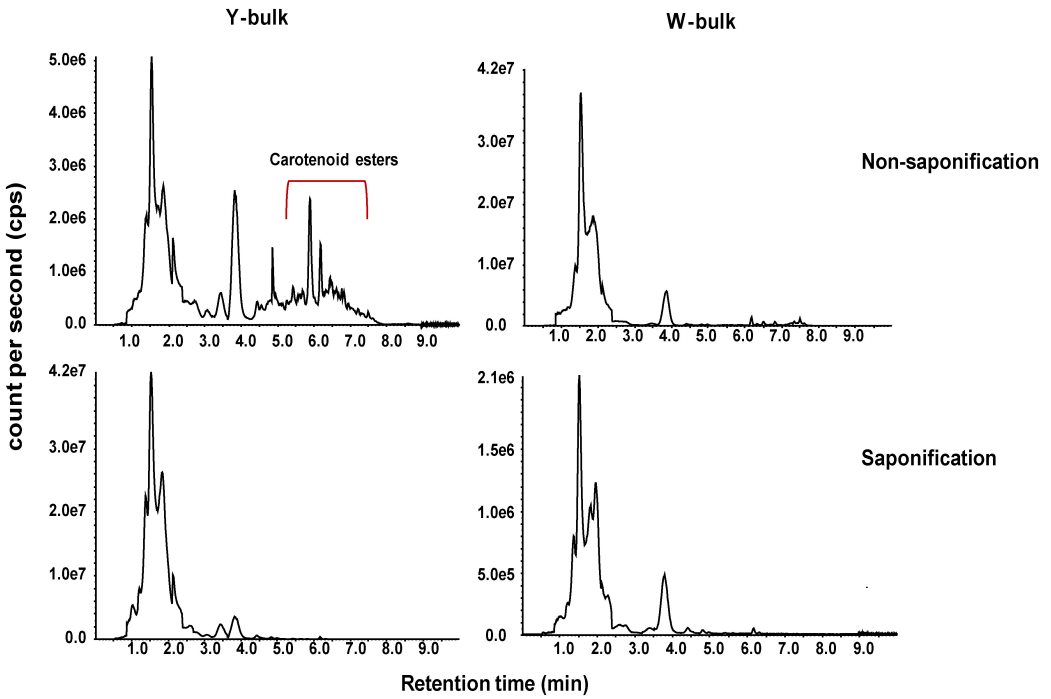


FIGURE 2 | Total ion chromatograms of carotenoids in yellow and white petals. Carotenoids extracts from mature yellow **(left column)** and white **(right column)** petals were subjected to LC-APCI-MS/MS analysis under non-saponification (up row) and saponification (down row) treatments.

TABLE 2 | Carotenoid contents and compositions in Y-bulk and W-bulk after saponification.

Compounds	Concentration [μ g/g DW]	
	Y-bulk	W-bulk
Carotenes		
(E/Z)-Phytoene	118.58 ± 13.91[a]	21.71 ± 1.75**
β-Carotene	5.68 ± 0.12	5.71 ± 0.38[no]
α-Carotene	0.94 ± 0.10	0.23 ± 0.03**
Xanthophylls		
Violaxanthin	745.70 ± 35.64	269.76 ± 27.74**
Lutein	386.27 ± 11.94	464.52 ± 27.45[no]
Neoxanthin	69.34 ± 0.99	18.40 ± 1.24**
Zeaxanthin	22.4 ± 0.80	10.51 ± 0.50**
Antheraxanthin	9.96 ± 0.56	8.81 ± 0.79[no]
α-Cryptoxanthin	2.98 ± 0.24	1.58 ± 0.07**
β-Cryptoxanthin	2.65 ± 0.14	2.09 ± 0.17[no]
Total content of carotenes	125.2	27.64**
Total content of xanthophylls	1239.31	775.67**
Total content of carotenoids	1364.51	803.31**

[a]*Mean ± SE (n = 3); ** Significant difference between Y-bulk and W-bulk within a component by Tukey's test (P < 0.01); *Significant difference at P < 0.05; [no]No significant difference at P < 0.05; n.d, not detected; DW: dry weight.*

TXBH64 and *BrWF3*. The genetic distances between the *BrWF3* locus and TXBH30 and TXBH64 were 0.7 and 2.0 cM, respectively (**Figure 3A**). The order of the markers in the genetic map is consistent with that in the physical map (**Figure 3A**).

To fine-map the *BrWF3* locus, we screened 596 white-flowered F_2 individuals using flanking markers (TXBH46 and TXBH31) and identified 52 recombinants. All the 52 recombinants were further genotyped using TXBH57, TXBH58, TXBH62, TXBH 30, TXBH64 and TXBH65, based on which 10 recombinants (type 2 and type 7) were identified (**Figure 3B**). Then, 21 new markers were developed, and seven of which exhibited polymorphism in the two parents (**Supplementary Table 4**). These seven new polymorphic markers were further used to screen all the 10 recombinants using KASP assay. The results delimited the *BrWF3* gene to a 105.6 kb interval between markers TXBH85 and TXBH84, each with one recombinant (type 3 and type 4) (**Figure 3B**). Three markers, SEQ14, TXBH76 and TXBH83, co-segregated with the *BrWF3* gene (**Figure 3B**).

Candidate Gene Analysis

Based on the fine mapping results for *BrWF3*, DNA sequences in the 105.6 kb interval were analyzed in the *Brassica* database[2] and comparative gene annotation in *Arabidopsis thaliana* was performed. As a result, 13 annotated or predicted genes were identified in the mapping region (**Table 3**). Four of these genes, *Bra032956*, *Bra032957*, *Bra032958*, and *Bra032959*, are homologs of *AT3G26840* (*PES2*) in *Arabidopsis thaliana*, which encodes a protein with phytyl ester synthesis and diacylglycerol acyltransferase activities and was previously reported to regulate carotenoid esterification (Zhang et al., 2018a,b; Kishimoto et al., 2020). Next, we examined the expression

[2]http://brassicadb.org/brad/

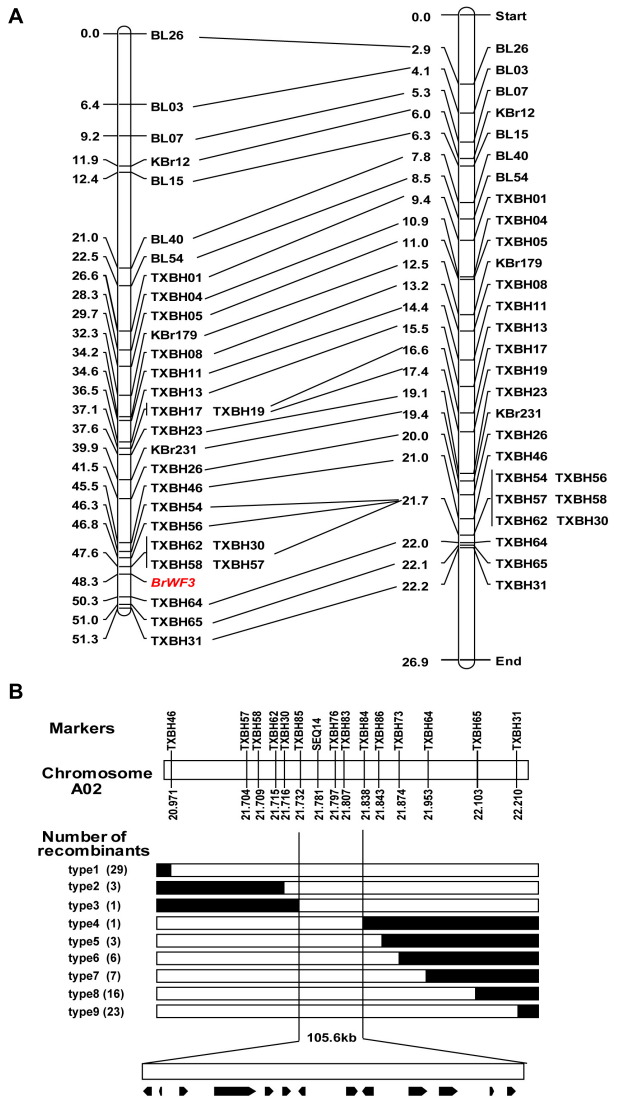


FIGURE 3 | Initial and fine mapping of the *BrWF3* gene in Chinese cabbage. **(A)** Initial mapping of *BrWF3*. Genetic map of *BrWF3* is on the left, with cM as the unit. The corresponding physical map (right, unit: Mb) are also shown. **(B)** Fine mapping of *BrWF3*. The *BrWF3* gene was delimited to an interval between TXBH85 and TXBH84, with an estimated length of 105.6 kb, and 13 genes were annotated in this region based on the reference genome sequence. The genetic structure of each recombinant type is depicted as white for homozygous white flower, black for heterozygous alleles, respectively. The number of each recombinant type is indicated in the brackets.

of these four candidate genes via RNA-seq and qRT-PCR analysis. RNA-seq analysis revealed that only *Bra032957* was differentially expressed among the four genes with the expression level decreasing approximately three fold in white petals compared with yellow petals (**Supplementary Table 10**). qRT-PCR analysis showed that the expression of *Bra032957* in petals was much higher than that of *Bra032956* and *Bra032959* and was significantly upregulated in yellow petals compared to white petals (**Figure 4A**). The results of qRT-PCR analysis and RNA-seq analysis were consistent. Subsequently, expression analysis in different tissues (root, stem, leaf, petal, sepal, stamen, and pistil) showed that *Bra032957* was predominantly expressed in petals (**Figure 4B**). Taken together, the results indicated that the *Bra032957* gene was the most likely candidate gene for *BrWF3*.

Sequence Analysis of *Bra032957* as a Candidate Gene for *BrWF3*

To characterize the sequence of the candidate genes in the white-flowered parental line Y640-288 and the yellow-flowered parental line Y641-87, the genomic sequence (gDNA) and coding sequence (CDS) of *Bra032957* were amplified and sequenced using specific primers (**Supplementary Figure 2** and **Supplementary Table 5**). The results showed that the *Bra032957* gene of Y641-87 was 5162 bp in length and contained 14 exons and 13 introns. The CDS of the *Bra032957* gene in Y641-87 was 2106 bp in length. Sequence alignment revealed a base deletion (G) at 477 bp of the CDS in the third exon in Y640-288 (**Supplementary Figure 3** and **Figure 5A**). The SNP deletion caused a frameshift mutation in the *Bra032957* gene and a premature stop codon in 168 a.a. residues (**Figure 5A**). Conserved domain analysis in NCBI showed that the *Bra032957* gene contained an α/β hydrolase-fold (amino acids 123-380) and a lysophospholipid acyltransferases (LPLAT) domain (amino acids 430-666), which can transfer acyl groups to acceptors, such as glycerol 3-phosphate (**Figure 5B**). The SNP deletion mutation caused a loss of the two conserved domains and ultimately caused the loss of function of the BrWF3 protein.

Based on the identified SNP deletion, we designed a KASP marker TXBH83 to screen the other three F_2 populations (Y640-288 × SD369-F_2, Y640-288 × Chiifu-F_2, Y66-83 × R16-11-F_2), including a total of 282 individuals. The results showed that TXBH83 co-segregated with the flower color phenotype (**Figures 6A-C** and **Supplementary Table 11**). Furthermore, 10 white-flowered and 10 yellow-flowered DH lines were genotyped for TXBH83, which also showed a 100% consistency between the flower color phenotype and genotype (**Figure 6D** and **Supplementary Table 11**). Overall, these findings suggest that the *Bra032957* gene is the most promising candidate gene for the white flower gene *BrWF3* in Chinese cabbage.

Transcriptome Analysis in Yellow and White Petals

To identify the gene regulatory networks involved in petal coloration, we performed comparative transcriptome analysis between W-bulk and Y-bulk. Approximately 300.8 million raw reads were generated for the six cDNA libraries, ranging from 43.1 to 57.4 Gb reads per library (**Supplementary Table 12**). All the raw reads were deposited in the NCBI Short Read Archive (SRA) database under accession number PRJNA682761. Among the clean reads, 85.2-87.46% were uniquely mapped to the reference genome (Wang et al., 2011). The Pearson correlation coefficients among the three replicates of each petal pool ranged from 0.98 to 0.99, indicating a high consistency among the three

TABLE 3 | Annotated genes in the candidate interval of the *BrWF3* locus.

Gene Name	Gene Position on A02	Arabidopsis thaliana homolog	Gene function
Bra032947	21730896-21732491	*AT3G26744*	MYC-like bHLH transcription factor
Bra032948	21734782-21735069	*AT3G26750*	HECT-like ubiquitin-conjugating enzyme (E2)-binding protein
Bra032949	21738466-21739774	*AT3G26760*	NAD(P)-binding Rossmann-fold superfamily protein
Bra032950	21743953-21753155	*AT3G17090*	Protein phosphatase 2C family protein;
Bra032951	21755514-21758199	*AT3G26770*	NAD(P)-binding Rossmann-fold superfamily protein
Bra032952	21762594-21764296	*AT3G26780*	Histidine phosphatase superfamily,clade-1
Bra032953	21768147-21770040	*AT3G26790*	B3 domain-containing transcription factor
Bra032954	21782268-21784649	*AT3G26810*	Auxin F box protein; AFB2
Bra032955	21786082-21787614	*AT5G14030*	translocon-associated protein beta(TRAPB) family protein
Bra032956	21797018-21801228	*AT3G26840*	Diacylglycerol acyltransferase; PES2
Bra032957	21805268-21810180	*AT3G26840*	Diacylglycerol acyltransferase; PES2
Bra032958	21833406-21833807	*AT3G26840*	Diacylglycerol acyltransferase; PES2
Bra032959	21836692-21839336	*AT3G26840*	Diacylglycerol acyltransferase; PES2

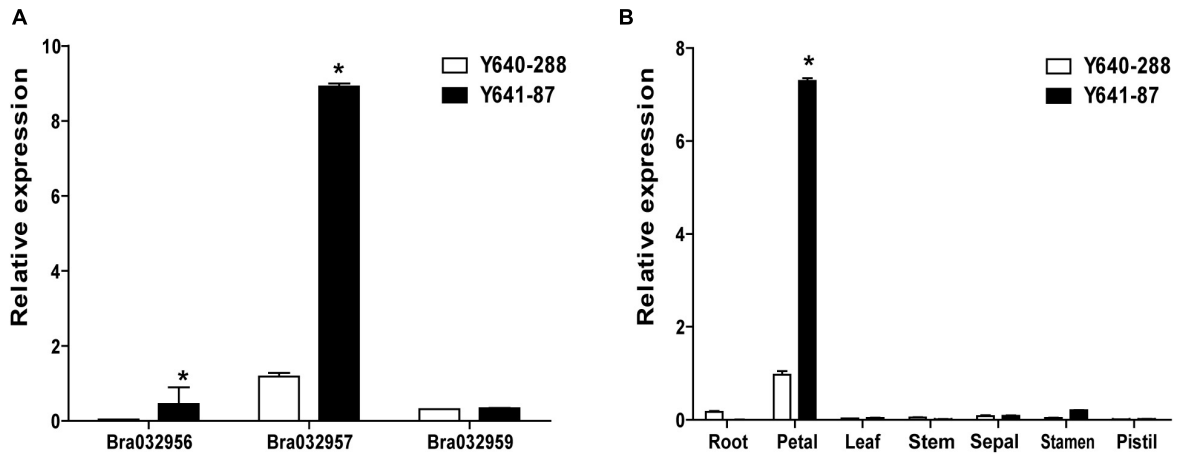


FIGURE 4 | Gene expression data analyses. **(A)** Relative expression of *Bra032956*, *Bra032957,* and *Bra032959* in petals of the two parents. **(B)** Relative expression of *Bra032957* in different tissues of the two parents. Error bars represent the standard error of three biological replicates, and asterisks indicate significant differences (*t*-test, $p < 0.05$).

replicates (**Supplementary Figure 4**). In total, we identified 6,009 differentially expressed genes (DEGs) between the W-bulk and Y-bulk, among which 2,913 genes were up-regulated and 3,816 were down-regulated in the Y-bulk compared with the W-bulk. Pathway enrichment analysis based on the KEGG database revealed that carotenoid biosynthesis was the most significantly enriched pathway (**Figure 7A**). In the Y-bulk, genes involved in carotenoid biosynthesis (**Figure 7B**), such as *PSY (Bra006391* and *Bra008569*), *PDS* (*Bra010751*), *BCH1* (*Bra013912*) and *ZEP* (*Bra037130*), were significantly up-regulated. Genes involved in carotenoid degradation (**Figure 7B**), such as *NCED3* (*Bra027336* and *Bra001552*) and *NCED4* (*Bra013378*), were also upregulated. Pathways of linoleic acid metabolism, alpha-linolenic acid metabolism, glycerophospholipid metabolism and arachidonic acid metabolism were also enriched, and most of the genes in these pathways were downregulated (**Figure 7A** and **Supplementary Table 12**). For example, genes encoding glycerophosphodiester phosphodiesterase (*Bra040704*, *Bra027481*, *Bra035967*, *Bra020395*, *Bra002676,* and *Bra006785*), phospholipase A (*Bra015531* and *Bra010327*), phosphatidate phosphatase (*Bra029774*) were downregulated (**Supplementary Figure 5** and **Supplementary Table 13**). However, genes participating in fatty acid elongation, such as 3-ketoacyl-CoA synthase (*Bra011936*, *Bra004513*, *Bra024749,* and *Bra004034*) and very-long-chain enoyl-CoA reductase (*Bra008657*), were significantly upregulated (**Supplementary Figure 5** and **Supplementary Table 13**). These results suggested that saturated but not unsaturated fatty acids might be the main acyl group donors for esterification of xanthophylls. FIBRILLIN (FBN) and ACTIVITY OF BC1 COMPLEX KINASE (ABC1K) are the most abundant proteins in PGs (van Wijk and Kessler, 2017). In this study, *FBN1b* (*Bra013602*) and *ABC1K8* (*Bra024339*) were significantly upregulated in Y-bulk (**Figure 7C**). Furthermore, *FBN1b* is highly expressed. The average fragments per kilobase million (FPKM) value of *FBN1b* was as high as 8383 in Y-bulk, 3148 in the W-bulk. *VITAMIN*

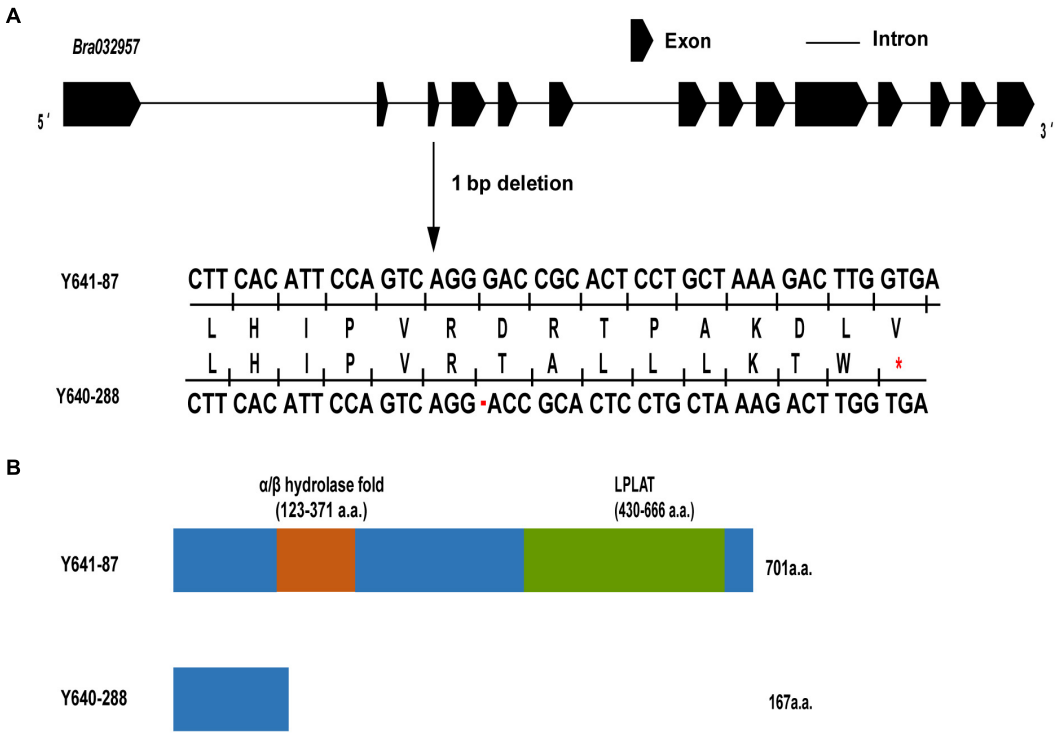


FIGURE 5 | Gene structure, amino acid analysis and predicted protein structure of *BrWF3*. **(A)** *BrWF3* includes 14 exons and 13 introns. A single nucleotide deletion (G) in the third exon of white flower plants results in a premature stop codon (indicated with red asterisks) due to a frameshift mutation. **(B)** Protein structures of BrWF3 in the two parents. The SNP deletion in Y640-288 causes the loss of two conserved protein domains of BrWF3.

E DEFICIENT 1 (*VTE1*), which encodes a key enzyme in tocopherol biosynthesis, was upregulated in Y-bulk but with no significant difference (**Figure 7C** and **Supplementary Table 13**). Lipoxygenase (LOX) are proteins recruited to chloroplast PGs and participated in jasmonate biosynthesis (van Wijk and Kessler, 2017). Genes encoding LOXs were all downregulated in yellow petals (**Figure 7C** and **Supplementary Table 13**), indicating that LOX proteins were not indispensible for PG development and formation in chromoplasts.

DISCUSSION

BSA-seq has been widely deployed for mapping agronomical traits in crops (Wang et al., 2017; Lee et al., 2020; Liang et al., 2020). In most cases, the genes controlling the target agronomic traits are located on the candidate regions detected by BSA-seq analysis. However, in one study, BSA-seq and traditional linkage analyses identified two different major loci for the purple leaf trait in *Brassica rapa*, one located on chromosome A07 and the other on A09 (Zhang X. et al., 2020). In our study, genetic analysis showed that the white flower trait in Y640-288 is controlled by a single recessive gene. However, our BSA-seq analysis identified five adjacent regions on A02, rather than only one sharp peak as noted in previous studies, and the results were not consistent with the genetic analysis. This phenomenon has also occurred in other studies from both our laboratory (unpublished data) and another laboratory (the Chinese cabbage research group at the Northwest A&F University). The two parental lines, Y640-288 and Y641-87, are both over-lapped head-type Chinese cabbage lines. While the reference material Chiifu-401-42 is a closed head-type Chinese cabbage line. Many significant structural variations were detected between the parental lines and the reference genome through single molecule real-time (SMRT) sequencing (**Supplementary Figure 6**). We suspected that the great structural variation between the parents and the reference genome might be responsible for the above discrepancy and reduce the efficiency of BSA-seq, which will be discussed in detail in a future study. Although the BSA-seq result was not perfect enough, the Δ(SNP-index) value for the other nine chromosomes were all close to 0, allowing the rapid mapping of the white flower gene on A02. The identified gene was named *BrWF3* because it was quite different from the other two genes reported previously (*Brwf1* on A01 and *Brwf2* on A09) (Zhang N. et al., 2020).

The present study successfully fine mapped the *BrWF3* gene to a physical interval of 105.6 kb. Functional annotation analysis of the 13 genes in the candidate region revealed that four genes, *Bra032956*, *Bra032957*, *Bra032958,* and *Bra032959*, which are homologous to *PES2* in *Arabidopsis*, might be candidate genes for *BrWF3*. In *Arabidopsis*, the *PES1* and *PES2* genes encode proteins that use medium-chain fatty acid-derived acyls to esterify phytyl released during chlorotic conditions. The resultant fatty acid phytyl esters accumulate in chloroplast PGs (Lippold et al., 2012). The *PYP1* (*Arabidopsis PES1* homolog)

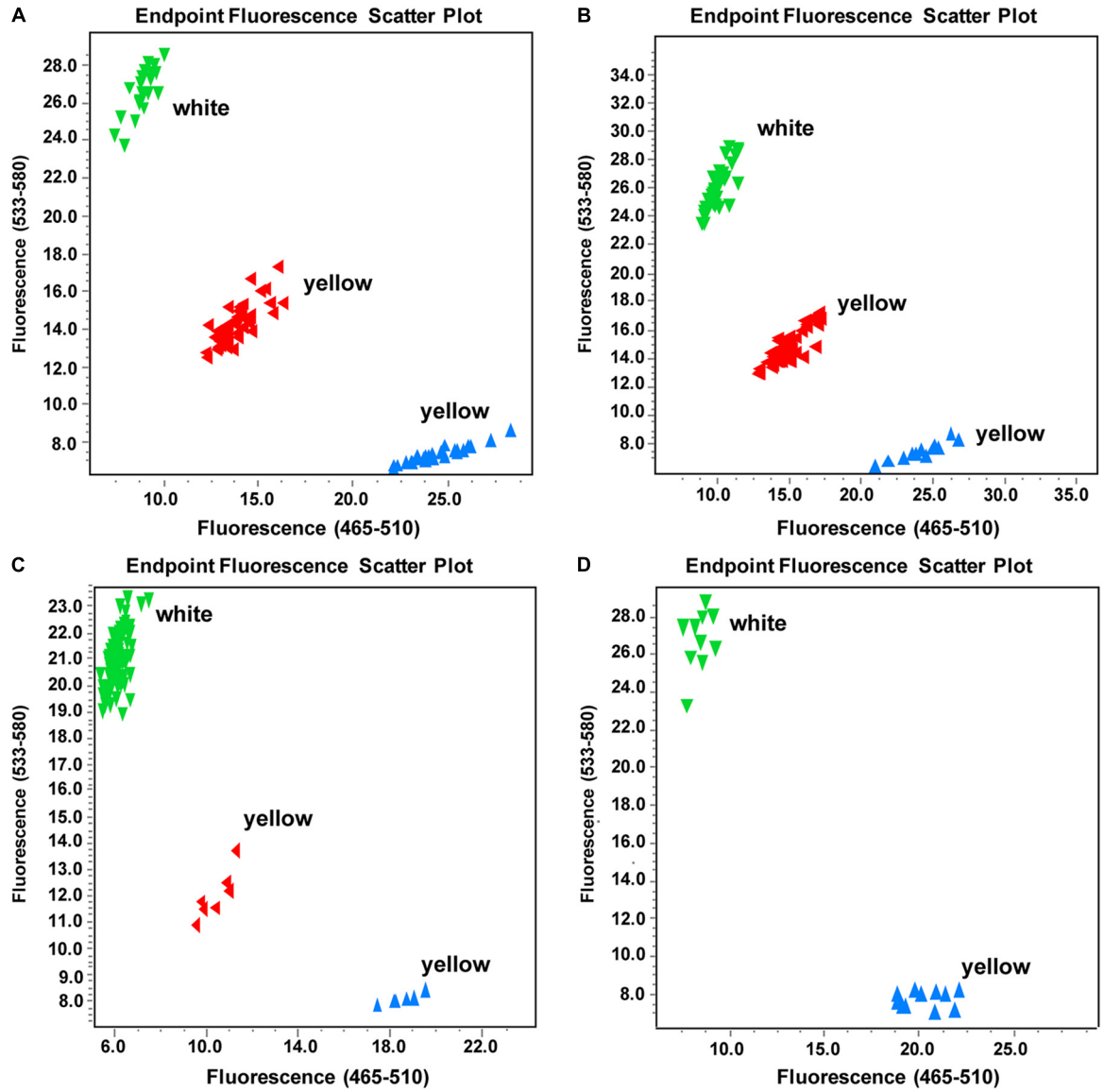


FIGURE 6 | Genotyping results of marker TXBH83 in part individuals of different populations. **(A–D)** Genotyping results of TXBH83 in the (Y640-288 × Chiifu)-F_2, (Y640-288 × SD369)-F_2, and (Y66-83 × R16-11)-F_2 populations and a natural population, respectively. More white-flowered plants were intentionally selected. The genotypes corresponding to those of Y640-288 are clustered to the Y axis, those matching Y641-87 genotypes are clustered to the X axis, and the heterozygous genotypes are clustered to the diagonal line. TXBH83 is totally co-segregated with the flower color phenotype.

gene in tomato (Ariizumi et al., 2014) and the *Bjpc1* and *Bjpc2* (*Arabidopsis PES2* homolog) genes in *B.juncea* are all responsible for flower color changes (Zhang et al., 2018a,b). Furthermore, RNA-seq and qRT-PCR analysis revealed that only *Bra032957* was significantly and differentially expressed, showing considerably increased expression in yellow petals. We also examined the sequence variation between white and yellow petals. No sequence variation was found in the *Bra032958* gene. However, *Bra032956*, *Bra032957* and *Bra032959* each possessed one SNP variation. The SNPs in *Bra032956* (marker TXBH76) and *Bra032959* (marker TXBH84) occurred in intron regions, and only the SNP of *Bra032957* was located an exon. Additionally, we developed KASP markers for these three SNPs, namely, TXBH76 in *Bra032956*, TXBH83 in *Bra032957* and TXBH84 in *Bra032959*. The TXBH84 marker showed one recombinant according to fine mapping. Thus, the *Bra032959* gene could be excluded as a candidate. The TXBH76 and TXBH83 markers cosegregated with the phenotype during fine mapping. However, the TXBH76 marker could not be used to differentiate the yellow and white flower phenotypes in another F_2 population (Y66-83 × R16-11) (**Supplementary Figure 7**). Accordingly, the possibility of *Bra032956* being the candidate gene was also eliminated. Moreover, we cloned the gDNA and cDNA sequences of *Bra032957* in Y641-87 and

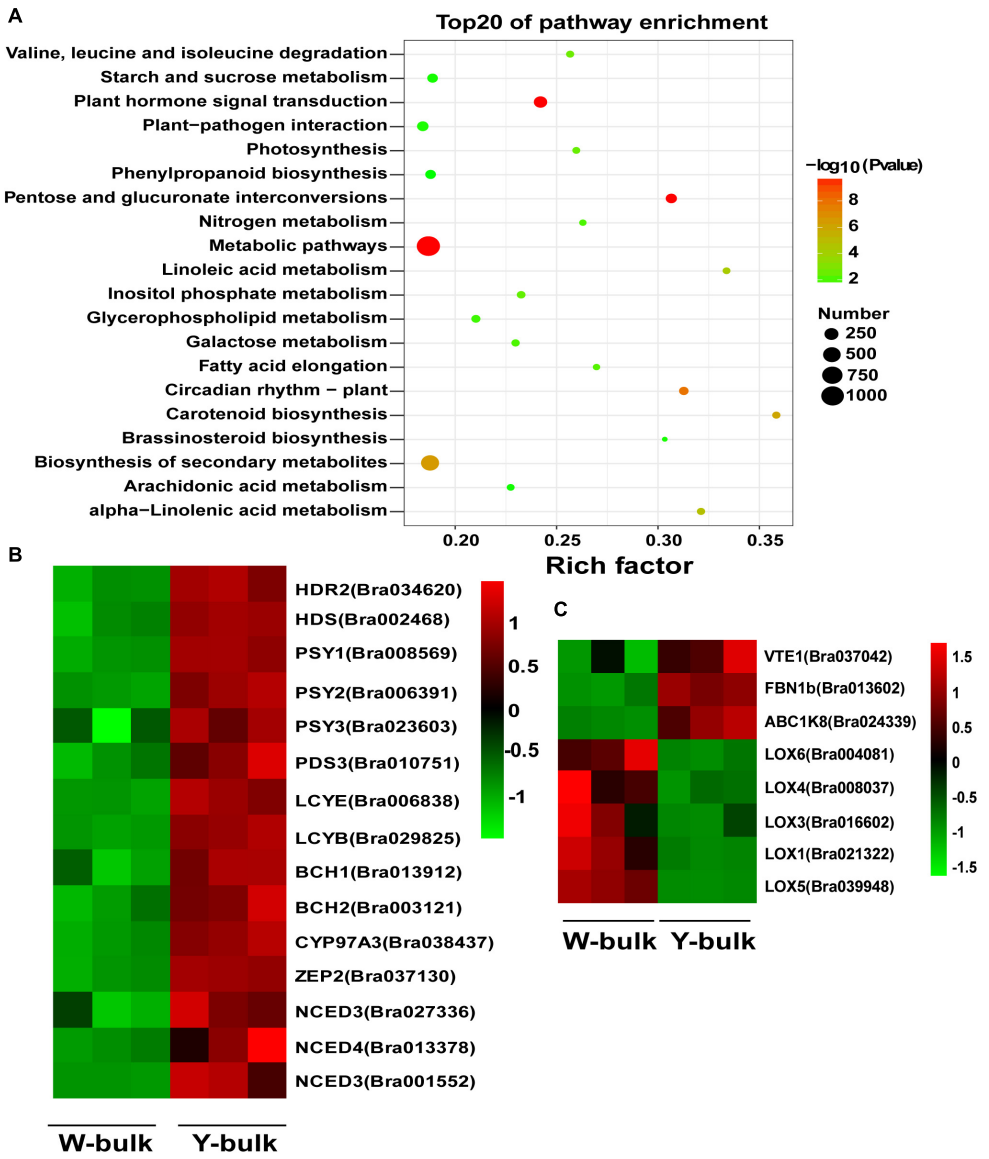


FIGURE 7 | Transcriptome analysis in yellow and white petals. **(A)** Scatter plot of top 20 enriched KEGG pathways. Rich factor is the ratio of the DEG number to the background number in a certain pathway. The size of the dots represents the number of genes, and the color of the dots represents the range of the -log10(*p*-value). **(B)** Differentially expressed genes involved in carotenoid biosynthesis and degradation. **(C)** Differentially expressed genes associated with proteins in PGs. The heatmap colors are shown in log10(FPKM). Three biological replicates of the W-bulk and Y-bulk are shown.

Y640-288. Sequence alignment revealed an SNP deletion in the third exon in the white flower parent Y640-288, which introduced a premature stop codon and caused enzymatic inactivity (**Supplementary Figure 3**). Taken together, the results indicated that the *Bra032957* gene was most likely responsible for the white flower trait in Chinese cabbage.

The KASP genotyping assay is one of the most efficient and cost-effective systems for SNP and small InDel genotyping (Semagn et al., 2013; Yang et al., 2020). In *B.juncea*, the white flower trait is collectively controlled by two recessive genes (*Bjpc1* and *Bjpc2*) (Zhang et al., 2018a,b). The *Bjpc1* gene, which is located on A02, is homologous to *Bra032956* in *B.rapa* (Zhang et al., 2018b). *Bjpc2* lies on B04 and is homologous to *Bra032957* in *B.rapa*. In our study, the *Bra032957* gene, which is located on A02, was the most promising candidate gene for the white flower trait in Chinese cabbage, suggesting a similar mechanism of flower color pigmentation in these two species. However, none of the closely linked and cosegregated markers of *Bjpc1* and *Bjpc2* showed polymorphisms in our F_2 population (Y640-288 × Y641-87) (**Supplementary Figures 8, 9**), suggesting divergence between *B.juncea* and *B.rapa*. For the SNP deletion in *BrWF3*, we developed the KASP marker TXBH83, which completely cosegregated with the flower color phenotype in three other F_2 populations and a small natural population. This marker

can be used as a functional marker for MAS breeding and the assessment of genetic resources for developing new ornamental varieties with visual appeal, which has profound significance.

Using LC-APCI-MS/MS analysis with saponification, we observed that violaxanthin and lutein were the two main carotenoids in petals, accounting for approximately 83 and 91% of the total carotenoids in the Y-bulk and the W-bulk, respectively. These results were consistent with another study conducted in Chinese cabbage (Zhang N. et al., 2020) but differed from those of studies in tomato and *B.juncea* (Ariizumi et al., 2014; Zhang et al., 2018a,b), which suggested that the quantities and composition of xanthophylls in yellow flowers display considerable diversity among different species (Ohmiya, 2011). Notably, in the saponified extracts, the violaxanthin content in the Y-bulk was approximately 2.76 times higher than that in the W-bulk, and the total carotenoids content in the Y-bulk was about 1.70 times higher than that in the W-bulk, but the lutein content between these two bulks was not significantly different (**Table 2**). Thus, we hypothesized that the lower pigmentation in white flowers was due to the decreased carotenoid content, which is particularly related to the reduced violaxanthin content (but not lutein content). Furthermore, comparison of the carotenoid profile in direct and saponified extracts revealed increased proportion xanthophyll esters in yellow petals. Taken together, we proposed that the *BrWF3* gene played an important role in the production of xanthophyll esters (particularly violaxanthin esters) for yellow color pigmentation in Chinese cabbage and the lower pigmentation in white flowers was due to lower xanthophyll esters. Interestingly, among the direct extracts, in addition to the carotenoid esters detected in the Y-bulk, we also detected a few carotenoid esters in the W-bulk. While in tomato, carotenoid esters cannot be detected in pale yellow flowers (Ariizumi et al., 2014). The potential reasons for this difference might be associated with the whole-genome triplication of *B.rapa* (Wang et al., 2011). Although *BrWF3* encodes inactive proteins in white petals, other paralogs might express several active proteins participating in carotenoid esterification, which was confirmed by the expression data in our study (**Figure 4**).

In colored organs, such as petals and fruit, carotenoids accumulate mainly in chromoplasts, particularly in carotenoid-lipoprotein sequestering structures (i.e., PGs and fibrils) (Li and Yuan, 2013; Li et al., 2016; Yuan et al., 2015). These carotenoid-lipoprotein sequestering structures enhance the sink strength of chromoplasts (Yuan et al., 2015; Li et al., 2016). Deruere et al. (1994) proposed an architectural model for fibrils or PGs in which carotenoids were located in the inner core protected by polar lipids and carotenoid-associated proteins, the carotenoid core interacted with the acyl residues of the polar lipids, and the polar lipids then interacted with carotenoid-associated proteins via polar head groups. Carotenoids, polar lipids and carotenoid-associated proteins are three indispensable components of carotenoid-lipoprotein-sequestering structures. Many studies have revealed that genes encoding carotenoid-associated proteins are positively associated with carotenoid accumulation and that the mutations in these genes hamper PG formation (Vishnevetsky et al., 1996; Pozueta-Romero et al., 1997). In our study, TEM analysis showed that numerous fully developed PGs could be observed in yellow petals, whereas only a few irregular and small PGs could be observed in white petals, which was consistent with results obtained in *B.juncea* (Zhang et al., 2018a,b) and tomato (Ariizumi et al., 2014). We suspect that the *BrWF3* gene from our study, *PYP1* in tomato (Ariizumi et al., 2014), and *Bjpc2* in *B.juncea* (Zhang et al., 2018a), encoding proteins with phytyl ester biosynthesis and diacylglycerol acyltransferase activities, construct the interaction between carotenoids and the polar lipids by transferring the acyl group from polar lipids to the hydroxy (-OH) group of xanthophylls. An SNP deletion in *BrWF3* causes loss of protein function, thereby disturbing the connection between carotenoids and the polar lipids and further hampering the assembly and formation of PGs. As expected, comparative transcriptome analysis between the Y-bulk and W-bulk showed that the genes involved in three indispensable components of PGs, carotenoids, polar lipids and carotenoid-associated proteins, were downregulated in white petals, in which the development and formation of PGs were hampered.

CONCLUSION

In the present study, we delimited the *BrWF3* gene responsible for the white flower trait in Chinese cabbage. BSA-seq and linkage analysis via KASP assays were employed to fine-map the *BrWF3* gene to an interval of 105.6 kb. Functional annotation analysis, expression analysis and sequence variation analysis revealed that *Bra032957*, which encodes a protein with phytyl ester synthesis and diacylglycerol acyltransferase activities, was the most likely candidate gene for *BrWF3*. *BrWF3* participated in the production of xanthophyll esters (particularly violaxanthin esters) and the formation of PGs. An SNP deletion in the third exon of *BrWF3* caused the loss of protein function and interfered with the formation of PGs, which subsequently reduced the activity of carotenoid metabolism and the content of carotenoids. Furthermore, we developed and validated the functional marker TXBH83 for *BrWF3*. These results not only provide valuable information for MAS breeding but also provide a significant contribution to research on the molecular mechanism underlying flower color pigmentation.

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 in the article/ **Supplementary Material**.

AUTHOR CONTRIBUTIONS

SY conceptualized and designed the experiments and drafted the manuscript. XT and ZW performed the experiments and analyzed the data. BT, YY and X-WZ directed the whole study and provided the funding resource. XW, YZ, HS, and XZ

participated in drafting the article and revising it critically. All authors contributed to the article and approved the submitted version.

FUNDING

This work was financially supported by Zhongyuan Scholar Program (202101510003), the Modern Agro-Industry Technology Research System (CARS-23-G-16), the National Science Foundation of China (31872945), and Sci-Tech Innovation Team of Henan Academy of Agricultural Sciences (2021TD06).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2021.646222/full#supplementary-material

REFERENCES

Abe, A., Kosugi, S., Yoshida, K., Natsume, S., Takagi, H., Kanzaki, H., et al. (2012). Genome sequencing reveals agronomically important loci in rice using MutMap. *Nat. Biotechnol.* 30, 174–178. doi: 10.1038/nbt.2095

Ariizumi, T., Hatakeyama, K., Hinata, K., Inatsugi, R., Nishida, I., Sato, S., et al. (2004). Disruption of the novel plant protein NEF1 affects lipid accumulation in the plastids of the tapetum and exine formation of pollen, resulting in male sterility in *Arabidopsis thaliana*. *Plant J.* 39, 170–181. doi: 10.1111/j.1365-313x.2004.02118.x

Ariizumi, T., Kishimoto, S., Kakami, R., Maoka, T., Hirakawa, H., Suzuki, Y., et al. (2014). Identification of the carotenoid modifying gene *PALE YELLOW PETAL 1* as an essential factor in xanthophyll esterification and yellow flower pigmentation in tomato (*Solanum lycopersicum*). *Plant J.* 79, 453–465. doi: 10.1111/tpj.12570

Benjamini, Y., and Hochberg, Y. (2000). On the adaptive control of the false discovery rate in multiple testing with independent statistics. *J. Educ. Behav. Stat.* 25, 60–83. doi: 10.2307/1165312

DellaPenna, D., and Pogson, B. J. (2006). Vitamin synthesis in plants: tocopherols and carotenoids. *Annu. Rev. Plant Biol.* 57, 711–738. doi: 10.1146/annurev.arplant.56.032604.144301

Deruere, J., Romer, S., D'Harlingue, A., Backhaus, R. A., Kuntz, M., and Camara, B. (1994). Fibril assembly and carotenoid overaccumulation in chromoplasts: a model for supramolecular lipoprotein structures. *Plant Cell* 6, 119–133. doi: 10.2307/3869680

Ding, B. Y., Niu, J., Shang, F., Yang, L., Chang, T. Y., and Wang, J. J. (2019). Characterization of the geranylgeranyl diphosphate synthase gene in *Acyrthosiphon pisum* (Hemiptera: Aphididae) and its association with carotenoid biosynthesis. *Front. Physiol.* 10:1398.

Fujimoto, R., Sasaki, T., and Nishio, T. (2006). Characterization of DNA methyltransferase genes in *Brassica rapa*. *Genes Genet. Syst.* 81, 235–242. doi: 10.1266/ggs.81.235

Grotewold, E. (2006). The genetics and biochemistry of floral pigments. *Annu. Rev. Plant Biol.* 57, 761–780. doi: 10.1146/annurev.arplant.57.032905.105248

Han, F. Q., Cui, H., Zhang, B., Liu, X., Yang, L., Zhuang, M., et al. (2019). Map-based cloning and characterization of *BoCCD4*, a gene responsible for white/yellow petal color in *B. oleracea*. *BMC Genomics* 20:242. doi: 10.1186/s12864-019-5596-2

Inbaraj, B. S., Lu, H., Hung, C. F., Wu, W. B., Lin, C. L., and Chen, B. H. (2008). Determination of carotenoids and their esters in fruits of *Lycium barbarum* Linnaeus by HPLC-DAD-APCI-MS. *J. Pharm. Biomed. Anal.* 47, 812–818. doi: 10.1016/j.jpba.2008.04.001

Islam, M. S., Zeng, L., Thyssen, G. N., Delhom, C. D., Kim, H. J., Li, P., et al. (2016). Mapping by sequencing in cotton (*Gossypium hirsutum*) line MD52ne identified candidate genes for fiber strength and its related quality attributes. *Theor. Appl. Genet.* 129, 1071–1086. doi: 10.1007/s00122-016-2684-4

Kevan, P. G., and Baker, H. G. (1983). Insects as flower visitors and pollinators. *Annu. Rev. Entomol.* 28, 407–453. doi: 10.1146/annurev.en.28.010183.002203

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

Kishimoto, S., Oda-Yamamizo, C., and Ohmiya, A. (2019). Comparison of petunia and calibrachoa in carotenoid pigmentation of corollas. *Breed. Sci.* 69, 117–126. doi: 10.1270/jsbbs.18130

Kishimoto, S., Oda-Yamamizo, C., and Ohmiya, A. (2020). Heterologous expression of xanthophyll esterase genes affects carotenoid accumulation in petunia corollas. *Sci. Rep.* 10:1299.

Koes, R. E., Quattrocchio, F., and Mol, J. N. M. (1994). The flavonoid biosynthetic pathway in plants: function and evolution. *BioEssays* 16, 123–132. doi: 10.1002/bies.950160209

Kosambi, D. D. (1944). The estimation of map distances from recombination values. *Ann. Eugen.* 12, 172–175. doi: 10.1111/j.1469-1809.1943.tb02321.x

Lee, S. B., Kim, J. E., Kim, H. T., Lee, G. M., Kim, B. S., and Lee, J. M. (2020). Genetic mapping of the *c1* locus by GBS-based BSA-seq revealed *Pseudo-Response Regulator 2* as a candidate gene controlling pepper fruit color. *Theor. Appl. Genet.* 133, 1897–1910. doi: 10.1007/s00122-020-03565-5

Li, H., and Durbin, R. (2010). Fast and accurate long-read alignment with Burrows-Wheeler transform. *Bioinformatics* 26, 589–595. doi: 10.1093/bioinformatics/btp698

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., et al. (2009). The sequence alignment/map format and SAMtools. *Bioinformatics* 25, 2078–2079. doi: 10.1093/bioinformatics/btp352

Li, L., and Yuan, H. (2013). Chromoplast biogenesis and carotenoid accumulation. *Arch. Biochem. Biophys.* 539, 102–109. doi: 10.1016/j.abb.2013.07.002

Li, L., Yuan, H., Zeng, Y., and Xu, Q. (2016). Plastids and carotenoid accumulation. *Subcell. Biochem.* 79, 273–293. doi: 10.1007/978-3-319-39126-7_10

Liang, T. M., Chi, W. C., Huang, L. K., Qu, M. Y., Zhang, S. B., Chen, Z. Q., et al. (2020). Bulked segregant analysis coupled with whole-genome sequencing (BSA-Seq) mapping identifies a novel *i21* haplotype conferring basal resistance to rice blast disease. *Int. J. Mol. Sci.* 21:2162. doi: 10.3390/ijms21062162

Lippold, F., vom Dorp, K., Abraham, M., Holzl, G., Wewer, V., Yilmaz, J. L., et al. (2012). Fatty acid phytyl ester synthesis in chloroplasts of *Arabidopsis*. *Plant Cell* 24, 2001–2014. doi: 10.1105/tpc.112.095588

Liu, L., Guo, W., Zhu, X., and Zhang, T. (2003). Inheritance and fine mapping of fertility restoration for cytoplasmic male sterility in *Gossypium hirsutum* L. *Theor. Appl. Genet.* 106, 461–469. doi: 10.1007/s00122-002-1084-0

Liu, Y., Lv, J., Liu, Z., Wang, J., Yang, B., Chen, W., et al. (2020). Integrative analysis of metabolome and transcriptome reveals the mechanism of color formation in pepper fruit (*Capsicum annuum* L.). *Food Chem.* 306:125629. doi: 10.1016/j.foodchem.2019.125629

Livak, K. J., and Schmittgen, T. D. (2001). Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-\Delta\Delta C_T}$ method. *Methods* 25, 402–408. doi: 10.1006/meth.2001.1262

Lopez, A. B., Van Eck, J., Conlin, B. J., Paolillo, D. J., O'Neill, J., and Li, L. (2008). Effect of the cauliflower *Or* transgene on carotenoid accumulation and chromoplast formation in transgenic potato tubers. *J. Exp. Bot.* 59, 213–223. doi: 10.1093/jxb/erm299

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.

Lu, S., Van Eck, J., Zhou, X. J., Lopez, A. B., O'Halloran, D. M., Cosman, K. M., et al. (2006). The cauliflower *Or* gene encodes a DnaJ Cysteine-rich domain-containing protein that mediates high levels of β-carotene accumulation. *Plant Cell* 18, 3594–3605. doi: 10.1105/tpc.106.046417

Nisar, N., Li, L., Lu, S., Khin, N. C., and Pogson, B. J. (2015). Carotenoid metabolism in plants. *Mol. Plant* 8, 68–82. doi: 10.1016/j.molp.2014.12.007

Ohmiya, A. (2011). Diversity of carotenoid composition in flower petals. *Jpn. Agric. Res. Q.* 45, 163–171. doi: 10.6090/jarq.45.163

Pozueta-Romero, J., Rafia, F., Houlne, G., Cheniclet, C., Carde, J. P., Schantz, M. L., et al. (1997). A ubiquitous plant housekeeping gene, *PAP*, encodes a major protein component of bell pepper chromoplasts. *Plant Physiol.* 115, 1185–1194. doi: 10.1104/pp.115.3.1185

Rahman, M. H. (2001). Inheritance of petal colour and its independent segregation from seed colour in *Brassica rapa*. *Plant Breed.* 120, 197–200. doi: 10.1046/j.1439-0523.2001.00607.x

Semagn, K., Babu, R., Hearne, S., and Olsen, M. (2013). Single nucleotide polymorphism genotyping using Kompetitive Allele Specific PCR (KASP): overview of the technology and its application in crop improvement. *Mol. Breed.* 33, 1–14. doi: 10.1007/s11032-013-9917-x

Simkin, A. J., Gaffe, J., Alcaraz, J. P., Carde, J. P., Bramley, P. M., Fraser, P. D., et al. (2007). Fibrillin influence on plastid ultrastructure and pigment content in tomato fruit. *Phytochemistry* 68, 1545–1556. doi: 10.1016/j.phytochem.2007.03.014

Singh, V. K., Mangalam, A. K., Dwivedi, S., and Naik, S. (1998). Primer premier: program for design of degenerate primers from a protein sequence. *Biotechniques* 24, 318–319. doi: 10.2144/98242pf02

Su, T. B., Yu, S. C., Zhang, J., Lan, W. F., Yu, Y. J., Zhang, D. S., et al. (2014). Loss of function of the carotenoid isomerase gene *BrCRTISO* confers orange color to the inner leaves of Chinese cabbage (*Brassica rapa* L. ssp. *pekinensis*). *Plant Mol. Biol. Rep.* 33, 648–659. doi: 10.1007/s11105-014-0779-0

Takada, S., Akter, A., Itabashi, E., Nishida, N., Shea, D. J., Miyaji, N., et al. (2019). The role of *FRIGIDA* and *FLOWERING LOCUS C* genes in flowering time of *Brassica rapa* leafy vegetables. *Sci. Rep.* 9:13843.

Takagi, H., Abe, A., Yoshida, K., Kosugi, S., Natsume, S., Mitsuoka, C., et al. (2013). QTL-seq: rapid mapping of quantitative trait loci in rice by whole genome resequencing of DNA from two bulked populations. *Plant J.* 74, 174–183. doi: 10.1111/tpj.12105

van Wijk, K. J., and Kessler, F. (2017). Plastoglobuli: plastid microcompartments with integrated functions in metabolism, plastid developmental transitions, and environmental adaptation. *Annu. Rev. Plant Biol.* 68, 253–289. doi: 10.1146/annurev-arplant-043015-111737

Vishnevetsky, M., Ovadis, M., and Vainstein, A. (1999). Carotenoid sequestration in plants: the role of carotenoid-associated proteins. *Trends Plant Sci.* 4, 232–235. doi: 10.1016/s1360-1385(99)01414-4

Vishnevetsky, M., Ovadis, M., Itzhaki, H., Levy, M., Libal-Weksler, Y., Adam, Z., et al. (1996). Molecular cloning of a carotenoid-associated protein from *Cucumis sativus* corollas: homologous genes invol-ved in carotenoid sequestration in chromoplasts. *Plant J.* 10, 1111–1118. doi: 10.1046/j.1365-313x.1996.10061111.x

Wang, H., Huang, Y., Xiao, Q., Huang, X., Li, C., Gao, X., et al. (2020). Carotenoids modulate kernel texture in maize by influencing amyloplast envelope integrity. *Nat. Commun.* 11:5346.

Wang, N., Liu, Z. Y., Zhang, Y., Li, C. Y., and Feng, H. (2017). Identification and fine mapping of a stay-green gene (*Brnye1*) in pakchoi *(Brassica campestris* L. ssp. *chinensis*). *Theor. Appl. Genet.* 131, 673–684. doi: 10.1007/s00122-017-3028-8

Wang, X. W., Wang, H. Z., Wang, J., Sun, R. F., Wu, J., Liu, S. Y., et al. (2011). The genome of the mesopolyploid crop species *Brassica rapa*. *Nat. Genet.* 43, 1035–1039.

Wu, J., Mao, X., Cai, T., Luo, J., and Wei, L. (2006). KOBAS server: a web-based platform for automated annotation and pathway identification. *Nucleic Acids Res.* 34, W720–W724.

Yang, S. J., Yu, W. T., Wei, X. C., Wang, Z. Y., Zhao, Y. Y., Zhao, X. B., et al. (2020). An extended KASP-SNP resource for molecular breeding in Chinese cabbage (*Brassica rapa* L. ssp. *pekinensis*). *PLoS One* 15:e0240042. doi: 10.1371/journal.pone.0240042

Yuan, H., Zhang, J., Nageswaran, D., and Li, L. (2015). Carotenoid metabolism and regulation in horticultural crops. *Hortic. Res.* 2:15036.

Zhang, B., Liu, C., Wang, Y., Yao, X., Wang, F., Wu, J., et al. (2015). Disruption of a *CAROTENOID CLEAVAGE DIOXYGENASE 4* gene converts flower colour from white to yellow in *Brassica* species. *New Phytol.* 206, 1513–1526. doi: 10.1111/nph.13335

Zhang, J. X., Yuan, H., Fei, Z. J., Pogson, B. J., Zhang, L. G., and Li, L. (2015). Molecular characterization and transcriptome analysis of orange head Chinese cabbage (*Brassica rapa* L. ssp. *pekinensis*). *Planta* 241, 1381–1394. doi: 10.1007/s00425-015-2262-z

Zhang, N., Chen, L., Ma, S., Wang, R. F., He, Q., Tian, M., et al. (2020). Fine mapping and candidate gene analysis of the white flower gene *Brwf* in Chinese cabbage (*Brassica rapa* L.). *Sci. Rep.* 10:6080.

Zhang, X. X., Li, R. H., Chen, L., Niu, S. L., Chen, L., Gao, J., et al. (2018a). Fine-mapping and candidate gene analysis of the *Brassica juncea* white-flowered mutant *Bjpc2* using the whole-genome resequencing. *Mol. Genet. Genomics* 293, 359–370. doi: 10.1007/s00438-017-1390-5

Zhang, X. X., Li, R. H., Chen, L., Niu, S. L., Li, Q., Xu, K., et al. (2018b). Inheritance and gene mapping of the white flower trait in *Brassica juncea*. *Mol. Breed.* 38:20.

Zhang, X., Liu, Z. Y., Wang, P., Wang, Q., Yang, S., and Feng, H. (2013). Fine mapping of *BrWax1*, a gene controlling cuticular wax biosynthesis in Chinese cabbage (*Brassica rapa* L. ssp. *pekinensis*). *Mol. Breed.* 32, 867–874. doi: 10.1007/s11032-013-9914-0

Zhang, X., Zhang, K., Wu, J., Guo, N., Liang, J. L., Wang, X. W., et al. (2020). QTL-Seq and sequence assembly rapidly mapped the gene *BrMYBL2.1* for the purple trait in *Brassica rapa*. *Sci. Rep.* 10:2328.

Zhou, W., Niu, Y., Ding, X., Zhao, S., Li, Y., Fan, G., et al. (2020). Analysis of carotenoid content and diversity in apricots (*Prunus armeniaca L.*) grown in China. *Food Chem.* 330:127223. doi: 10.1016/j.foodchem.2020.127223

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.

frontiers in Plant Science

ORIGINAL RESEARCH
published: 14 May 2021
doi: 10.3389/fpls.2021.662684

Characterization of *BoaCRTISO* Reveals Its Role in Carotenoid Biosynthesis in Chinese Kale

Min Jiang[1†], Fen Zhang[1†], Qiao Yuan[1], Peixing Lin[1], Hao Zheng[1], Sha Liang[1], Yue Jian[1], Huiying Miao[2], Huanxiu Li[1], Qiaomei Wang[2] and Bo Sun[1*]*

[1] College of Horticulture, Sichuan Agricultural University, Chengdu, China, [2] Key Laboratory of Horticultural Plant Growth, Development and Quality Improvement, Ministry of Agriculture, Department of Horticulture, Zhejiang University, Hangzhou, China

Carotenoids are organic pigments that play an important role in both plant coloration and human health; they are a critical subject in molecular breeding due to growing demand for natural molecules in both food and medicine. In this study, we focus upon characterizing *BoaCRTISO*, the carotenoid isomerase gene before the branch of the carotenoid biosynthetic pathway, which is expressed in all organs and developmental stages of Chinese kale, and BoaCRTISO, which is located in the chloroplast. The expression of *BoaCRTISO* is induced by strong light, red and blue combined light, and gibberellic acid treatment, but it is suppressed by darkness and abscisic acid treatment. We obtained *BoaCRTISO*-silenced plants via virus-induced gene silencing technology, and the silence efficiencies ranged from 52 to 77%. The expressions of most carotenoid and chlorophyll biosynthetic genes in *BoaCRTISO*-silenced plants were downregulated, and the contents of carotenoids and chlorophyll were reduced. Meanwhile, *BoaCRTISO*-silenced plants exhibited phenotypes of yellowing leaves and inhibited growth. This functional characterization of *BoaCRTISO* provides insight for the biosynthesis and regulation of carotenoid in Chinese kale.

Keywords: *BoaCRTISO*, carotenoid biosynthesis, Chinese kale, color, gene expression, VIGS

OPEN ACCESS

Edited by:
Renata Rivera-Madrid, Scientific Research Center of Yucatán (CICY), Mexico

Reviewed by:
Tianhu Sun, Cornell University, United States
Sarah Frusciante, Italian National Agency for New Technologies, Energy and Sustainable Economic Development (ENEA), Italy

***Correspondence:**
Qiaomei Wang
qmwang@zju.edu.cn
Bo Sun
bsun@sicau.edu.cn

[†]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: *15 February 2021*
Accepted: *09 April 2021*
Published: *14 May 2021*

Citation:
Jiang M, Zhang F, Yuan Q, Lin P, Zheng H, Liang S, Jian Y, Miao H, Li H, Wang Q and Sun B (2021) Characterization of BoaCRTISO Reveals Its Role in Carotenoid Biosynthesis in Chinese Kale. Front. Plant Sci. 12:662684. doi: 10.3389/fpls.2021.662684

INTRODUCTION

Chinese kale (*Brassica oleracea* var. *alboglabra*) is a member of Brassicaceae, or the mustard family, which originated in South China and later spread throughout Southeast Asia (Lei et al., 2017). Its main edible parts are the highly nutritious tender leaves and bolting stems (Sun et al., 2018b). Our previous studies have shown that Chinese kale is rich in health-promoting compounds including vitamin C, glucosinolates, and carotenoids (Sun et al., 2012a,b). Carotenoids are the second largest group of natural pigments and are known to play crucial roles in plant color, photosynthesis, and human health (Xiao et al., 2012). The different types and contents of carotenoids lead to the rich and colorful red, yellow, and orange pigments of vegetables, fruits, and flowers such as tomato (Ye et al., 2015), citrus (Kato et al., 2004), watermelon (Lv et al., 2015), and chrysanthemum (Yamagishi et al., 2010). In green plant tissues, carotenoids are also light-harvesting pigments that are closely associated with chlorophylls in the photosynthetic membranes of plants where they protect chlorophyll against photooxidative damage (Guo et al., 2018; Llorente et al., 2020). In humans, carotenoids are necessary to maintain body health. Carotenoids are the precursors of retinoid synthesis and are theorized to act as antioxidants that reduce the risk of macular degeneration (Nisar et al., 2015). It is also reported that carotenoids could treat chronic

diseases, such as type 2 diabetes and cardiometabolic diseases, and lower the risk of cancer and cardiovascular disease (Gu et al., 2015; Liu et al., 2015). Given the growing list of human health benefits, horticulturists and others have a deep interest in developing carotenoid-rich food crops through breeding and metabolic engineering. Recent research generally uses molecular breeding methods to enhance carotenoids and to focus on changing biosynthetic pathways to alter metabolic flux (Saini et al., 2014; Kaur et al., 2017), such as golden rice (Beyer et al., 2002) and multivitamin corn (Naqvia et al., 2009).

Currently, the carotenoid biosynthetic pathway in complex and vascular plants has been clarified (Nisar et al., 2015). Several major carotenoids in plants appear in their *trans* form through the activity of isomerases, which are enzymes that function to convert molecules from one isomer to another. For example, carotenoid isomerase (CRTISO) is a key isomerase in the carotenoid biosynthetic pathway, which can catalyze pro-lycopene to lycopene (Pinheiro et al., 2019). In *Arabidopsis* (Park et al., 2002), tomato (Isaacson et al., 2002), and cabbage (Su et al., 2015), the loss of *CRTISO* function could cause the accumulation of pro-lycopene, which delays the greening of *Arabidopsis*, and changes the tomato color from red to orange and the cabbage color from yellow to orange. A *crtiso* mutant in rice produces a so-called zebra phenotype with alternating green and chlorotic crossbands on the mutant leaves (Chai et al., 2011). Besides the aesthetic effect on plant color, *CRTISO* also has a significant functional effect on photosynthetic efficiency. For instance, chlorophyll accumulation during photomorphogenesis was delayed markedly in a *ccr2* (*Atcrtiso*) mutant (Park et al., 2002), while the photosynthetic rate of yellow leaf sectors produced from a *crtiso* mutant in rice was approximately 27–42% of that of the wild-type unaffected leaves (Chai et al., 2011). In addition, carotenoid contents and related gene expressions can be regulated by exogenous phytohormone and light treatments in most plants (Divya et al., 2018). The exogenous methyl jasmonate (MeJA) spray treatment significantly enhanced the accumulation of lycopene in tomato (Liu et al., 2012) and other carotenoids in leafy vegetables (Divya et al., 2014; Saini et al., 2014), and abscisic acid (ABA) also had the similar effect in tomato (Sun et al., 2012), whereas gibberellic acid (GA_3) treatment had a marked effect on preventing the carotenoid accumulation in Navelate oranges (Rodrigo and Zacarias, 2007). It has also been found that the expression levels of *BoaPDS1* and *BoaPDS2* genes in Chinese kale were both regulated by red and blue combined light, and salicylic acid treatments (Sun et al., 2019). We suggest that these abiotic environmental conditions could regulate carotenoid biosynthetic pathways through *CRTISO*.

At present, the development of Chinese kale in China is relatively slow compared with other brassica vegetables. One of the key limiting reasons is that its color is not rich enough, and most of the germplasm resources are green (Lei et al., 2017), which makes it very difficult to breed new varieties with other colors via crossbreeding. In this study, we investigated the expression patterns of *BoaCRTISO* under different conditions; and we obtained *BoaCRTISO*-silenced plants using virus-induced gene silencing (VIGS) technology, and then we analyzed changes in the phenotype, carotenoid contents, and the related biosynthetic gene expressions of *BoaCRTISO*-silenced plants. Our findings allow a better understanding of the function of *BoaCRTISO* in regulating the carotenoid contents and coloring in Chinese kale, and this work will help in the creation of new varieties of Chinese kale from the perspective of molecular breeding.

MATERIALS AND METHODS

Plant Materials

In this study, we used the white-flower Chinese kale, *Brassica oleracea* var. *alboglabra* "Sijicutiao." The plants were grown in trays containing a mixture of peat and vermiculite (3:1) in an incubator with a light intensity of 80 μmol m^{-2} s^{-1}, a temperature of 25/20°C (day/night), a 12/12-h (day/night) light cycle, and 75% humidity. Seedlings with five to six true leaves were transplanted to the greenhouse of Sichuan Agricultural University for cultivation. Fertilizer and water were applied as needed.

Chinese kale materials were sampled at different developmental stages (germinating seeds, cotyledons, fifth to sixth true leaves, and mature leaves), as were organs at different stages of maturation (roots, bolting stems, petioles, leaves, inflorescence, fruit pod, and young fruit), and floral organs at the flower buds stage and the opening flowers stage (sepals, petals, stamens, and pistils). The materials were frozen by liquid nitrogen and stored at −80°C for subsequent studies (Sun et al., 2019).

Light and Phytohormone Treatments

The 30-day-old Chinese kale seedlings, raised in identical growth conditions, were selected for light intensity, light quality, and phytohormone treatments (Sun et al., 2019). The light quality treatments were divided into red light (R:B = 10:0), blue light (R:B = 0:10), and red and blue combined light (R:B = 5:5) treatments, while white light treatment was used as a control. Light intensity treatments were divided into darkness (0 μmol m^{-2} s^{-1}), weak light (40 μmol m^{-2} s^{-1}), and strong light (120 μmol m^{-2} s^{-1}) groups; the control group was exposed to white light (80 μmol m^{-2} s^{-1}). Phytohormones were sprayed on the leaf surface with 5 μM of GA_3 and 0.5 μM of ABA, and distilled water was used as control. When the leaves ceased dripping, the plants were moved into an artificial intelligence light incubator. In addition, 100 μM of MeJA was used to fumigate the Chinese kale seedlings in a transparent and closed container. The cultural conditions of the control and treated Chinese kale seedlings were as described in section "Plant Materials." The fifth to sixth true leaves were sampled at 0, 1, 3, 6, 12, 24, 48, and 72 h after each treatment, and the samples were immediately frozen in liquid nitrogen and stored at −80°C until RNA extraction.

RNA Extraction and qPCR Analysis

Total RNA was extracted from Chinese kale using a cetyl trimethylammonium bromide (CTAB) method (Chen et al., 2011). The well-integrated total RNA was used for cDNA

synthesis using the PrimeScript^{TM} 1st Strand cDNA Synthesis Kit (TAKARA, Japan). The qPCR primers for carotenoid and chlorophyll biosynthetic genes in Chinese kale were synthesized based on the qPCR primer database[1], and β-*actin* (Büchert et al., 2011) was used as the reference gene (**Supplementary Table 1**). The qPCR was performed referring to the TB Green Premix Ex Taq II (Tli RNaseH Plus) (TAKARA, Japan) kit instructions by using the Bio-Rad iCycler thermocycler (Bio-Rad, United States). Gene expression analysis was performed using the $2^{-\Delta\Delta CT}$ method (Livak and Schmittgen, 2001).

Molecular Cloning and Sequence Analysis

According to the sequences of the *CRTISO* gene and promoters of homologous species such as cabbage and Chinese cabbage published by Brassica database (BRAD), specific primers for the *BoaCRTISO* gene and promoters were designed (**Supplementary Table 1**). The method of gene cloning refers to Sun et al. (2018b). The CRTISO amino acid sequence of other species was downloaded from the National Center for Biotechnology Information (NCBI) and subjected to multiple sequence alignment using DNAMAN software (Lynnon Biosoft, United States). The phylogenetic tree was generated using the neighbor-joining method by MEGA 6.0 software (Tamura et al., 2013). The *cis*-acting elements on the promoter sequences of *BoaCRTISO* were predicted using the PlantCARE online software[2].

Subcellular Localization

Subcellular localization of BoaCRTISO was performed using the methods described by Sun et al. (2018a). The complete coding sequence (CDS) of *BoaCRTISO* was amplified using the primers BoaCRTISO-GFP-F/R (**Supplementary Table 1**), in which a *Bam*HI site at the 5′-end and a *Spe*I site at the 3′-end of the gene were incorporated. The BoaCRTISO and the pC2300-35S-eGFP plasmid digested with *Bam*HI and *Spe*I were mixed to generate pC2300-35S-*BoaCRTISO*-eGFP plasmid. Chinese kale mesophyll protoplasts were isolated and purified, and then the pC2300-35S-*BoaCRTISO*-eGFP and pC2300-35S-eGFP plasmids were transformed into Chinese kale protoplasts. The protoplasts expressing green fluorescent protein (GFP)-fusion protein were observed, and images were captured using a BX-51 fluorescence microscope equipped with a DP70 camera (Olympus, Japan).

Virus-Induced Gene Silencing of BoaCRTISO

A turnip yellow mosaic virus (TYMV)-induced gene silencing system was used to functionally characterize *BoaCRTISO* according to the previously described methods (Muntha et al., 2018). The pTY vector was digested with *Sna*BI, and the resulting linearized vector was analyzed by gel electrophoresis to confirm specificity. The linearized pTY vector and the annealed 80-base-pair-specific nucleotide of *BoaCRTISO* were ligated using T4 ligase to form a pTY-BoaCRTISO vector. The amplification of a *TYMV-CP* gene of the expected size (520 nt) was used to identify positive clones. Details of the pTY-CP F/R primer pairs are provided in **Supplementary Table 1**. For the virus infiltration, 5 μg of purified pTY-S carrying the target gene was diluted in 25 μl of double-distilled H_2O, which was then used to infiltrate two to four leaves of each plant. The infiltration was repeated every 5 days for 20 days, for a total of four interventions. The control plants were infiltrated with water, and the plants infiltrated with an empty pTY-S vector were used as the reference. Infiltrated plants were incubated in a growth chamber set at 22/20°C (day/night) with an 8/16-h (light/dark) cycle. One week after the last infiltration, plant phenotype analysis was performed, and the newly grown leaves were sampled for carotenoid composition, content, and related gene expression analysis.

Color

Color analysis of the transgenic plants was conducted using an NR110 chroma meter (3 nh, China). Three positions on the sampled leaves of each infiltrated plant were randomly selected, and the color values L^*, a^*, and b^* were obtained.

Determination of Carotenoid and Chlorophyll Composition and Contents

Carotenoid and chlorophyll concentrations were determined using the methods of Sun et al. (2021) with a slight modification. Approximately 200 mg of leaves was ground and extracted with 25 ml of acetone. The samples were sonicated for 20 min and then centrifuged at 4,000 *g* at room temperature for 5 min. The supernatant was filtered through 0.22-μm cellulose acetate filters and analyzed by high-performance liquid chromatography (HPLC). HPLC analysis of carotenoids and chlorophyll was carried out using an Agilent 1260 instrument with a VWD detector (Agilent, United States). Samples (10 μl) were separated at 30°C on a Waters C18 column (150 × 3.9-mm id; 4-μm particle size) using isopropanol and 80% acetonitrile–water at a flow rate of 0.5 ml min^{-1}. Absorbance was detected at 448 and 428 nm.

Statistical Analysis

All the results are shown as the means of three replicates. Statistical analysis was performed using SPSS version 18 (SPSS Inc., United States). The data were analyzed using one-way analysis of variance (ANOVA), and differences were compared using the least significant difference (LSD) test at a significance level of 0.05.

RESULTS

Isolation and Characterization of the BoaCRTISO

Analysis of the *Brassica oleracea* genome data revealed that the *CRTISO* gene has only one member. The results of cloning of *BoaCRTISO* also showed that there is only one *BoaCRTISO* in

[1]https://biodb.swu.edu.cn/qprimerdb/

[2]http://bioinformatics.psb.ugent.be/webtools/plantcare/html/

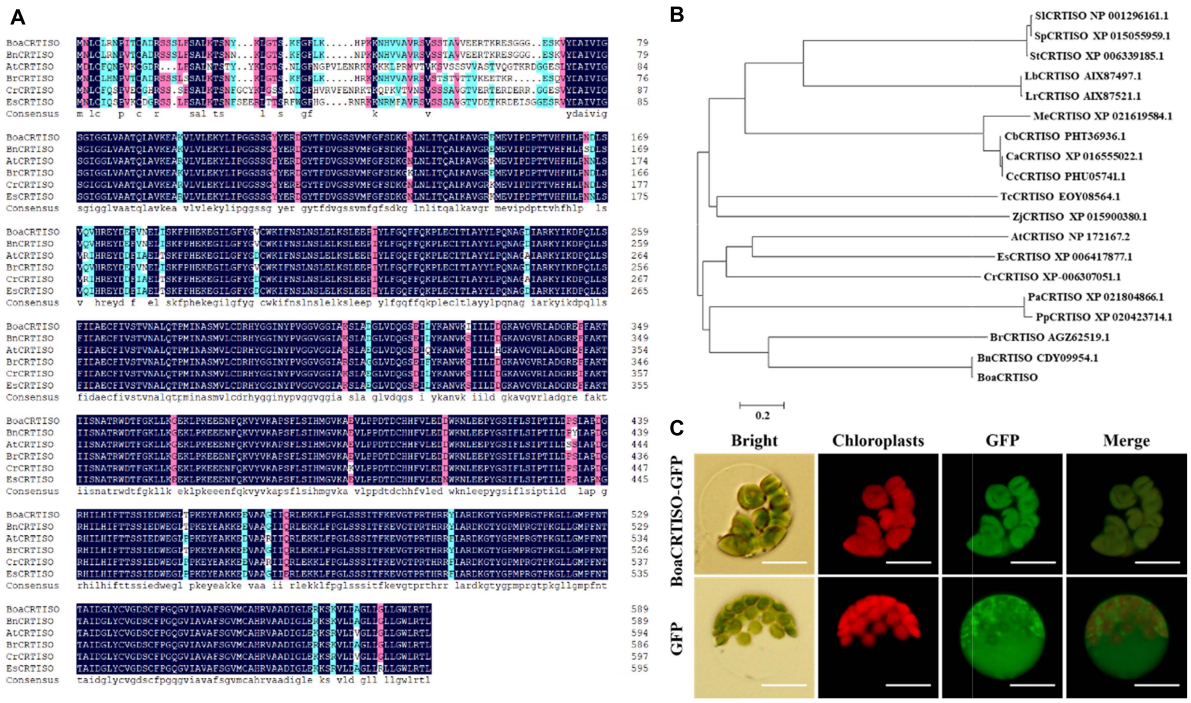


FIGURE 1 | Sequence characteristics of BoaCRTISO. **(A)** Alignment of the protein sequence of BoaCRTISO with selected homologs. The alignment was performed using DNAMAN software. The amino acids with 100% identity are shown with a black background, those with ≥75% identity are shown in red, and those with ≥50% identity are shown in blue. The species and their accession numbers in GenBank [*Brassica napus* (Bn): BnCRTISO (CDY09954.1), *Brassica rapa* (Br): BrCRTISO (AGZ62519.1), *Arabidopsis thaliana* (At): AtCRTISO (NP_172167.2), *Capsella rubella* (Cr): CrCRTISO (XP_006307051.1), and *Eutrema salsugineum* (Es): EsCRTISO (XP_006417877.1)] are listed here. **(B)** Phylogenetic analysis of BoaCRTISO and selected CRTISO from other plant species. The phylogenetic tree was generated using the neighbor-joining method by MEGA 6.0 software. The bar indicates an evolutionary distance of 0.2%. The species [*B. napus* (Bn), *B. rapa* (Br), *A. thaliana* (At), *C. rubella* (Cr), *E. salsugineum* (Es), *Prunus persica* (Pp), *Lycium barbarum* (Lb), *Lycium ruthenicum* (Lr), *Prunus avium* (Pa), *Manihot esculenta* (Me), *Ziziphus jujube* (Zj), *Solanum lycopersicum* (Sl), *Theobroma cacao* (Tc), *Solanum pennellii* (Sp), *Solanum tuberosum* (St), *Capsicum baccatum* (Cb), *Capsicum annuum* (Ca), and *Capsicum chinense* (Cc)] are listed here. **(C)** Subcellular localization of the BoaCRTISO-GFP fusion protein in Chinese kale protoplasts. Free green fluorescent protein served as a control. Bars = 30 μm.

Chinese kale. The CDS of the *BoaCRTISO* was cloned with sequence lengths of 1,773 bp (GenBank accession MN810158), which encodes a 590-amino-acid protein. Alignment results showed that the changes in amino acid sequences of CRTISO among different species mainly occurred at the N-terminus (**Figure 1A**). A phylogenetic analysis showed that BoaCRTISO clustered with other *Cruciferae* CRTISO (**Figure 1B**), which indicates that CRTISO is highly conserved in cruciferous plants. Moreover, a construct encoding BoaCRTISO fused to GFP was transformed into Chinese kale protoplasts, and strong fluorescence from GFP-BoaCRTISO was detected in the chloroplast (**Figure 1C**), which clearly demonstrates that BoaCRTISO localizes to the chloroplast.

Temporal and Spatial Expressions of BoaCRTISO

The *BoaCRTISO* gene was expressed in all developmental stages and organs (**Figure 2**). During the development of Chinese kale, the highest level of *BoaCRTISO* expression was in the cotyledon stage, followed by the germination and maturity stages, and it was lowest in the true leaves stage (**Figure 2A**). In different organs during the mature period, the expression levels of *BoaCRTISO* in inflorescences, seed pods, and young seeds were relatively high and were notably more than twice those of other organs (**Figure 2B**). During the flower buds and opening flowers stages, the expression levels of *BoaCRTISO* in both pistils and sepals were more than twice those observed in petals and stamens. Compared with the flower buds stage, the expression of *BoaCRTISO* in all flower organs was significantly downregulated at the opening flowers stage, except for stamens.

Light and Phytohormone Treatments Affected BoaCRTISO Expression

The promoter of the *BoaCRTISO* from Chinese kale leaves was cloned with sequence lengths of 1,789 bp (GenBank accession MN810159). The *cis*-elements of *BoaCRTISO* promoter mainly contained light- and phytohormone-responsive elements (**Figure 3A**). The light-responsive elements included AE-box, Box 4, TCT-motif, chs-Unit 1 m 1, G-box, GT1-motif, and MRE, while phytohormone-responsive elements included ABA-responsive element, MeJA response elements (CGTCA-motif and TGACG-motif), and GA response element (GARE-motif).

Under the treatment of different light intensities, it can be seen that the expression of *BoaCRTISO* is significantly induced

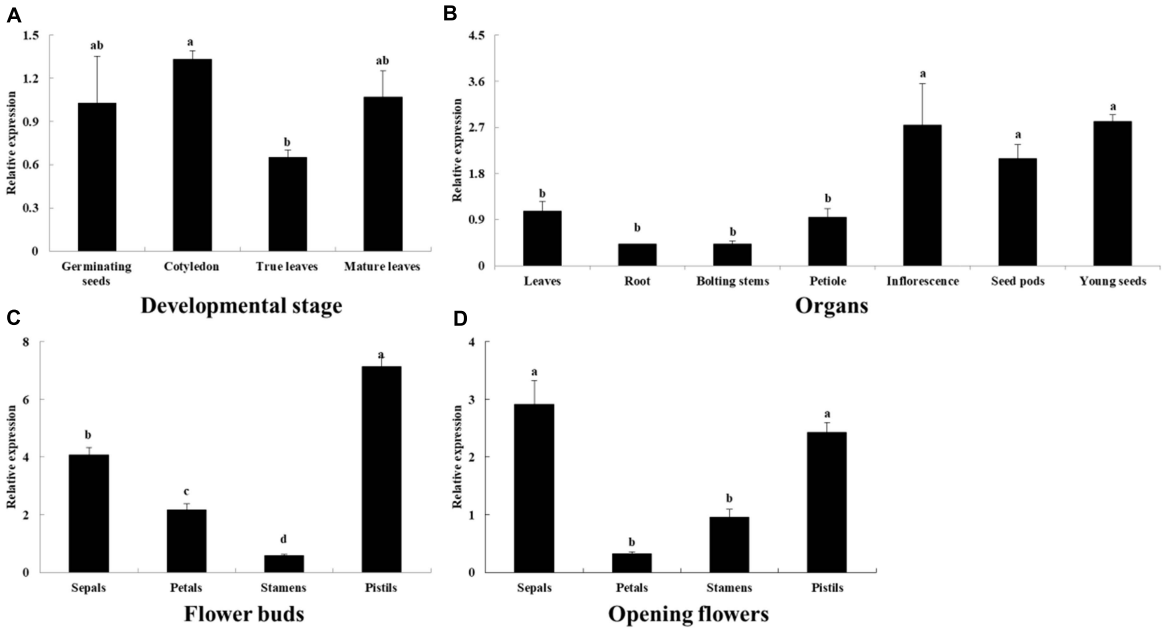


FIGURE 2 | Expression levels of *BoaCRTISO* in different developmental stages: **(A)** organs, **(B)** flower organs in flower buds stage **(C)**, and opening flowers stage **(D)** of Chinese kale. The *BoaCRTISO* expression of germinating seeds was set as 1. The samples were a mixture from three individual plants. Data are expressed as mean ± standard deviation. The same letter in the same histogram indicates that there is no significant difference between the values tested by least significant difference (LSD) ($p < 0.05$).

by strong light and significantly inhibited in the dark, while weak light has little effect on it (**Figure 3B** and **Supplementary Table 2**). Under the combined red and blue light treatments, the expression level of *BoaCRTISO* was significantly higher than that of the control group after 3 h, and its peak value was about three times that of the control group at 6 h. Compared with that of the control group, the expression of *BoaCRTISO* was significantly lower under red light treatment, but there was no significant difference under blue light treatment (**Figure 3C** and **Supplementary Table 2**). Under ABA treatment, the expression level of *BoaCRTISO* was inhibited before 24 h, but the inhibition gradually weakened afterward (**Figure 3D** and **Supplementary Table 2**). Under GA_3 treatment, the expression level of *BoaCRTISO* was significantly induced (**Figure 3E** and **Supplementary Table 2**), while the expression level of *BoaCRTISO* had no obvious regularity under MeJA treatment (**Figure 3F** and **Supplementary Table 2**).

BoaCRTISO Silencing Affected the Color and Inhibited the Growth of Chinese Kale

The expressions of *BoaCRTISO* in all pTY-BoaCRTISO plants were significantly reduced compared with control and the pTY plants, except for pTY-BoaCRTISO 1 plant (**Figure 4A**). Then, we observed that the pTY-BoaCRTISO plants were yellow-green compared with the dark-green of the control and the pTY plants, except for the pTY-BoaCRTISO 1 plant (**Figures 4B,C**). The changes of color between the pTY-BoaCRTISO plants, control, and pTY plants were then analyzed (**Figure 4D**). The yellow-blue value b^* of pTY-BoaCRTISO 2–6 plants were significantly higher than that of the control and pTY plants, which indicated that pTY-BoaCRTISO 2–6 plants were yellowing. The b^* value of the pTY-BoaCRTISO 2 plant with the most remarkable color change was 18, which was more than three times of that of the control and pTY plants. The red-green value a^* of pTY-BoaCRTISO 2–6 plants was significantly lower than that of control and pTY plants. The L^* value is similar among different groups. However, the L^*, a^*, and b^* values of pTY-BoaCRTISO 1 plant were all similar to those of control and pTY plants. In addition, the growth of pTY-BoaCRTISO plants were significantly inhibited, and the plant heights were significantly reduced compared with control and pTY plants, except for pTY-BoaCRTISO 1 plant (**Figures 4B,E**). Therefore, control 1, pTY 3, and pTY-BoaCRTISO 2, 5, and 6 with typical phenotypes and relatively high silencing efficiencies were chosen for subsequent analyses.

BoaCRTISO Silencing Reduced Carotenoids and Chlorophyll Accumulation in Chinese Kale

After infection with the pTY-BoaCRTISO plasmid, the expression levels of *CRTISO* of pTY-BoaCRTISO 2, 5, and 6 were significantly reduced and were 23%, 45%, and 48% of the control plants, while the silencing efficiencies were 77%, 55%, and 52%, respectively (**Supplementary Figure 1**). Next, the total and individual carotenoid contents in pTY-BoaCRTISO plants were analyzed, and we found that their contents were decreased compared with control and pTY plants, except that there was no significant change in β-carotene (**Figure 5**). Specifically, lutein, violaxanthin, and neoxanthin in pTY-BoaCRTISO 2 plant were

A

Function	*Cis*-element	Sequence(5'-3')	Location of *BoaCRTISO*
Light responsiveness	AE box	AGAAACTT	-1622(+)
	chs-Unit 1 m1	ACCTAACCCGG	-1076(+)
	TCT-motif	TCTTAC	-1655(-)
	GT1-motif	GGTTAA	-315(+)
	Box 4	ATTAAT	-1290(+), -554(-), -249(-)
	G-box	CACGTC	-1729(+)
	MRE	AACCTAA	-52(-)
Abscisic acid responsiveness	ABRE	AACCCGG	-1071(+)
		ACGTG	-1725(-)
		TACGGTC	-784(+)
MeJA responsiveness	CGTCA-motif	CGTCA	-262(-), -20(-)
	TGACG-motif	TGACG	-266(+), -24(+)
Gibberellin responsiveness	GARE-motif	TCTGTTG	-277(-)

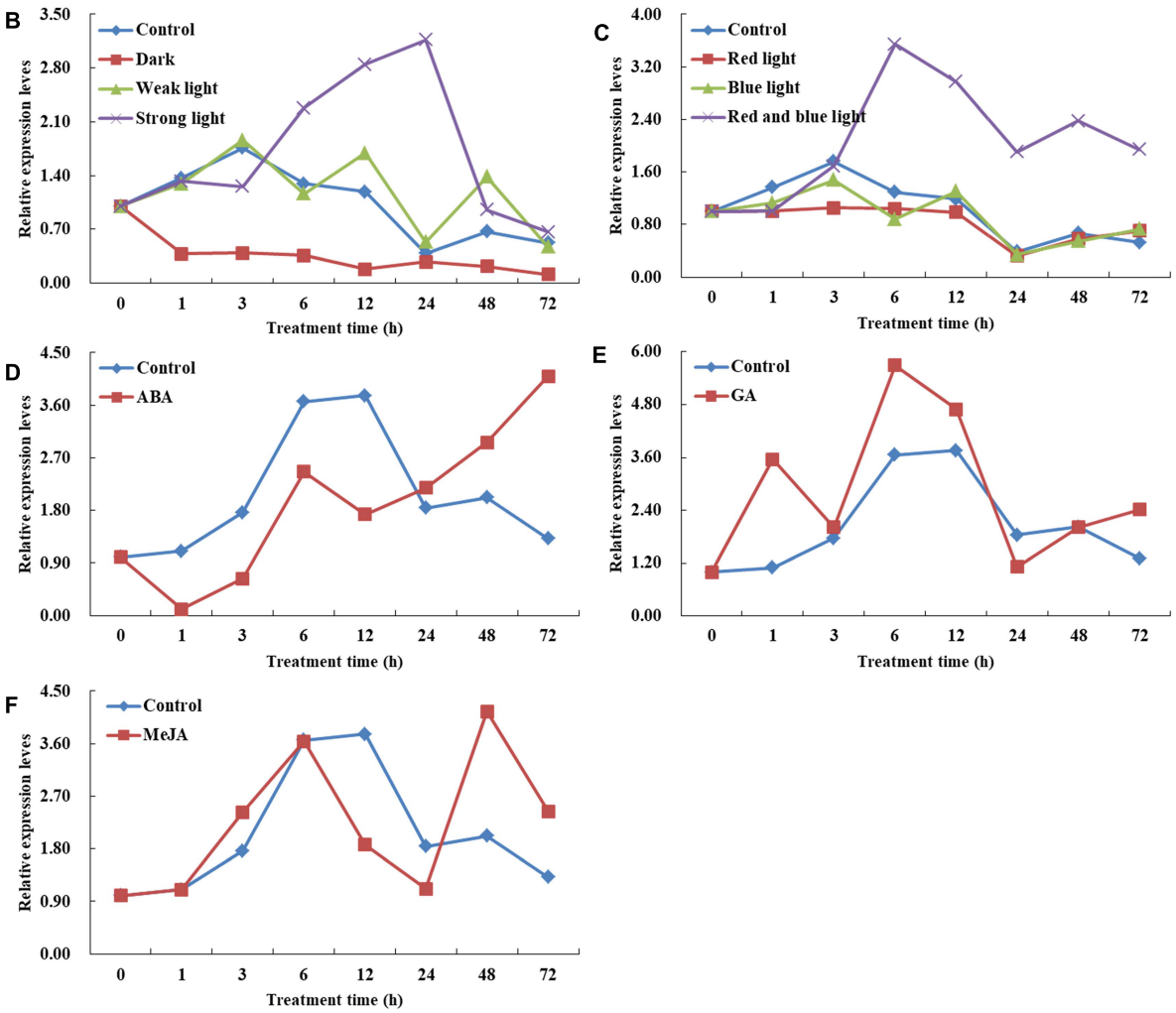
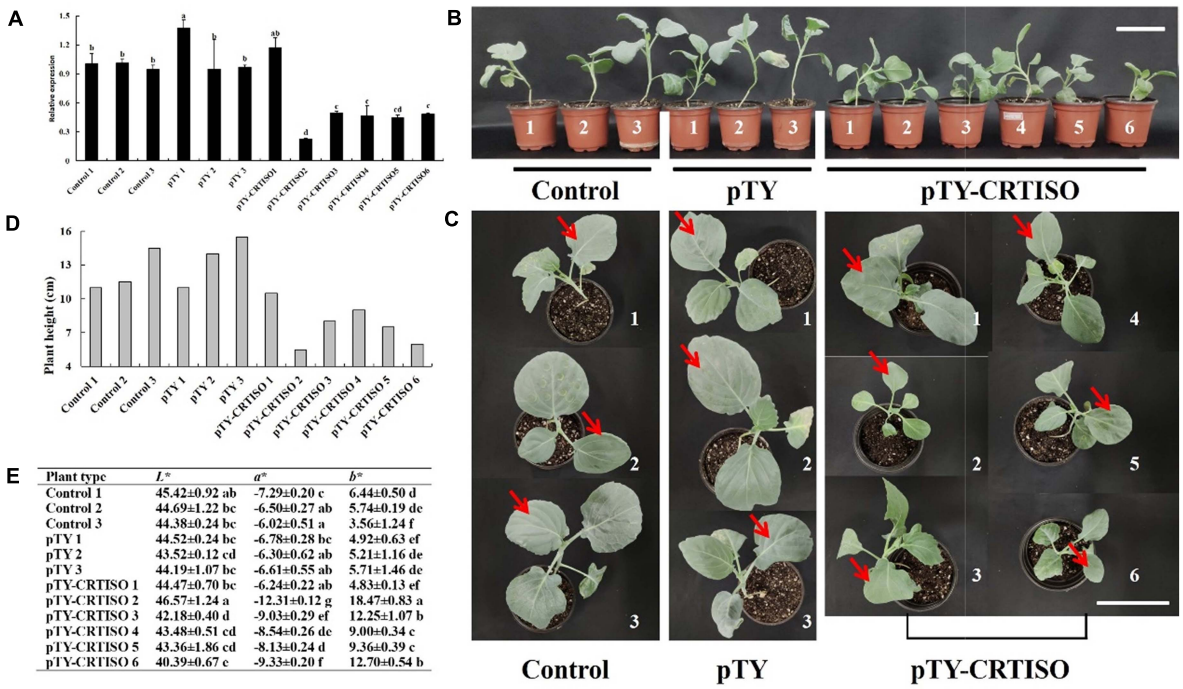


FIGURE 3 | Response of *BoaCRTISO* to different exogenous treatments. **(A)** *Cis*-acting elements respond to light and phytohormone in the promoter regions of *BoaCRTISO*. The + and − in brackets represent sense strand and antisense strand, respectively. **(B)** Expression levels of *BoaCRTISO* after treatments with darkness, weak light, and strong light. **(C)** Expression levels of *BoaCRTISO* after treatments with red light, blue light, and combined red and blue lights. **(D)** Expression level of *BoaCRTISO* after abscisic acid (ABA) treatment. **(E)** Expression level of *BoaCRTISO* after GA_3 treatment. **(F)** Expression level of *BoaCRTISO* after methyl jasmonate (MeJA) treatment.



Plant type	*L**	*a**	*b**
Control 1	45.42±0.92 ab	-7.29±0.20 c	6.44±0.50 d
Control 2	44.69±1.22 bc	-6.50±0.27 ab	5.74±0.19 de
Control 3	44.38±0.24 bc	-6.02±0.51 a	3.56±1.24 f
pTY 1	44.52±0.24 bc	-6.78±0.28 bc	4.92±0.63 ef
pTY 2	43.52±0.12 cd	-6.30±0.62 ab	5.21±1.16 de
pTY 3	44.19±1.07 bc	-6.61±0.55 ab	5.71±1.46 de
pTY-CRTISO 1	44.47±0.70 bc	-6.24±0.22 ab	4.83±0.13 ef
pTY-CRTISO 2	46.57±1.24 a	-12.31±0.12 g	18.47±0.83 a
pTY-CRTISO 3	42.18±0.40 d	-9.03±0.29 ef	12.25±1.07 b
pTY-CRTISO 4	43.48±0.51 cd	-8.54±0.26 de	9.00±0.34 c
pTY-CRTISO 5	43.36±1.86 cd	-8.13±0.24 d	9.36±0.39 c
pTY-CRTISO 6	40.39±0.67 c	-9.33±0.20 f	12.70±0.54 b

FIGURE 4 | *BoaCRTISO* silencing affects the color and growth of Chinese kale. **(A)** Expressions of *BoaCRTISO* in control, pTY, and pTY-BoaCRTISO plants. Data are expressed as mean ± standard deviation. The same letter indicates that there is no significant difference between the values tested by least significant difference (LSD) ($p < 0.05$). **(B)** Front view of control, pTY, and pTY-BoaCRTISO plants. Bar = 10 cm. **(C)** Top view of control, pTY, and pTY-BoaCRTISO plants. The red arrow points to the sampled leaves. Bar = 10 cm. **(D)** Plant heights of control, pTY, and pTY-BoaCRTISO plants at 1 week after the last infiltration. **(E)** The color parameters of control, pTY, and pTY-BoaCRTISO plants at 1 week after the last infiltration. Data are expressed as a mean ± SD. The same letter in the same column means no significant differences among values ($p < 0.05$) according to a least significant difference (LSD) test.

only approximately 70% of the control. Moreover, the contents of lutein, neoxanthin, and violaxanthin in pTY-BoaCRTISO 5 and 6 plants were all reduced, but only significantly so in neoxanthin and lutein. Since lutein content accounted for more than 60% of the total carotenoid levels in Chinese kale, the total carotenoid contents in pTY-BoaCRTISO 2 and 6 plants were significantly reduced when their lutein contents were significantly inhibited.

Chlorophyll, as a pigment closely related to carotenoids, has also been analyzed for its content. The individual and total chlorophyll of the pTY-BoaCRTISO 2 plant with the lowest *BoaCRTISO* expression decreased significantly. The changes in chlorophyll of pTY-BoaCRTISO 5 and 6 plants whose expression level of *BoaCRTISO* decreased by half were different. Chlorophyll *a* and total chlorophyll in pTY-BoaCRTISO 6 plants decreased significantly, whereas the individual and total chlorophylls of pTY-BoaCRTISO 5 plants did not change significantly.

BoaCRTISO Silencing Suppressed the Expression of Carotenoid and Chlorophyll Biosynthetic Genes

To investigate whether *BoaCRTISO* affects carotenoid accumulation by regulating the transcription of carotenoid biosynthetic genes, their expressions in the pTY-BoaCRTISO plants were analyzed by using RT-qPCR (**Figure 6** and **Supplementary Figure 1**). We found that the expression levels of all carotenoid biosynthetic genes in the pTY-BoaCRTISO 2 plant were significantly downregulated except *ZDS*, which was significantly upregulated, and β-*OHase*, which was unchanged. Regarding pTY-BoaCRTISO 5 and 6 plants, the expressions of most of their carotenoid biosynthetic genes were also downregulated, while the expressions of *ZDS*, *LCYe2*, and ε-*OHase* genes of pTY-BoaCRTISO 5 plant did not change significantly, and the expressions of *ZDS* and slightly increased while *NXS* did not change in the pTY-BoaCRTISO 6 plant.

In pTY-BoaCRTISO plants, chlorophyll, as the main contributor to the color of the leaves of Chinese kale, content decreased significantly, which made us pay attention to the expression of chlorophyll biosynthetic genes. The results indicated that the expression levels of the most chlorophyll biosynthetic genes detected were significantly decreased in the pTY-BoaCRTISO plants (**Supplementary Figure 2**). These findings suggest that the downregulation of *BoaCRTISO* expression has an inhibitory effect on the entire carotenoid and chlorophyll biosynthetic pathway in Chinese kale.

DISCUSSION

The short-duration conventional VIGS has many advantages, including easy and rapid results, unnecessary use of stable plant transformation, and the ability to apply multiple copies or family

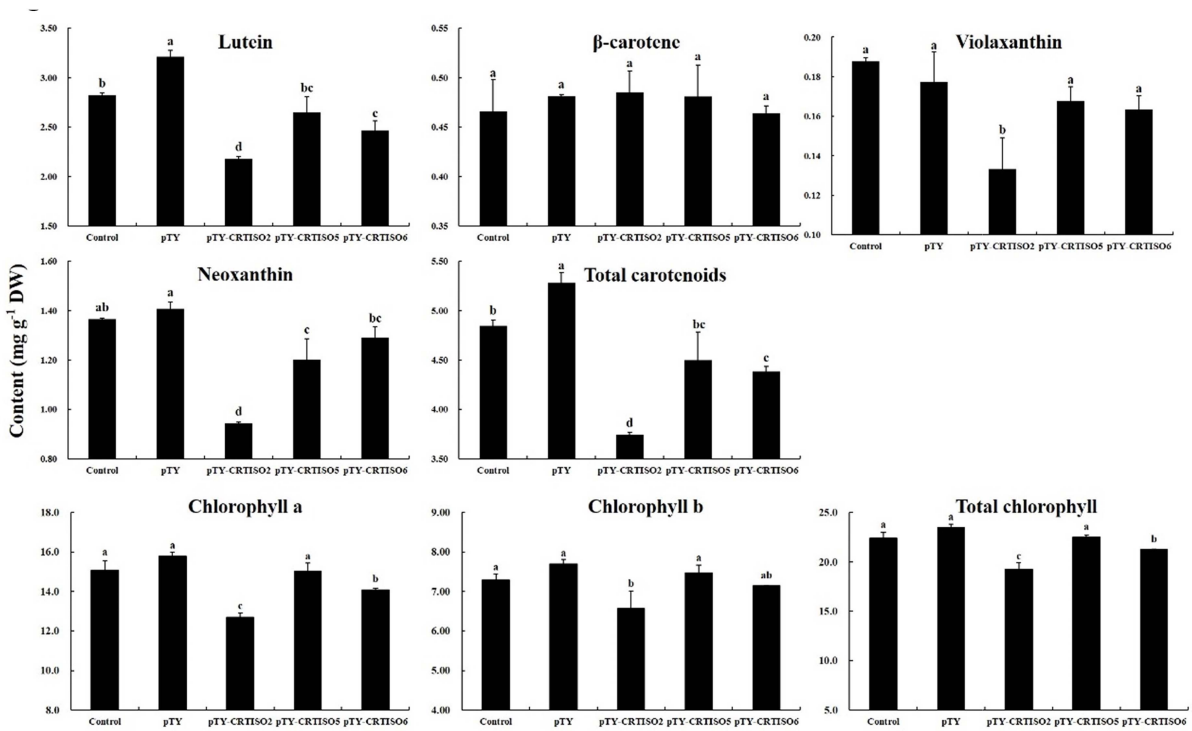


FIGURE 5 | Carotenoid composition and contents in control, pTY, and pTY-BoaCRTISO plants. Samples of leaves were taken from control, pTY, and pTY-BoaCRTISO plants at 1 week after the last infiltration. Data are expressed as mean ± standard deviation. The same letter in the same histogram indicates that there is no significant difference between the values tested by least significant difference (LSD) ($p < 0.05$).

members (Senthil-Kumar and Mysore, 2011). In this study, the short, inverted repeat coding for *BoaCRTISO* was inserted into the VIGS system of TYMV to silence the targeted gene, and the resulting *BoaCRTISO*-silenced plants were successfully obtained with a silencing efficiency of 52–77%. The efficiency of VIGS can be affected by multiple factors, such as crop species and virus types (Senthil-Kumar and Mysore, 2011). Several previous studies indicated that the silencing efficiencies of *Brassica juncea*, *Catharanthus roseus*, and strawberry were 90% (Muntha et al., 2018), 70% (Kumar et al., 2020), and 45% (Xie et al., 2020), respectively, which demonstrates the varying degrees of effectiveness across species. We have previously used tobacco rattle virus-mediated VIGS technology for gene silencing in Chinese kale; however, the efficiency of gene silencing was poor (data available upon request). Therefore, we suggest selecting an appropriate virus type when conducting VIGS assays for different crop species.

The role of carotenoids in plant coloration is well studied, and the biosynthetic pathways have been generally described. In this study, it was found that the leaves of Chinese kale color turned yellow and the b^* value increased in pTY-BoaCRTISO-treated plants that silenced *BoaCRTISO* expression. This result is consistent with several previous results but inconsistent with a few others. Our findings are similar to those on seedlings of melons with EMS mutagenesis, which leads to CRTISO being downregulated, or silenced, which appear yellow-green (Galpaz et al., 2013). In a separate study, the leaves showed a yellow-green zebra phenotype in a *crtiso* mutant in rice (Chai et al., 2011). However, the mutation of *CRTISO* in tomatoes changes the fruit phenotype to orange (Isaacson et al., 2002). These findings indicate that the phenotypes of leaves and other organs are expressed differently according to species after *CRTISO* is silenced. Besides, after the *BoaCRTISO* gene was target-edited in Chinese kale, the plant showed a distinct yellow color, although the gene expression level was similar to that of the VIGS plants (Sun et al., 2020). This shows that stable plant transformation can make plants show a more obvious phenotype than transient VIGS technology. In addition to leaves and fruits, color changes of the floral organs are also related to *BoaCRTISO* expression. We found that expression of *BoaCRTISO* in the flower buds stage of Chinese kale was significantly higher than that in the flowering stage, and this finding was contrary to the results of studies on lilies (Yamagishi et al., 2010) and chrysanthemums (Kishimoto and Ohmiya, 2006). This is likely because the Chinese kale flowering process is one of carotenoid degradation, and the opposite process occurs in the chrysanthemum and lily.

After determining that *BoaCRTISO* partial suppression can yellow Chinese kale leaves, we further explored the relationship among its phenotype, gene expression, and carotenoid contents. Prior work showed that the "Micro Tom" tomato was silenced using an inverted repeat of a fragment from the *Citrus CRTISO* that resulted in an overall decrease in transcript accumulation of all genes from the carotenoid biosynthetic

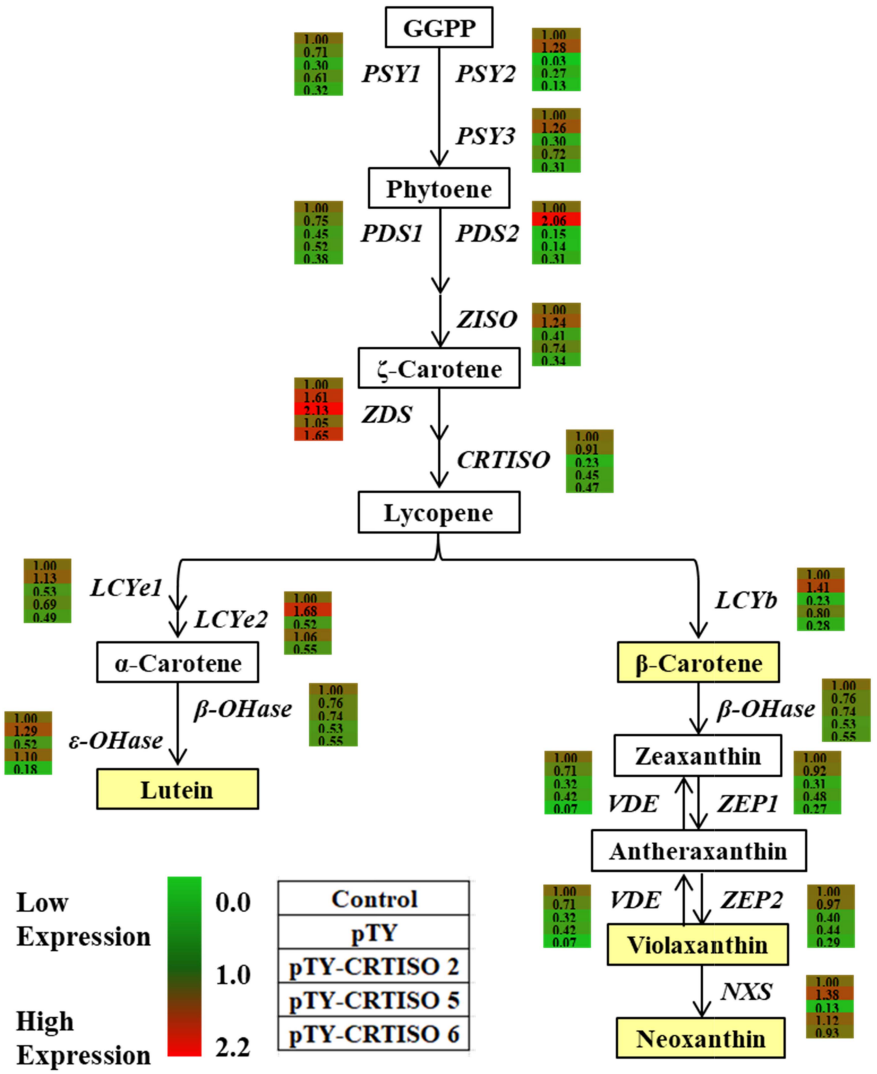


FIGURE 6 | Heat map of carotenoid biosynthetic gene expressions in control, pTY, and pTY-BoaCRTISO plants. Samples of leaves were taken from control, pTY, and pTY-BoaCRTISO plants at 1 week after the last infiltration. GGPP, geranylgeranyl diphosphate; PSY, phytoene synthase; PDS, phytoene desaturase; ZDS, ζ-carotene desaturase; Z-ISO, ζ-carotene isomerase; CRTISO, carotenoid isomerase; LCYe, lycopene ε-cyclase; LCYb, lycopene β-cyclase; ε-OHase, ε-carotene hydroxylase; β-OHase, β-carotene hydroxylase; VDE, violaxanthin de-epoxidase; ZEP, zeaxanthin epoxidase; NXS, neoxanthin synthase.

pathway (Pinheiro et al., 2019), whereas downregulation of *CRTISO* expression in tomato (Isaacson et al., 2002) and cabbage (Su et al., 2015) can lead to upregulation of upstream gene expression and downregulation of downstream gene expression in the carotenoid biosynthetic pathway, causing the accumulation of pro-lycopene. In our study, we found that after *BoaCRTISO* expression was downregulated, almost all carotenoid biosynthetic gene expressions in Chinese kale were also downregulated; this finding is consistent with results reported on citrus (Pinheiro et al., 2019) but inconsistent with that known from tomato (Isaacson et al., 2002) and cabbage (Su et al., 2015). Lutein is the primary component of carotenoids in Chinese kale, while prolycopene and lycopene are not detected. Therefore, the downregulation of *BoaCRTISO* expression led to the overall decline of carotenoid contents rather than the accumulation of prolycopene, as has been reported in other species. These results suggest that *CRTISO* might have variable functions in different species and that it could regulate the entire suite of carotenoid biosynthetic genes and carotenoid contents in Chinese kale.

We also found that *BoaCRTISO* expression can be regulated by exogenous light and hormone treatments. In our study, *BoaCRTISO* expression was induced under strong light, but it was inhibited in darkness compared with the control plants; this indicates that *CRTISO* is a light-inducible gene, which is consistent with the expression changes of citrus *CRTISO* under light and dark conditions (Gao et al., 2011). A few previous reports suggested that mixed (red and blue) or white lights could stimulate carotenoid contents in plants in comparison with applications of only red or only blue light (Tamura et al., 2013; Mastropasqua et al., 2020). Similarly,

we found that the red and blue combination light treatment significantly induced *BoaCRTISO* expression compared with applications of only red, blue, or white light. In addition to light, *BoaCRTISO* expression was also regulated by ABA, GA_3, and MeJA exogenous treatments. After ABA treatment, we observed that *BoaCRTISO* expression decreases significantly in the early stage; this may be due to the effect of feedback inhibition where endogenous ABA decreased because the content of exogenous ABA increased, and then the biosynthesis of carotenoids (which are an ABA precursor) decreased accordingly. In a similar fashion, it has been reported that gibberellin synthesis is inhibited after overexpression of a lycopene β-cyclase gene in tobacco, which indicates that gibberellin could negatively regulate carotenoid biosynthesis (Moreno et al., 2016). However, we found that *BoaCRTISO* expression was positively regulated by GA_3 treatment, which may be because the tested carotenoid biosynthetic genes differed in the previous study. In addition, MeJA has a dosage effect of high concentration inhibition and low concentration promotion on the regulation of carotenoids in tomato (Liu et al., 2012), which is consistent with our results. These results suggest that external environmental changes could be manipulated to regulate carotenoid biosynthesis.

Meanwhile, CRTISO plays an important role in photosynthesis and plant growth (Xie et al., 2019). In rice, the reduced photosynthetic rate of *crtiso* mutant leaves proved that *CRTISO* has an effect on photosynthesis (Chai et al., 2011). In our study, the expression of *BoaCRTISO* in leaves was significantly higher than that in root organs, and the subcellular localization of BoaCRTISO was in the chloroplast; this is a key organelle for photosynthesis, and the localization we observed may have occurred because leaves are the main photosynthetic organs and require more carotenoids than other plant tissues (Nisar et al., 2015). These results proved that *BoaCRTISO* plays an important role in photosynthesis by Chinese kale. In addition, the temporal and spatial expression patterns of *BoaCRTISO* were constitutive, and the heights of pTY-BoaCRTISO 2, 5, and 6 plants were significantly reduced, which indicates that *BoaCRTISO* is a necessary gene for the growth and development of Chinese kale. When the expression of carotenoid biosynthetic genes decreased, the growth and development of fleshy roots of carrots were also restricted (Flores-Ortiz et al., 2020), which is consistent with our results. In addition, the chlorophyll content and biosynthetic gene expressions in pTY-BoaCRTISO plants were significantly reduced, which would also affect photosynthesis to inhibit plant growth. This suggested that *BoaCRTISO* may affect plant growth by directly or indirectly affecting photosynthesis; nevertheless, the specific mechanism needs further study.

In summary, the carotenoid isomerase gene *BoaCRTISO* and promoter of Chinese kale were cloned, and BoaCRTISO was located in the chloroplast via a subcellular localization assay. The expression of *BoaCRTISO* was detected in all growth and development periods, and organs, of Chinese kale. VIGS-mediated *BoaCRTISO* silencing suppressed the expression of carotenoid and chlorophyll biosynthetic genes, decreased carotenoid and chlorophyll contents, yellowed Chinese kale, and inhibited Chinese kale growth. In addition, *BoaCRTISO* expression can be induced by strong light, combined red and blue light, and GA_3 treatments but inhibited by darkness and ABA treatments. These findings collectively mean that *BoaCRTISO* is a viable gene to target for improvements of the appearance and nutritional quality of Chinese kale for the betterment of human health.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**, further inquiries can be directed to the corresponding authors.

AUTHOR CONTRIBUTIONS

MJ: investigation, data curation, and writing – original draft. FZ: methodology and writing – original draft. QY and PL: data curation. HZ, SL, and YJ: investigation. HM: methodology and data curation. QW and HL: conceptualization, funding acquisition, and writing – review and editing. BS: conceptualization, funding acquisition, writing – review and editing, and supervision. All authors review and editing.

FUNDING

This work was supported by the National Natural Science Foundation of China (32072586 and 31500247), Zhejiang Provincial Ten-thousand Program for Leading Talents of Science and Technology Innovation (2018R52026), Sichuan Science and Technology Program (2018NZ0081), Project of New Varieties Breeding of Sichuan Vegetable Innovation Team (sccxtd-2020-05), and Technology Project of Zhishi Supply Chain Technology Co., Ltd. (2020008).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2021.662684/full#supplementary-material

Supplementary Figure 1 | Expressions of carotenoid biosynthetic genes in control, pTY, and pTY-BoaCRTISO plants. Samples of leaves were taken from control, pTY, and pTY-BoaCRTISO plants at 1 week after the last infiltration. The carotenoid biosynthetic gene expressions were calculated based on the respective expression of the respective genes in control plant. Data are expressed as mean ± standard deviation. The same letter in the same histogram indicates that there is no significant difference between the values tested by LSD ($p < 0.05$).

Supplementary Figure 2 | Expressions of chlorophyll biosynthetic genes in control, pTY, and pTY-BoaCRTISO plants. Samples of leaves were taken from control, pTY, and pTY-BoaCRTISO plants at 1 week after the last infiltration. The chlorophyll biosynthetic gene expressions were calculated based on the respective expression of the respective genes in control plant. Data are expressed as mean ± standard deviation. The same letter in the same histogram indicates that there is no significant difference between the values tested by LSD ($p < 0.05$). ALAD, 5-aminolevulinic acid dehydratase; HemE1, glutamyl tRNA reductase; Chl, magnesium-chelatase; CS, chlorophyll synthase.

REFERENCES

Beyer, P., Al-Babili, S., Ye, X. D., Lucca, P., Schaub, P., Welsch, R., et al. (2002). Golden rice: introducing the β-carotene biosynthesis pathway into rice endosperm by genetic engineering to defeat vitamin a deficiency. *J. Nutr.* 2, 506S–510S.

Büchert, A. M., Civello, P. M., and Martínez, G. A. (2011). Effect of hot air, UV-C, white light and modified atmosphere treatments on expression of chlorophyll degrading genes in postharvest broccoli (*Brassica oleracea* L.) florets. *Sci. Hortic.* 127, 214–219. doi: 10.1016/j.scienta.2010.11.001

Chai, C. L., Fang, J., Liu, Y., Tong, H. N., Gong, Y. Q., Wang, Y. Q., et al. (2011). ZEBRA2, encoding a carotenoid isomerase, is involved in photoprotection in rice. *Plant Mol. Biol.* 75, 211–221. doi: 10.1007/s11103-010-9719-z

Chen, Q., Yu, H. W., Wang, X. R., Xie, X. L., Yue, X. Y., and Tang, H. R. (2011). An alternative cetyltrimethylammonium bromide-based protocol for RNA isolation from blackberry (Rubus L.). *Genet. Mol. Res.* 11, 1773–1782. doi: 10.4238/2012.june.29.10

Divya, P., Puthusseri, B., and Neelwarne, B. (2014). The effect of plant regulators on the concentration of carotenoids and phenolic compounds in foliage of coriander. *LWT Food Sci. Technol.* 56, 101–110. doi: 10.1016/j.lwt.2013.11.012

Divya, P., Puthusseri, B., Lokesh, M. A., and Neelwarne, B. (2018). Effects of methyl jasmonate and carotenogenic inhibitors on gene expression and carotenoid accumulation in coriander (*Coriandrum sativum* L.) foliage. *Food Res. Int.* 111, 11–19. doi: 10.1016/j.foodres.2018.04.040

Flores-Ortiz, C., Alvarez, L. M., Undurraga, A., Arias, D., Durán, F., Wegener, G., et al. (2020). Differential role of the two ζ-carotene desaturase paralogs in carrot (*Daucus carota*): ZDS1 is a functional gene essential for plant development and carotenoid synthesis. *Plant Sci.* 291:110327. doi: 10.1016/j.plantsci.2019.110327

Galpaz, N., Burger, Y., Lavee, T., Tzuri, G., Sherman, A., Melamed, T., et al. (2013). Genetic and chemical characterization of an EMS induced mutation in Cucumis melo CRTISO gene. *Arch. Biochem. Biophys.* 539, 117–125. doi: 10.1016/j.abb.2013.08.006

Gao, H. J., Xu, J., Liu, X., Giuliano, G., Batschauer, A., and Kleinig, H. (2011). Light effect on carotenoids production and expression of carotenogenesis genes in citrus callus of four genotypes. *Acta Physiol. Plant.* 33, 2485–2492. doi: 10.1007/s11738-011-0793-x

Gu, H. H., Wang, J. S., Zhao, Z. Q., Sheng, X. G., Yu, H. F., and Huang, W. B. (2015). Characterization of the appearance, health promoting compounds, and antioxidant capacity of the florets of the loose-curd cauliflower. *Int. J. Food Prop.* 18, 392–402. doi: 10.1080/10942912.2013.831445

Guo, X. Y., Song, C. K., Ho, C. T., and Wan, X. C. (2018). Contrbution of L-theanine to the formation of 2, 5-dimetheylpyrazine, a key roasted peanutty flavor in Oolong tea during manufacturing processes. *Food Chem.* 263, 18–28. doi: 10.1016/j.foodchem.2018.04.117

Isaacson, T., Ronen, G., Zamir, D., and Hirschberg, J. (2002). Cloning of tangerine from tomato reveals a carotenoid isomerase essential for the production of β-carotene and xanthophylls in plants. *Plant Cell* 14, 333–342. doi: 10.1105/tpc.010303

Kato, M., Ikoma, Y., Matsumoto, H., Sugiura, M., Hyodo, H., and Yano, M. (2004). Accumulation of carotenoids and expression of carotenoid biosynthetic genes during maturation in citrus fruit. *Plant Physiol.* 134, 824–837. doi: 10.1104/pp.103.031104

Kaur, N., Pandey, A. S., Kumar, P., Pandey, P., Kesarwani, A. K., Mantri, S. S., et al. (2017). Regulation of banana phytoene synthase (MaPSY) expression, characterization and their modulation under various abiotic stress conditions. *Front. Plant Sci.* 8:462.

Kishimoto, S., and Ohmiya, A. (2006). Regulation of carotenoid biosynthesis in petals and leaves of chrysanthemum (*Chrysanthemum morifolium*). *Physiol. Plant* 128, 437–447.

Kumar, S. R., Rai, A., Bomzan, D. P., Kumar, K., Hemmerlin, A., Dwivedi, V., et al. (2020). A plastid-localized bona fide geranylgeranyl diphosphate synthase plays a necessary role in monoterpene indole alkaloid biosynthesis in *Catharanthus roseus*. *Plant J.* 103, 248–265. doi: 10.1111/tpj.14725

Lei, J. J., Chen, G. J., Chen, C. M., and Cao, B. H. (2017). Germplasm diversity of Chinese kale in China. *Hortic. Plant J.* 3, 101–104. doi: 10.1016/j.hpj.2017.07.006

Liu, L. H., Shao, Z. Y., Zhang, M., and Wang, Q. M. (2015). Regulation of carotenoid metabolism in tomato. *Mol. Plant* 8, 28–39. doi: 10.1016/j.molp.2014.11.006

Liu, L. H., Wei, J., Zhang, M., Zhang, L. P., Li, C. Y., and Wang, Q. M. (2012). Ethylene independent induction of lycopene biosynthesis in tomato fruits by jasmonates. *J. Exp. Bot.* 63, 5751–5761. doi: 10.1093/jxb/ers224

Livak, K. J., and Schmittgen, T. D. (2001). Analysis of relative gene expression data using real-time quantitative PCR and the 2-ΔΔCT method. *Methods* 25, 402–408. doi: 10.1006/meth.2001.1262

Llorente, B., Torres-Montilla, S., Morelli, L., Florez-Sarasa, I., Matus, J. T., Ezquerro, M., et al. (2020). Synthetic conversion of leaf chloroplasts into carotenoid-rich plastids reveals mechanistic basis of natural chromoplast development. *P. Natl. Acad. Sci. USA* 117, 21796–21803. doi: 10.1073/pnas.2004405117

Lv, P., Li, N., Liu, H., Gu, H. H., and Zhao, W. E. (2015). Changes in carotenoid profiles and in the expression pattern of the genes in carotenoid metabolisms during fruit development and ripening in four watermelon cultivars. *Food Chem.* 174, 52–59. doi: 10.1016/j.foodchem.2014.11.022

Mastropasqua, L., Dipierro, N., and Paciolla, C. (2020). Effects of darkness and light spectra on nutrients and pigments in Radish, Soybean, Mung Bean and Pumpkin Sprouts. *Antioxidants* 9:558. doi: 10.3390/antiox9060558

Moreno, J. C., Cerda, A., Simpson, K., Lopez-Diaz, I., Carrera, E., Handford, M., et al. (2016). Increased *Nicotiana tabacum* fitness through positive regulation of carotenoid, gibberellin and chlorophyll pathways promoted by *Daucus carota* lycopene β-cyclase (Dclcyb1) expression. *J. Exp. Bot.* 67, 2325–2338. doi: 10.1093/jxb/erw037

Muntha, S. T., Zhang, L. L., Zhou, Y. F., Zhao, X., Hu, Z. Y., Yang, J. H., et al. (2018). Phytochrome a signal transduction 1 and CONSTANS-LIKE 13 coordinately orchestrate shoot branching and flowering in leafy Brassica juncea. *Plant Biotechnol. J.* 17:13057.

Naqvia, S., Zhua, C. F., Farrea, G., Ramessara, K., Bassiea, L., Breitenbachb, J., et al. (2009). Transgenic multivitamin corn through biofortification of endosperm with three vitamins representing three distinct metabolic pathways. *P. Natl. Acad. Sci. USA.* 106, 7762–7767. doi: 10.1073/pnas.0901412106

Nisar, N., Li, L., Lu, S., Khin, N. C., and Pogson, B. J. (2015). Carotenoid metabolism in plants. *Mol. Plant* 8, 68–82. doi: 10.1016/j.molp.2014.12.007

Park, H., Kreunen, S. S., Cuttriss, A. J., DellaPenna, D., and Pogson, B. J. (2002). Identification of the carotenoid isomerase provides insight into carotenoid biosynthesis, prolamellar body formation, and photomorphogenesis. *Plant Cell* 14, 321–332. doi: 10.1105/tpc.010302

Pinheiro, T. T., Peres, L. E. P., Purgatto, E., Latado, R. R., Maniero, R. A., Martins, M. M., et al. (2019). Citrus carotenoid isomerase gene characterization by complementation of the "Micro-Tom" tangerine mutant. *Plant Cell Rep.* 38, 623–636. doi: 10.1007/s00299-019-02393-2

Rodrigo, M. J., and Zacarias, L. (2007). Effect of postharvest ethylene treatment on carotenoid accumulation and the expression of carotenoid biosynthetic genes in the flavedo of orange (*Citrus sinensis* L. Osbeck) fruit. *Postharvest Biol. Tech.* 43, 14–22.

Saini, R. K., Prashanth, K. V. H., Shetty, N. P., and Giridhar, P. (2014). Elicitors, SA and MJ enhance carotenoids and tocopherol biosynthesis and expression of antioxidant related genes in *Moringa oleifera* Lam. leaves. *Acta Physiol. Plant.* 36, 2695–2704. doi: 10.1007/s11738-014-1640-7

Senthil-Kumar, M., and Mysore, K. S. (2011). New dimensions for VIGS in plant functional genomics. *Trends Plant Sci.* 16, 656–665. doi: 10.1016/j.tplants.2011.08.006

Su, T. B., Yu, S. C., Wang, J., Zhang, F. L., Yu, Y. J., Zhang, D. S., et al. (2015). Loss of function of the carotenoid isomerase gene BrCRTISO confers orange color to the inner leaves of Chinese cabbage (*Brassica rapa* L. ssp. pekinensis). *Plant Mol. Biol. Rep.* 33, 648–659. doi: 10.1007/s11105-014-0779-0

Sun, B., Di, H. M., Zhang, J. Q., Xia, P. X., Huang, W. L., Jian, Y., et al. (2021). Effect of light on sensory quality, health-promoting phytochemicals and antioxidant capacity in post-harvest baby mustard. *Food Chem.* 339:128057. doi: 10.1016/j.foodchem.2020.128057

Sun, B., Jiang, M., Liang, S., Zheng, H., Chen, Q., Wang, Y., et al. (2019). Functional differences of BaPDS1 and BaPDS2 genes in Chinese kale. *Roy. Soc. Open Sci.* 6:190260. doi: 10.1098/rsos.190260

Sun, B., Jiang, M., Zheng, H., Jian, Y., Huang, W. L., Yuan, Q., et al. (2020). Color-related chlorophyll and carotenoid concentrations of Chinese kale can be altered through CRISPR/Cas9 targeted editing of the carotenoid isomerase gene BoaCRTISO. *Hortic. Res.* 7:161.

Sun, B., Yan, H. Z., Liu, N., Wei, J., and Wang, Q. M. (2012a). Effect of 1-MCP treatment on postharvest quality characters, antioxidants and glucosinolates

of Chinese kale. *Food Chem.* 131, 519–526. doi: 10.1016/j.foodchem.2011.09.016

Sun, B., Yan, H. Z., Zhang, F., and Wang, Q. M. (2012b). Effects of plant hormones on main health-promoting compounds and antioxidant capacity of Chinese kale. *Food Res. Int.* 48, 359–366. doi: 10.1016/j.foodres.2012.04.021

Sun, B., Zhang, F., Xiao, N., Jiang, M., Yuan, Q., Xue, S. L., et al. (2018a). An efficient mesophyll protoplast isolation, purification and PEG-mediated transient gene expression for subcellular localization in Chinese kale. *Sci. Hortic.* 241, 187–193. doi: 10.1016/j.scienta.2018.07.001

Sun, B., Zhang, F., Xue, S. L., Chang, J. Q., Zheng, A. H., Jiang, M., et al. (2018b). Molecular Cloning and Expression Analysis of the ζ-Carotene Desaturase Gene in Chinese kale (*Brassica oleracea* var. alboglabra Bailey). *Hortic. Plant J.* 4, 94–102. doi: 10.1016/j.hpj.2018.03.005

Sun, L., Yuan, B., Zhang, M., Wang, L., Cui, M., Wang, Q., et al. (2012). Fruit-specific RNAi-mediated suppression of SlNCED1 increases both lycopene and beta-carotene contents in tomato fruit. *J. Exp. Bot.* 63, 3097–3108. doi: 10.1093/jxb/ers026

Tamura, K., Stecher, G., Peterson, D., Filipski, A., and Kumar, S. (2013). MEGA6: molecular evolutionary genetics analysis version 6.0. *Mol. Biol. Evol.* 30, 2725–2729. doi: 10.1093/molbev/mst197

Xiao, Z. L., Lester, G. E., Luo, Y. G., and Wang, Q. (2012). Assessment of vitamin and carotenoid concentrations of emerging food products: edible microgreens. *J. Agric. Food Chem.* 60, 7644–7651. doi: 10.1021/jf300459b

Xie, J., Yao, S. X., Ming, J., Deng, L. L., and Zeng, K. F. (2019). Variations in chlorophyll and carotenoid contents and expression of genes involved in pigment metabolism response to oleocellosis in citrus fruits. *Food Chem.* 272, 49–57. doi: 10.1016/j.foodchem.2018.08.020

Xie, Y. G., Ma, Y. Y., Bi, P. P., Wei, W., Li, J., Hu, Y., et al. (2020). Transcription factor FvTCP9 promotes strawberry fruit ripening by regulating the biosynthesis of abscisic acid and anthocyanins. *Plant Physiol. Biochem.* 146, 374–383. doi: 10.1016/j.plaphy.2019.11.004

Yamagishi, M., Kishimoto, S., and Nakayama, M. (2010). Carotenoid composition and changes in expression of carotenoid biosynthetic genes in tepals of Asiatic hybrid lily. *Plant Breed.* 129, 100–107. doi: 10.1111/j.1439-0523.2009.01656.x

Ye, J. H., T. X., Yang, C. M., Li, H. X., Yang, M. Z., et al. (2015). Transcriptome profiling of tomato fruit development reveals transcription factors associated with ascorbic acid, carotenoid and flavonoid biosynthesis. *PLoS One* 10:e0130885. doi: 10.1371/journal.pone.0130885

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.

frontiers in Plant Science

ORIGINAL RESEARCH
published: 11 June 2021
doi: 10.3389/fpls.2021.685136

Unraveling the Regulatory Mechanism of Color Diversity in *Camellia japonica* Petals by Integrative Transcriptome and Metabolome Analysis

*Mingyue Fu[1], Xu Yang[2], Jiarui Zheng[1], Ling Wang[1], Xiaoyan Yang[1], Yi Tu[1], Jiabao Ye[1], Weiwei Zhang[1], Yongling Liao[1], Shuiyuan Cheng[3] and Feng Xu[1]**

[1] College of Horticulture and Gardening, Yangtze University, Jingzhou, China, [2] Department of Forestry Ecology, Hubei Ecology Polytechnic College, Wuhan, China, [3] National R&D Center for Se-Rich Agricultural Products Processing, Wuhan Polytechnic University, Wuhan, China

OPEN ACCESS

Edited by:
Xiumin Fu,
South China Botanical Garden, Chinese Academy of Sciences, China

Reviewed by:
Lingyun Zhang,
South China Agricultural University, China
Xin Mei,
Qiannan Normal College for Nationalities, China

***Correspondence:**
Feng Xu
xufeng198@126.com

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

Received: *24 March 2021*
Accepted: *17 May 2021*
Published: *11 June 2021*

Citation:
Fu M, Yang X, Zheng J, Wang L, Yang X, Tu Y, Ye J, Zhang W, Liao Y, Cheng S and Xu F (2021) Unraveling the Regulatory Mechanism of Color Diversity in Camellia japonica Petals by Integrative Transcriptome and Metabolome Analysis. Front. Plant Sci. 12:685136. doi: 10.3389/fpls.2021.685136

Camellia japonica petals are colorful, rich in anthocyanins, and possess important ornamental, edible, and medicinal value. However, the regulatory mechanism of anthocyanin accumulation in *C. japonica* is still unclear. In this study, an integrative analysis of the metabolome and transcriptome was conducted in five *C. japonica* cultivars with different petal colors. Overall, a total of 187 flavonoids were identified (including 25 anthocyanins), and 11 anthocyanins were markedly differentially accumulated among these petals, contributing to the different petal colors in *C. japonica*. Moreover, cyanidin-3-*O*-(6″-*O*-malonyl) glucoside was confirmed as the main contributor to the red petal phenotype, while cyanidin-3-*O*-rutinoside, peonidin-3-*O*-glucoside, cyanidin-3-*O*-glucoside, and pelargonidin-3-*O*-glucoside were responsible for the deep coloration of the *C. japonica* petals. Furthermore, a total of 12,531 differentially expressed genes (DEGs) and overlapping DEGs (634 DEGs) were identified by RNA sequencing, and the correlation between the expression level of the DEGs and the anthocyanin content was explored. The candidate genes regulating anthocyanin accumulation in the *C. japonica* petals were identified and included 37 structural genes (especially *CjANS* and *Cj4CL*), 18 keys differentially expressed transcription factors (such as *GATA*, *MYB*, *bHLH*, *WRKY*, and *NAC*), and 16 other regulators (mainly including transporter proteins, zinc-finger proteins, and others). Our results provide new insights for elucidating the function of anthocyanins in *C. japonica* petal color expression.

Keywords: *Camellia japonica*, petal, anthocyanin, color diversity, transcriptome, transcription factor, structural genes, metabolome

Abbreviations: PAL, phenylalanine ammonia-lyase; 4CL, 4-coumaroyl CoA ligase; CYP73A, *trans*-cinnamate 4-monooxygenase; CHS, chalcone synthase; CHI, chalcone isomerase; F3H, flavonone 3-hydroxylase; F3′H, flavonoid 3′-monooxygenase; FLS, flavonol synthase; F3′5′H, flavonoid 3′,5′-hydroxylase; DFR, dihydroflavonol 4-reductase; LAR, leucoanthocyanidin reductase, ANS, anthocyanidin synthase; UFGT, anthocyanidin 3-*O*-glucosyltransferase; GSTF, glutathione S-transferase; UGT, UDP-glycosyltransferase; ERD6L, sugar transporter ERD6-like; GRP2, glycine-rich RNA-binding protein 2; VTC, GDP-L-galactose phosphorylase/guanyltransferase; HSP, heat-shock protein; CST12, cytosolic sulfotransferase 12; lecRLK3, lectin S-receptor-like serine/threonine-protein kinase 3; BCCP, biotin carboxyl carrier protein of acetyl-CoA carboxylase.

INTRODUCTION

As a visible trait of plants, color is of great significance to plant growth and development. Anthocyanins are water-soluble pigments belonging to the flavonoid family that mainly contribute to red and blue coloring. In recent years, a series of studies on anthocyanins has revealed their contribution to typical coloring in plants. For example, anthocyanins are involved in color formation in the skin of *Ziziphus jujuba* (Zhang Q. et al., 2020) and *Malus domestica* (Fang et al., 2019a,b). Anthocyanins are responsible for flower color and petal blotches in *Paeonia suffruticosa* (Gu et al., 2019), *Salvia miltiorrhiza* (Jiang et al., 2020), and *Pleione limprichtii* (Zhang et al., 2019). Anthocyanins also contribute to fruit color, such as in *Hylocereus spp.* (Fan et al., 2020) and *Morus atropurpurea* (Huang et al., 2020). Moreover, anthocyanin has various other functions in plants, including in the protection against abiotic stresses, resistance to pathogens, tolerance to environmental stresses, and pollinator attraction (Schaefer et al., 2004; Zhang et al., 2015). Anthocyanins are also used as food colorants (Shahid et al., 2013). Compared with other food colorants, anthocyanins have strong antioxidant and free radical-scavenging properties, and thus they are associated with a variety of health benefits, and foods rich in anthocyanins are popular among consumers (Alvarez-Suarez et al., 2014; Smeriglio et al., 2016). However, artificially synthetic anthocyanins are inevitably used, and these contain various metal ions that are harmful to human health (Carocho et al., 2014). Thus, natural anthocyanins have great market potential. Anthocyanin-rich plant extracts have been used as substitutes for synthetic pigments in a wide variety of products; for instance, in dairy products, including cheese, fermented milk, and milk (Kitts and Tomiuk, 2013; de Mejia et al., 2015; Pineda-Vadillo et al., 2017).

Anthocyanins are synthesized *via* the flavonoid pathway and are regulated by many structural genes, including *CHS*, *CHI*, *DFR*, *F3H*, *UFGT*, and *ANS* (Hichri et al., 2011). Moreover, some transcription factors (TFs) affect the synthesis of anthocyanins by binding to the promoter regions of structural genes, such as bZIP, MYB, bHLH, MADS-box, and WD (An et al., 2017; Lu et al., 2018; Jian et al., 2019). Among them, the regulatory function of the MBW (MYB-bHLH-WD) protein complex consisting of MYB, bHLH, and WD in anthocyanin synthesis is well stablished (Feng et al., 2020). Besides, some zinc-finger proteins play a key role in the accumulation of anthocyanins (Shi et al., 2018; Bai et al., 2019). Some external environmental factors can also induce the formation of anthocyanins, such as UV-B, temperature, and drought (An et al., 2019; Yang et al., 2020).

Common sources of natural anthocyanins mainly include red grape marc, black carrot, purple potato, and onion (Ersus Bilek et al., 2017; Mourtzinos et al., 2018), but these cannot meet the increasing demand for natural food pigments. Therefore, finding new extraction sources of natural anthocyanins is necessary. The flowers of *Camellia japonica*, as one of the top 10 national flowers in China, are well-known for their striking colors. The flower color of *C. japonica* is undoubtedly caused by the hyperaccumulation of anthocyanins, and thus this flower might constitute a new natural anthocyanin source (Pan et al., 2020). However, the anthocyanins in the petals of *C. japonica* have not been sufficiently identified or quantified. In this study, we detected and quantified the composition and content of anthocyanins and clarified the regulatory network of anthocyanin biosynthesis in *C. japonica* using an integrated metabolome and transcriptome analysis. This study identified the candidate genes regulating the mechanism of *C. japonica* petal coloration, thus providing a foundation for the metabolic engineering of anthocyanin biosynthesis in the petals of *C. japonica*.

MATERIALS AND METHODS

Plant Materials

Five *C. japonica* varieties were planted in Wunao Mountain National Forest Park (31°13′44″N, 114°59′17″E) in Macheng, Hubei Province, China. All collections were approved by the head of Wunao Mountain National Forest Park in Macheng. The petal colors of the five *C. japonica* varieties, namely "Niuxiaomeiyu" (CK), "Sishaluo" (T1), "Zaohongyang" (T2), "Dahongmudan" (T3), and "Huangjiatianerong" (T4) were white, pink, deep pink, red, and crimson, respectively (**Figure 1A**). Sixty petals were sampled and pooled from three individual plants, and three replicates were set for each *C. japonica* variety. Fresh petals were collected on April 30, 2020. All samples were sampled, immediately frozen in liquid nitrogen, and then stored at –80°C.

Measurement of Total Anthocyanin Content

Approximately 0.1 g of *C. japonica* petal material was used to determine the total anthocyanin content as described by Mehrtens et al. (2005) with some modifications. Fresh petals were ground in 10 mL 95% ethanol (0.1 mol L^{-1} HCl) and extracted twice at 60°C for 1 h. The final volume of extraction solution was 25 mL (including ethanol-HCl), and the absorption at 520, 620, and 650 nm was measured using a sense microplate reader (435–301, HIDEX, Finland). The total anthocyanin content was calculated using the following formula: $Q = A\lambda \times V \times 1{,}000 / 489.72M$ (mmol/g FW), where $A\lambda$ = (A530–A620)–0.1 (A650–A620), V represents the volume of the extraction solution, and M represents the weight of the fresh sample. Ninety-five percent ethanol (0.1 mol L^{-1} HCl) was used as the blank control.

Separation and Detection of Total Flavonoids in *Camellia japonica* Petals

Petal powder was obtained by crushing freeze-dried petals in a mixer mill (MM 400, Retsch). The extraction of total flavonoids was as follows: 100 mg powder was accurately weighed and extracted in 1.0 mL 70% aqueous methanol solution overnight at 4°C. The extract was centrifuged at a high speed, and the sediment was removed. The supernatant was absorbed using a solid-phase extraction cartridge (CNWBOND Carbon-GCB SPE Cartridge, 250 mg, 3 mL; ANPEL, Shanghai, China,[1]) and filtered with a nylon syringe filter (SCAA-104, 0.22 μm pore

[1]www.anpel.com.cn/cnw

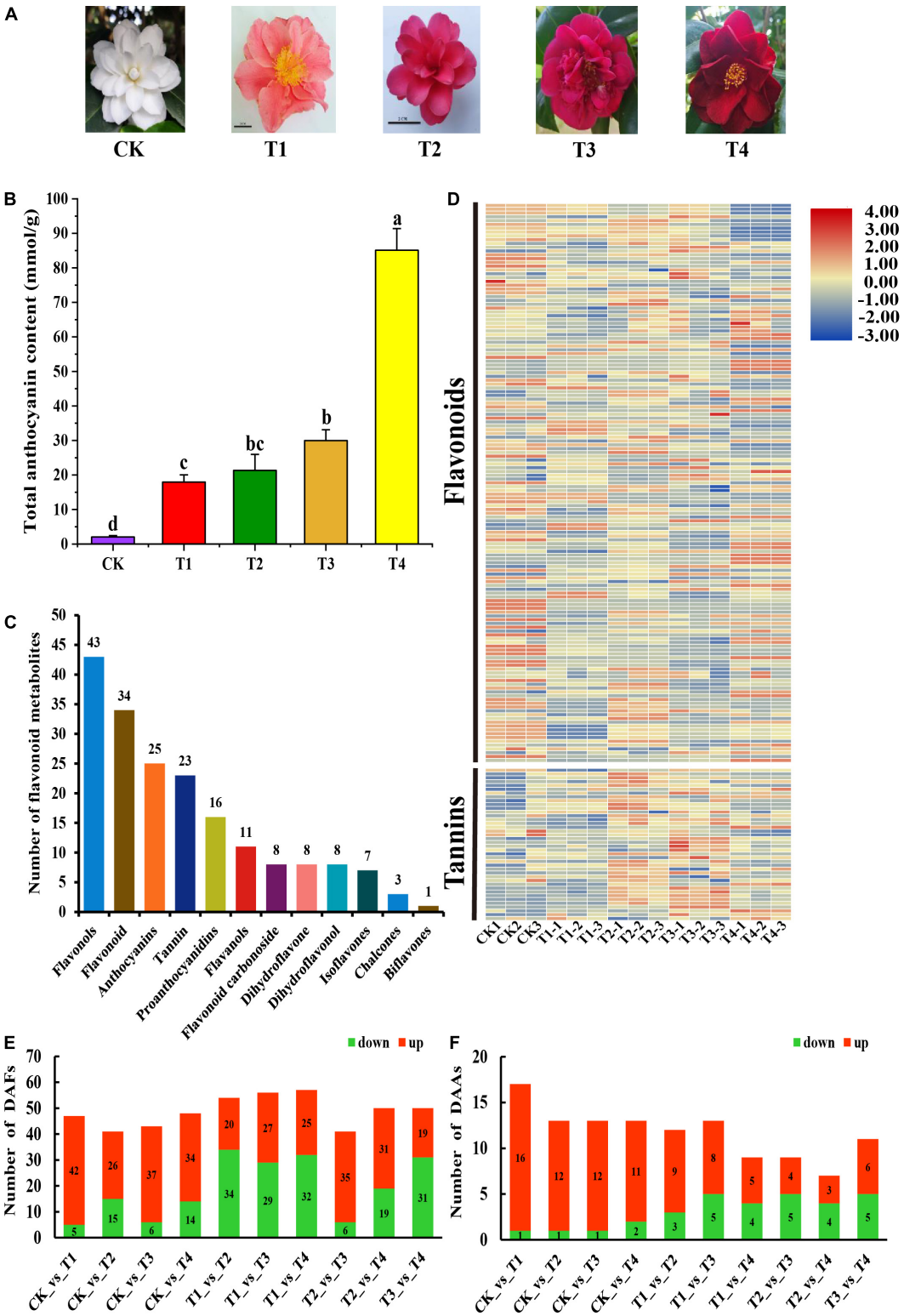


FIGURE 1 | Phenotypes Comparison of phenotypes, composition and content of anthocyanins among five *C. japonica* petals. **(A)** Phenotypes of *C. japonica* petals. **(B)** Total anthocyanin content of *C. japonica* petals. All data were shown as mean ± SE ($n = 3$). Means with different letters at each treatment represented a significant difference at $p \leq 0.05$. **(C)** Classification and statistical analysis of all metabolites detected. **(D)** The heatmap analysis of all metabolites relative content by TBtools. **(E)** Number of DAFs among all comparison units. **(F)** Number of DAAs among all comparison units.

size; ANPEL, Shanghai, China,[2]) for liquid chromatography-mass spectrometry (LC-MS) analysis. The extraction solution was analyzed by an LC-electrospray ionization-tandem mass spectrometry (ESI-MS/MS) system (HPLC, Shim-pack UFLC SHIMADZU CBM30A system, Shimadzu, Japan,[3]; MS, 4500 Q TRAP, Applied Biosystems, United States,[4]). The HPLC and MS conditions were described by Chen J. et al. (2020).

Identification and Quantitative Analysis of Metabolites

The metabolites eluted from the HPLC were monitored for each period by scheduled multiple reaction monitoring (MRM). The MRM signals were converted and analyzed using the software Analyst 1.6.3 (Metware Biotechnology Co., Ltd., Wuhan, China). Metabolite identification and quantification in this study were conducted following the commercially available standard Metabolites Database (Metware Biotechnology Co., Ltd., Wuhan, China) (Jiang et al., 2020) and public metabolite databases (Wishart et al., 2013; Zhu et al., 2013). The identified peak area of each compound was used for principal component analysis and orthogonal partial least squares-discriminant analysis. The differentially accumulated flavonoids (DAFs) and differentially accumulated anthocyanins (DAAs) were identified by R statistical software, and the screening criteria were as follows: metabolites with fold change ≥ 2 and fold changes ≤ 0.5 and variable importance in the project (VIP) score ≥ 1 were considered significantly differentially accumulated.

RNA Extraction and Sequencing

All freeze-dried petals were ground on dry ice to extract the total RNA. After extraction with a Trizol reagent kit (Invitrogen, Carlsbad, CA, United States), the quality and integrity of the total RNA were assessed and checked using an Agilent 2100 Bioanalyzer (Agilent Technologies, Palo Alto, CA, United States) and RNase-free agarose gel electrophoresis, respectively. The mRNA was enriched by Oligo(dT) beads and fragmented into short fragments using fragmentation buffer. The short fragments were reversed transcribed into cDNA with random primers and synthesized to second-strand cDNA using DNA polymerase I, RNase, dNTP, and buffer. The QiaQuick PCR extraction kit (Qiagen, Venlo, Netherlands) was used to purify the cDNA fragments. Following end-repair and the addition of poly(A), and fragments were ligated to Illumina sequencing adapters. After selection by agarose gel electrophoresis and PCR amplification, all products were sequenced on the Illumina HiSeq2500 platform of Gene Denovo Biotechnology Co. (Guangzhou, China).

RNA-seq Data Analysis and Annotation

To acquire high-quality clean reads for assembly and analysis, reads obtained from the sequencing platform were filtered by fastp (version 0.18.0) (Chen et al., 2018). To reduce the influence of the rRNA in the sample on the results, the software bowtie2 (Gene Denovo Biotechnology Co., Guangzhou, China) was used to align clean reads to the ribosome database (Langmead and Salzberg, 2012). The remaining unmapped reads were used for subsequent transcriptome analysis by removing the reads that could be mapped to the ribosomes. The clean reads were mapped to the reference genome (Xia et al., 2019; Tea Plant Information Archive,[5]) using HISAT2. 2.4 (Kim et al., 2015). The mapped reads of the five groups of samples were assembled by StringTie v1.3.1, and the fragments per kilobase of transcript per million mapped reads (FPKM) value was counted to quantify the expression (Pertea et al., 2015, 2016). All transcripts were annotated from databases, including the Gene Ontology (GO) database, Kyoto Encyclopedia of Genes and Genomes (KEGG) database, NCBI non-redundant (Nr) database, Swiss-Prot protein database, and Pfam database. The genes featuring a false discovery rate (FDR) below 0.05 and absolute fold change ≥ 2 and fold change ≤ 0.5 were considered as differentially expressed genes (DEGs). The DEGs among the five group samples were identified by DESeq2 for subsequent analysis.

Weighted Gene Co-expression Network Analysis

The overlapping DEGs and DAAs were selected for co-expression network analysis *via* weighted gene co-expression network analysis (WGCNA) tools from the BMKCloud platform[6]. The content of overlapping DAAs was used as a trait for the WGCNA, and modules were obtained through WGCNA analysis with default settings. Furthermore, the correlation coefficients between the hub genes in the module and the DAAs were calculated using the OmicShare tools (Gene Denovo Co., Ltd., Guangzhou, China, [7]). The hub genes with DAAs Pearson's correlation coefficient (PCC) values ≥ 0.95 or ≤ -0.9 were selected to draw the directed interaction network diagram.

Transcription Factor Analysis

Considering the important role of TFs in anthocyanin synthesis, the TFs expressed in all samples were identified. All putative TFs were retrieved by searching against the transcription annotation file with the keyword "transcription factor" and the numbers and types of TFs were counted. Moreover, the expression levels of all TFs in the sample were analyzed by TBtools (Chen C. et al., 2020), and the PCC between these differentially expressed TFs and total anthocyanin content of samples was calculated. The TFs with $|PCC| \geq 0.9$ were selected for subsequent analysis.

Validation RNA-seq by Quantitative Real Time-PCR

Total RNA isolation was conducted by using an RNAprep Pure Plant Plus Kit (DP441, Tiangen, Beijing, China,[8]). First-strand cDNA was synthesized with HiScript III RT SuperMix for qPCR (+gDNA wiper) (R323, Vazyme, Nanjing, China,[9]) and the extracted RNA was used as the template. The LineGene 9600 Plus

[2]http://www.anpel.com.cn/

[3]www.shimadzu.com.cn/

[4]www.appliedbiosystems.com.cn/

[5]http://tpia.teaplant.org/

[6]https://international.biocloud.net/zh/dashboard

[7]http://www.omicshare.com/tools

[8]https://en.tiangen.com/about/index.html

[9]http://www.vazyme.com/Home.html

Fluorescent Quantitative PCR System (Bioer, Hangzhou, China) was used to perform quantitative real time (qRT)-PCR and AceQ Universal SYBR qPCR Master Mix (Q511, Vazyme) was selected as the fluorochrome. The primers for qRT-PCR were designed using the Integrated DNA Technologies tool[10] and are listed in Table S1. To validate the results of the RNA-seq, 33 DEGs were randomly selected for qRT-PCR, and the *18S* gene was used as an internal control (Sun et al., 2014). The results of the qRT-PCR were calculated according to the $2^{-\Delta\ \Delta\ Ct}$ comparative Ct method (Schmittgen and Livak, 2008).

Statistical Analysis

All raw RNA-seq data in the present study were uploaded to the Genome Sequence Archive in the BIG Data Center, Beijing Institute of Genomics (BIG), Chinese Academy of Sciences. The accession number is CRA003840, and the data are publicly accessible at https://bigd.big.ac.cn/gsa. Significant differences were calculated by SPSS 22.0 (SPSS Inc., Chicago, IL, United States) software for a one-way ANOVA analysis with a Turkey test, and significance was assessed at $P \leq 0.05$. All data in this paper were expressed as means ± standard deviations (SD).

RESULTS

Metabolic Differences Among the Petals of *Camellia japonica* Cultivar

To detect the metabolic mechanism of the color phenotype in the *C. japonica* petals, the total anthocyanins in the petals were measured. The total anthocyanin content of the T4 sample was 85.09 mM g^{-1}, which was significantly higher than that of the other samples (**Figure 1B**). This was followed by the T1, T2, and T3 samples, the total anthocyanin contents of which were 17.96, 21.32, and 29.98 mmol g^{-1}, respectively. The total anthocyanin content of the petals of the CK cultivar was significantly lower than the other varieties, with an anthocyanin content of 2.02 mmol g^{-1} (**Figure 1B**). These results corresponded with the color intensity of the *C. japonica* petals. Anthocyanins are a class of varied flavonoids with similar structural units. To further analyze the differences in flavonoid metabolites in the *C. japonica* petals of different colors, the anthocyanin and flavonoid metabolites were detected by UHPLC-ESI-MS/MS. A total of 187 flavonoids were identified from these five sample groups, and the relative content of all 187 flavonoids was analyzed (**Figure 1D**). The annotation, relative content, and classification of the 187 flavonoids are detailed in **Supplementary Table 2**. These flavonoids were classified into 12 categories, including flavonols, flavonoid, anthocyanins, tannin, proanthocyanidins, flavonols, flavonoid carbonoside, dihydroflavone, dihydroflavonol, isoflavones, chalcones, and biflavones. These categories contained 1 to 43 types of metabolites (**Figure 1C**).

Based on the data obtained from the metabolome, 47, 41, 43, 48, 54, 56, 57, 41, 50, and 50 types of DAFs were identified between the CK vs T1, CK vs T2, CK vs T3, CK vs T4, T1 vs T2, T1 vs T3, T1 vs T4, T2 vs T3, T2 vs T4, and T3 vs T4 groups, respectively (**Figure 1E**). Subsequently, the DAAs among these samples were obtained by analyzing the relative anthocyanin quantification results (**Figure 1F**). The results indicated that there were 16, 12, 12, and 11 types of anthocyanins that were up-regulated in T1, T2, T3, and T4, respectively, compared with CK, and these might be the key metabolites that influencing petal coloration in *C. japonica* (**Figure 1F**).

A total of 80 DAFs were identified among all samples and the relative contents of these DAFs were illustrated in a heatmap (**Figure 2A**). Among these DAFs, 18 DAAs were included, and their relative contents are detailed in **Figure 2B**. Moreover, the overlapping metabolites were identified. The results showed that 20 DAFs were shared among CK vs T1, CK vs T2, CK vs T3, and CK vs T4 (**Figure 2C**). These 20 metabolites included five flavonoids, three flavonols, one dihydroflavonols, and 11 of anthocyanins. The classification results of these 11 anthocyanins indicated that most were cyanidin glucosides and a small amount were pelargonidin glucosides and peonidin glucosides (**Figure 2D**).

Overview of the Transcriptome Data

To further study the molecular mechanism of *C. japonica* petal coloring, RNA-seq analysis was performed. Fifteen libraries were established using the five varieties of samples (three biological replicates for each variety), and a total of 39.19 to 58.82 Gb raw reads for each sample were obtained by sequencing. Clean reads were obtained after filtering low-quality reads and accounted for 99.72–99.83% of the raw data (**Supplementary Table 3**). The Q20 and Q30 values of each library were greater than or equal to 96.72 and 91.43%, respectively (**Supplementary Table 4**). The GC contents of each sample ranged from 46.21–47.73% (**Supplementary Table 4**). HISAT2. 2.4 was used to map clean reads to the reference genome, and the mapping ratio ranged between 73.40 and 76.89% (**Supplementary Table 5**). In addition, more than 86.77% of the reads were mapped to the exon region of the reference genome (**Supplementary Table 6**). Ultimately, the sequence and expression information of 57,410 genes was obtained for subsequent analysis. Heatmap analysis of the samples based on the FPKM values showed that all the biological replicates exhibited similar expression patterns, indicating the high reliability of our sequencing data (**Supplementary Figure 1A**). All these data proved that the sequencing quality was sufficient for further analysis.

Identification of DEGs in *Camellia japonica* Petals

A total of 12,531 DEGs were characterized among all samples by analyzing the FPKM value of transcripts obtained from the transcriptome data. In detail, there were 4,984, 4,374, 6,139, 5,969, 5,475, 4,371, 6,784, 5,894, 7,337, and 8,308 DEGs, respectively in the 10 comparison groups, including CK vs T1, CK vs T2, CK vs T3, CK vs T4, T1 vs T2, T1 vs T3, T1 vs T4, T2 vs T3, T2 vs T4, and T3 vs T4 (**Figure 2C**). Among these comparison groups, the T3 vs T4 group had the largest number of up-regulated and down-regulated DEGs (**Figure 2C**). Furthermore,

[10]https://sg.idtdna.com/pages

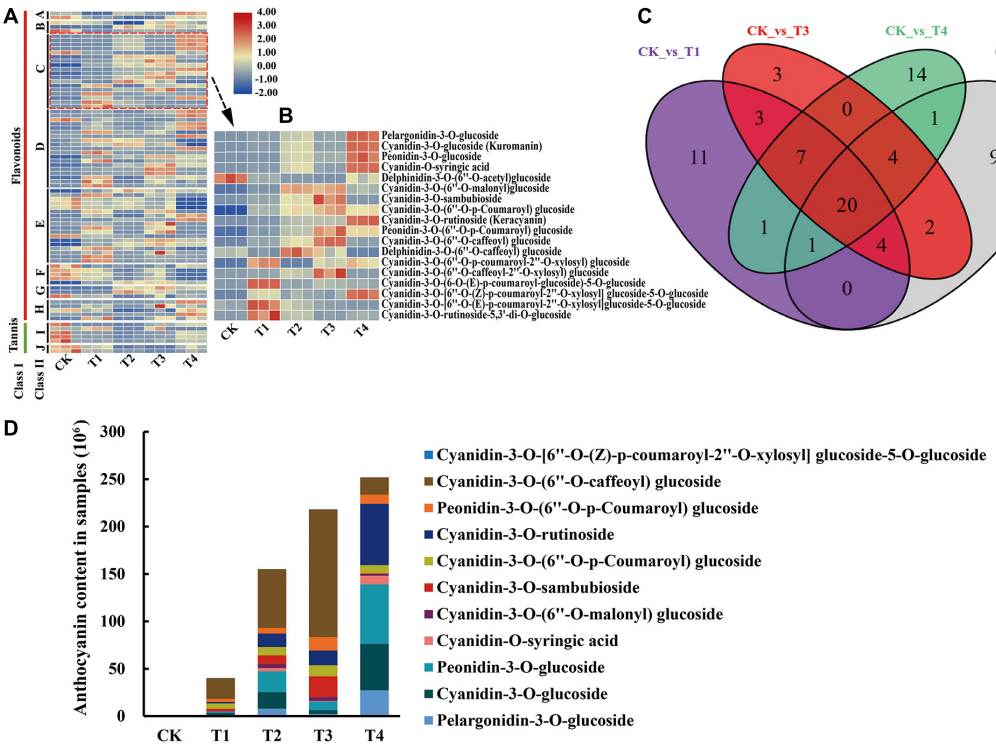


FIGURE 2 | Differentially accumulated metabolites analysis of metabolome. **(A)** The heatmap analysis of all differentially accumulated metabolites by TBtools. A, chalcones; B, dihydroflavonol; C, anthocyanins; D, flavonoid; E, flavonols; F, flavonoid carbonoside; G, flavanols; H, isoflavones; I, tannin; J, proanthocyanidins. **(B)** The heatmap analysis of all DAAs according to the relative content in samples. **(C)** Venn analysis among CK vs T1, CK vs T2, CK vs T3, and CK vs T4. **(D)** The relative content of overlapping DAAs (11 anthocyanins) in five *C. japonica* cultivar petals.

the overlap DEGs among the CK vs T1, CK vs T2, CK vs T3, and CK vs T4 groups were screened using the venn diagram function in TBtools (Chen C. et al., 2020). The results indicated that 634 genes were differentially expressed in these groups, which indicated that these DEGs might have key functions in the color expression of different petals (**Supplementary Figure 1B**).

Functional Annotation of DEGs

All DEGs were mapped to the GO database and further classified into three classifications, including cellular component, molecular function, and biological process. In the cellular component category, more than 71.2% of DEGs were enriched in cell, cell part, organelle, and membrane terms (**Figure 3A**). In the biological process category, most of the DEGs were mapped to metabolic process (4,781, 38.15%), cellular process (4,426, 35.3%), and single-organism process (3,929, 31.4%) terms (**Figure 3A**), but for molecular function, nearly 69.36% of DEGs were mapped to catalytic activity and binding terms (**Figure 3A**). For the overlapping DEGs (634 DEGs), the GO annotation has the similar results (**Supplementary Figure 2A**). For the KEGG annotation results, 12,531 DEGs among all samples were mapped to 129 KEGG pathways. Among these pathways, biosynthesis of secondary metabolites, metabolic, and phenylpropanoid biosynthesis pathways were significantly enriched (**Figure 3B**). The KEGG annotation results of the overlapping DEGs (634 DEGs) showed that most enrichment pathway were photosynthesis-antenna proteins, ascorbate and aldarate metabolism, zeatin biosynthesis and protein export (**Supplementary Figure 2B**).

The Key DEGs Responsible for the Anthocyanin Biosynthesis Pathway

To explore the difference in anthocyanin biosynthesis among the petals from the five cultivars, DEGs in the anthocyanin synthesis pathway were identified, including the phenylpropanoid biosynthesis, flavonoid biosynthesis, and anthocyanin biosynthesis pathways. The results revealed that there were 37 DEGs enriched in the anthocyanin synthesis pathway, including *CjPAL*, *Cj4CL*, *CjCYP73A*, *CjCHS*, *CjCHI*, *CjF3H*, *CjFLS*, *CjF3′5′H*, *CjDFR*, *CjLAR*, and *CjANS* (**Figure 4A**). The PCC between the expression level of these DEGs and the total anthocyanin content was further calculated. The results showed that 10 DEGs positively regulated the anthocyanin synthesis, whereas 27 DEGs negatively regulated anthocyanin synthesis. Among these DEGs, the expression level of two DEGs, namely *CjANS* (*CSS0029211*) and *Cj4CL* (*CSS0016246*), indicated a significant positive correlation with the total anthocyanin content in the samples (PCC > 0.9, **Figure 4B**), suggesting that these two DEGs may have an essential role in anthocyanin accumulation.

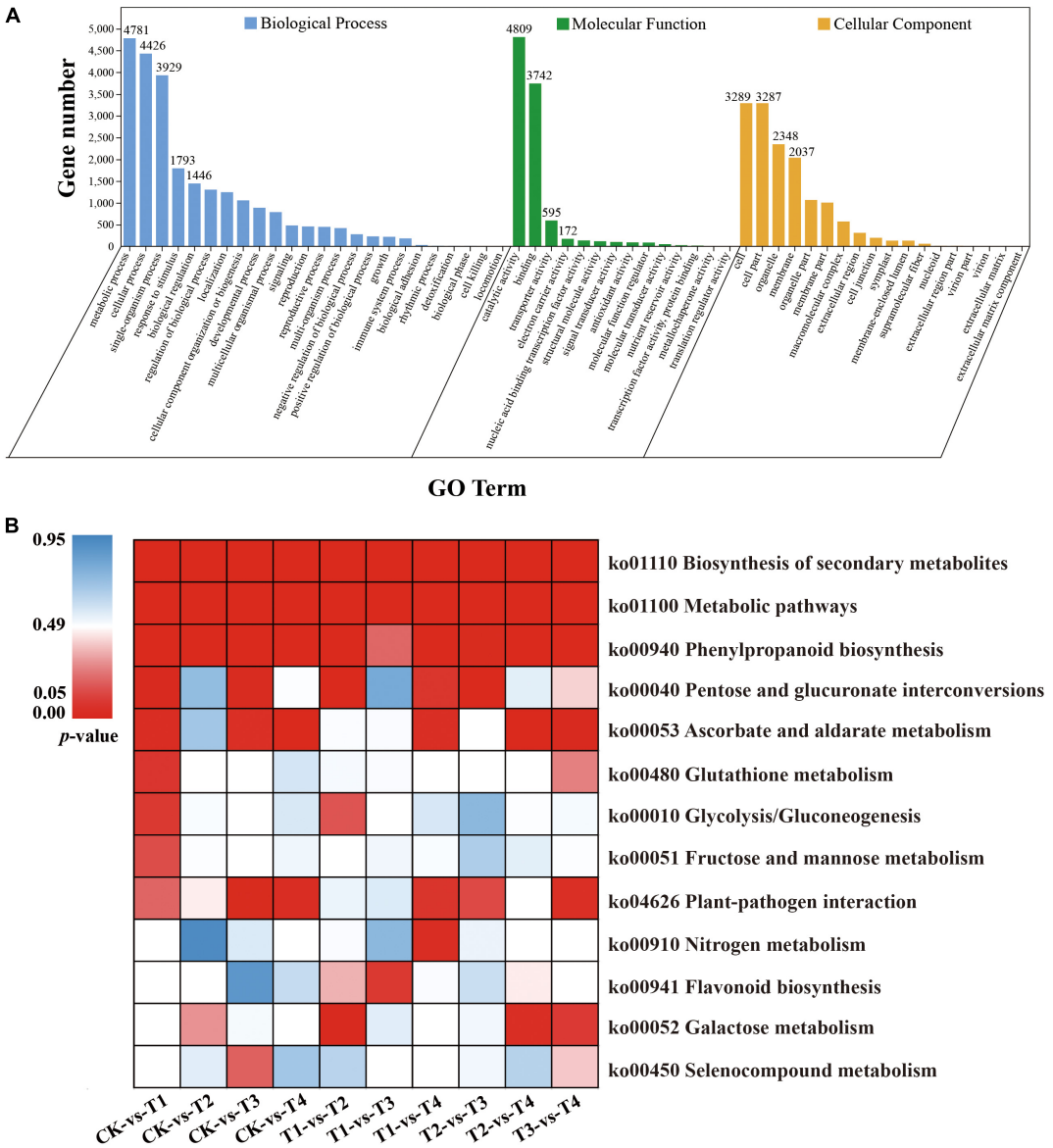


FIGURE 3 | GO and KEGG enrichment analysis of all DEGs. **(A)** GO enrichment results of all DEGs **(B)** Enrichment of the top 13 KEGG pathways of all DEGs according to the *p*-value. The red represents significant enrichment.

Identification TFs Related to Anthocyanin Biosynthesis

Transcription factors are important regulators of anthocyanin accumulation and they typically act by controlling the expression of structural genes in the anthocyanin biosynthesis pathway. In this study, a total of 883 TFs were identified by searching the transcriptome annotation results. The classified results indicated that most of these TFs belonged to the *AP2/ERF*, *bHLH*, *MYB*, *WRKY*, *GATA*, *NAC*, and *bZIP* family (**Supplementary Figure 3A**). Furthermore, the differentially expressed TFs were characterized by analyzing their FPKM values (**Supplementary Figure 3A**). By calculating the PCC between the expression level of these TFs and the total anthocyanins content, 19 key TFs ($|PCC| \geq 0.9$) involved in the accumulation of anthocyanins were identified, including six negative regulators and 13 positive regulators. These six negative regulators, including one *MYB*, two *bHLHs*, one *WRKY*, one *TCP*, and one *AP2/ERF* gene, likely act as repressors in anthocyanin synthesis. However, the 13 positive regulators, including one *GATA*, two *MYBs*, three *bHLHs*, two *TCPs*, three *AP2/ERFs*, one *RAV*, and one *NAC* gene, might act as promotors of anthocyanin accumulation (**Table 1**). Furthermore, the expression level of these key TFs was illustrated in a heatmap, and the results showed that the positive regulators exhibited the highest expression level in T4, whereas the negative regulators exhibited the highest expression level in T1 (**Supplementary Figure 3B**).

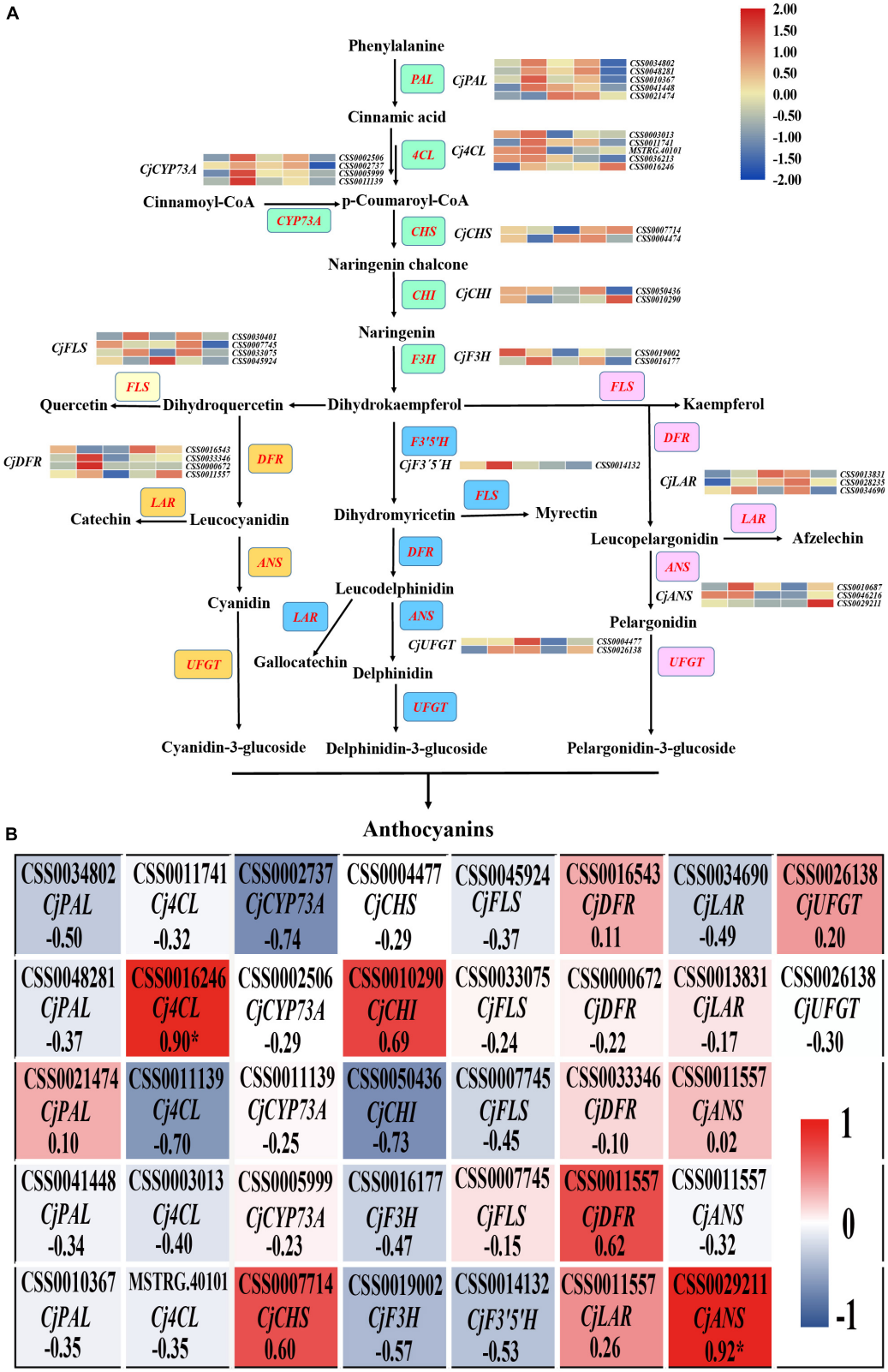


FIGURE 4 | The analysis of DEGs in anthocyanin biosynthesis pathway. **(A)** The identification of all DEGs in anthocyanin biosynthesis pathway. The red font represents DEGs and the colored boxes represent the different branches of anthocyanin synthesis. **(B)** The heatmap analysis of all DEGs in anthocyanin biosynthesis pathway according to the FPKM value.

TABLE 1 | The candidate TFs involved in anthocyanin accumulation.

Transcription factor	Gene ID	Annotation	Correlation with total anthocyanin	*p*-value
GATA	CSS0049672	GATA transcription factor 1	0.93	0.0212
MYB	CSS0015300	MYB family transcription factor	−0.97	0.0073
	CSS0014780	transcription factor R2R3-MYB1	0.93	0.0206
	CSS0045979	PREDICTED: transcription factor MYB86	0.93	0.0209
bHLH	MSTRG.36794	transcription factor bHLH51-like	−0.96	0.0110
	CSS0025568	PREDICTED: transcription factor bHLH106-like	−0.94	0.0157
	CSS0019162	PREDICTED: transcription factor bHLH137	0.97	0.0060
	CSS0029144	PREDICTED: transcription factor bHLH71	0.97	0.0071
	CSS0018697	PREDICTED: transcription factor bHLH11	0.95	0.0121
WRKY	CSS0016070	PREDICTED: WRKY transcription factor 44	−0.95	0.0124
TCP	CSS0042452	PREDICTED: transcription factor TCP20-like	−0.90	0.0353
	CSS0010873	PREDICTED: transcription factor TCP23-like	0.96	0.0094
	MSTRG.27463	transcription factor TCP4-like	0.95	0.0142
AP2/ERF	CSS0018330	transcription factor APETALA2	−0.92	0.0256
	CSS0002584	PREDICTED: ethylene-responsive transcription factor ERF023	0.98	0.0030
	CSS0009504	PREDICTED: ethylene-responsive transcription factor ERF014-like	0.98	0.0046
	CSS0048980	PREDICTED: AP2/ERF and B3 domain-containing transcription factor At1g50680-like isoform X2	0.95	0.0141
RAV	CSS0037887	RAV transcription factor	0.96	0.0090
NAC	CSS0036947	NAC transcription factor 037	0.99	0.0011

Identification of Anthocyanin Synthesis-Related DEGs by WGCNA

To investigate the gene regulatory network of anthocyanin accumulation in the *C. japonica* petals, WGCNA was conducted using the FPKM values of the overlapping DEGs (634 DEGs) and DAAs (11 anthocyanins) as source data. Five modules were identified in the cluster dendrogram, donated as turquoise, brown, blue, yellow, and gray modules (**Figure 5A**). Analysis of the module-anthocyanin relationships revealed that the brown, turquoise, and yellow modules have high PCCs with the content of anthocyanins (**Figure 5B**). To explore the expression patterns of the genes in these modules, heatmaps were conducted using the FPKM value of the genes in these modules. The heatmap results of the brown module proposed that the co-expressed genes had a higher expression level only in T3, showing significant sample specificity (**Supplementary Figure 4A**). In the turquoise module, the co-expressed genes were highly expressed in CK (**Supplementary Figure 4B**), whereas in the yellow module, the co-expressed genes had the highest expression level in T3 and the lowest expression level in CK (**Supplementary Figure 4C**). Moreover, a directed interaction network diagram was constructed using the PCCs between the hub genes in the brown, turquoise, and yellow module (PCC $\geq$ 0.95 or PCC $\leq$ −0.9) and the content of the 11 anthocyanins. As shown in **Figure 5C**, 16 genes in the modules were selected to construct the regulatory network, including two *CjGSTFs*, three *CjUGTs*, *CjGT*, two *CjERD6Ls*, *CjGRP2*, two *CjVTCs*, *CjHSP*, two *CjBCCPs*, *CjlecRLK3*, and *CjCST12*. All the anthocyanins were positively regulated by these genes, except for cyanidin-3-*O*-(6″-*O*-p-Coumaroyl) glucoside. The above results indicate that these 16 genes may play essential roles in anthocyanin synthesis in *C. japonica*.

Verification of the Results in RNA-seq by qRT-PCR

To validate the accuracy of the RNA-seq data, 33 genes were randomly selected for qRT-PCR. The relative expression levels of these 33 genes were normalized to the expression of *18S* rRNA (**Supplementary Figure 5**). Further analysis revealed that the expression levels of these genes were significantly correlated with the RNA-seq results (**Supplementary Figure 6**, $R^2 > 0.75$), indicating that the RNA-seq data were credible and accurate.

DISCUSSION

Anthocyanin Identification in the *Camellia japonica* Petals

Anthocyanins are key metabolites for coloration in plant organs, contributing to the purple leaves in *Brassica rapa* and *Brassica oleracea* (Zhang et al., 2014b; Horiuchi et al., 2020), and the purple skin in *Capsicum annuum* and *Solanum melongena* (Zhang et al., 2014c; Tang et al., 2020). Considering the health value of anthocyanins, foods rich in anthocyanins and anthocyanin extracts have broad market application prospects, and finding new sources of anthocyanin extraction has become a new problem. Similar to the flowers of other species, the petals of *C. japonic*a rich in anthocyanins. Thus, *C. japonica* may be served as a new source of natural anthocyanins extraction. The content of anthocyanins in the petals of *C. japonica* was determined in this study, and the results showed that the color intensity of the *C. japonica* petals was changed with the different anthocyanin contents (**Figure 1**). Here, in the red to dark-red petals of *C. japonica*, the anthocyanin content was greater than 20 mmol/g and even reached 85 mmol/g, which is significantly

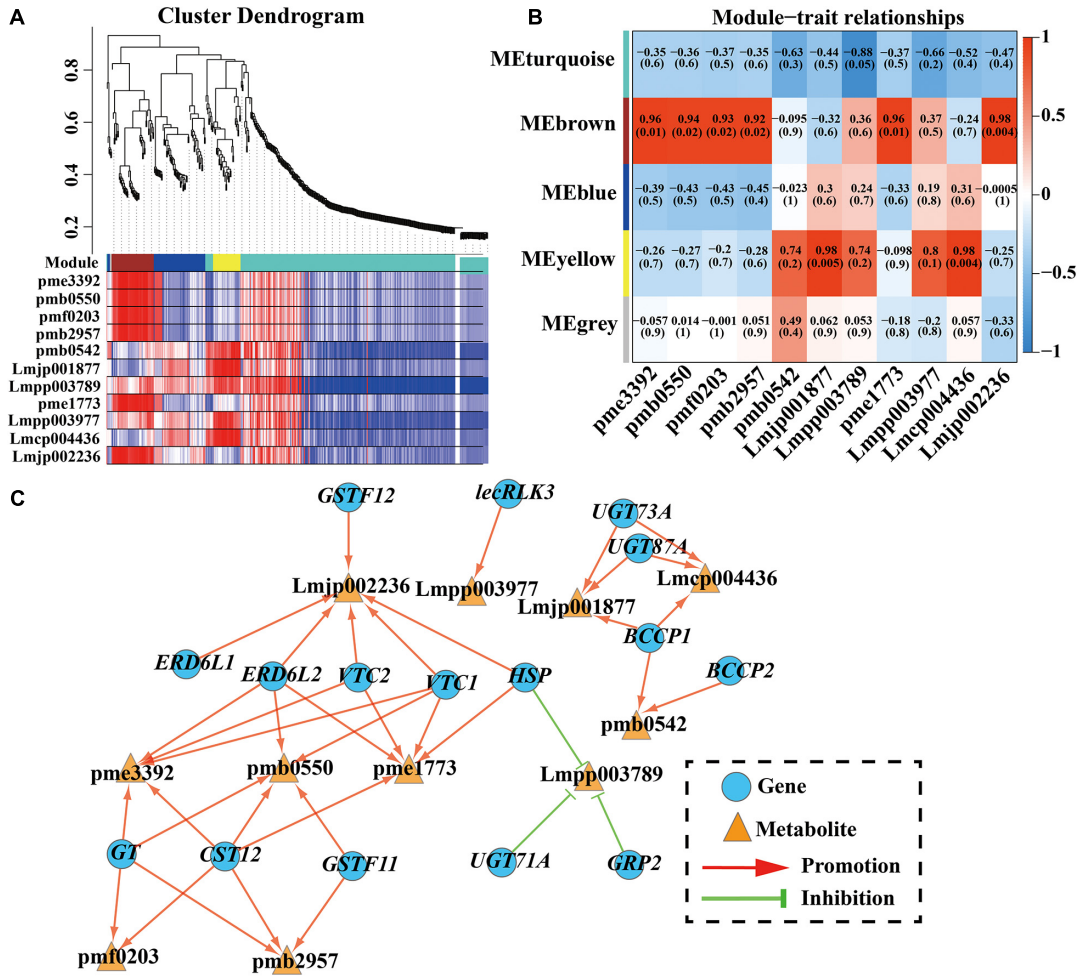


FIGURE 5 | WGCNA results of 634 DEGs (with 11 DAAs PCC ≥ 0.95 or ≤ –0.90). **(A)** Hierarchical clustering tree (cluster dendrogram) results showed 5 expression modules, labeled with different colors. **(B)** Module-anthocyanin relationship analysis. The value inside each box represents Pearson's correlation coefficient between the module with anthocyanin, and the number in each parentheses represents *p*-value. The color scale on the right represents the degree of correlation between modules and anthocyanins and the red represent high correlation. **(C)** The directed interaction network diagram between DEGs and anthocyanin. Red line represents the promotion of synthesis, while green "T" represents inhibition of accumulation. Blue circles represent DEGs and orange triangles represent anthocyanin. pme3392, pelargonidin-3-*O*-glucoside; pmb0550, cyanidin-3-*O*-glucoside; pmf0203, peonidin-3-*O*-glucoside; pmb2957, cyanidin-*O*-syringic acid; pme1773, cyanidin-3-*O*-(6″-*O*-malonyl) glucoside; Lmjp002236, cyanidin-3-*O*-sambubioside; Lmpp003789, cyanidin-3-*O*-(6″-*O*-p-Coumaroyl) glucoside; pmb0542, cyanidin-3-*O*-rutinoside; Lmjp001877, peonidin-3-*O*-(6″-*O*-p-Coumaroyl) glucoside; Lmpp003977, cyanidin-3-*O*-(6″-*O*-caffeoyl) glucoside; Lmcp004436, cyanidin-3-*O*-[6″-*O*-(Z)-p-coumaroyl-2″-*O*-xylosyl] glucoside-5-*O*-glucoside.

higher than the anthocyanin content of the peel extract of *S. melongena* (Todaro et al., 2009; Boulekbache-Makhlouf et al., 2013), indicating that *C. japonica* may have better application prospects as a natural food pigment.

There are many species of anthocyanins in colorful plants, of which six are well-known, namely cyanidin, delphinidin, peonidin, malvidin, pelargonidin, and petunidin (De Rosso and Mercadante, 2007). The molecular structures and chemical characteristics of these anthocyanins have been extensively investigated. Generally, cyanidin contributes to red-purple, delphinidin contributes to blue-red or purple, and pelargonidin contributes to orange and red (Khoo et al., 2017). In our study, only four aglycones of anthocyanin were detected in the *C. japonica* petals, namely cyanidin, pelargonidin, delphinidin, and peonidin. Malvidin and petunidin were not detected, indicating that these are not involved in the coloration of *C. japonica* petals. These findings corroborate the results of Xue et al. (2019), who demonstrated that cyanidin, pelargonidin, and delphinidin were the main components of the red flowers in strawberries. Interestingly, the results of this study differ from those of Zhou et al. (2020), who found that cyanidin, malvidin, petunidin, and pelargonidin were the main anthocyanin components in the pink flowers of *Camellia sinensis* (same genus species), which may be the results of domestication because *C. sinensis* is used as a drink rather than a decoration plant.

Anthocyanins isolated in nature are highly unstable and susceptible to degradation, resulting in the loss of biological activity and discoloration. Therefore, anthocyanins in organisms

are bound to glycosides to enhance their stability (Ongkowijoyo et al., 2018). In this study, five types of glycoside were detected by analyzing the molecular structure of the anthocyanins, including two monoglycosides (glucose and arabinose) and three polyglycosides (di-glucoside, rutinose, and sambubioside). The data showed that anthocyanins containing polyglycosides accounted for 17.14–39.11% of the total anthocyanins in the *C. japonica* samples. Moreover, anthocyanins in the form of syringic acid (cyanidin-*O*-syringic acid) accounted for a maximum of 3.25% of total anthocyanins, which have also been reported to result in the rubellis mutant tepals of *Michelia maudiae* (Lang et al., 2019). However, in strawberries, although the content of anthocyanins varies among species, most anthocyanins in strawberries (more than 94%) are monoglycosides (Kelebek and Selli, 2011). The binding positions of the monoglycoside and polyglycoside anthocyanins are different. According to the report of Du et al. (2018), monoglycoside anthocyanins naturally occur at the 3- position, but when more than one glucose, they can also be associated with the 5-hydroxyl position. Anthocyanin degradation begins with the separation of a sugar molecule from the anthocyanin structure, and glycan substituents of anthocyanins are more readily lost from the 5- position than the 3- position (De Rosso and Mercadante, 2007). Therefore, polyglycoside anthocyanins exhibit stronger stability than monoglycoside anthocyanins. These findings suggest that polyglycoside anthocyanins are abundant and potentially of greater commercial value in *C. japonica*.

Cyanidin-3-*O*-(6″-*O*-malonyl) Glucoside Content Contributes to Flower Color Variation in *Camellia japonica*

Differences in flower color are typically caused by different anthocyanin types and contents. High anthocyanin content in petals can make flowers darker in color. For example, anthocyanins are not present in white flower mutants of *Gentiana triflora*, but do accumulate at high levels in the blue flowers of *G. triflora* (mainly delphinidin) (Nakatsuka et al., 2005). In this study, 11 DAAs among CK vs T1, CK vs T2, CK vs T3, and CK vs T4 were identified according to the anthocyanin content in each sample, which indicated that these 11 anthocyanins contribute to the different petal colors in *C. japonica*. By analyzing the relative contents of these anthocyanins in the five groups, the results found that the content of cyanidin-3-*O*-(6″-*O*-malonyl) glucoside varied greatly among T1, T2, T3, and T4. However, cyanidin-3-*O*-(6″-*O*-malonyl) glucoside was almost undetectable in CK, which indicated that cyanidin-3-*O*-(6″-*O*-malonyl) glucoside is the main source of pink, red, and crimson phenotypes (**Figure 2D**). This finding supports the results of Cheng et al. (2014), who proposed that cyanidin-3-glucoside is the primary component in the red flowers of *Prunus persica*. Reports on *C. sinensis* document that cyanidin *O*-syringic acid, petunidin 3-*O*-glucoside, and pelargonidin 3-*O*-β-d-glucoside accumulate in pink flower but not in white flower, which indicates that these compounds play important role in the color expressing of the pink flower (Zhou et al., 2020). Moreover, cyanidin-3-*O*-rutinoside, peonidin-3-*O*-glucoside, cyanidin-3-*O*-glucoside, and pelargonidin-3-*O*-glucoside greatly accumulated in T1, T2, T3, and T4, implying that these anthocyanins play important roles in deepening the color of *C. japonica* petals (**Figure 2D**).

Key Structural Genes Responsible for Anthocyanin Synthesis in *Camellia japonica* Petals

The anthocyanin biosynthesis pathway was reported as early as 1995 (Holton and Cornish, 1995), and the functions of key genes in the pathway have been explored in recent studies. For example, Zhang Y. et al. (2020) found that the six-base insertion mutation of *PsF3′H* upregulated the expression of *PsFLS*, resulting in difference in anthocyanin accumulation and acyanic flowers in *P. suffruticosa*. Considering the difference in anthocyanin contents among these *C. japonica* petals, two genes (*CjANS*, *CSS0029211*; *Cj4CL*, *CSS0016246*) among the 37 DEGs in the anthocyanin synthesis pathway were identified. We speculate that these two genes may play key roles in anthocyanin accumulation (**Figure 4**). As one of the key enzyme genes in the downstream of the anthocyanin biosynthesis pathway, *ANS* catalyzes the transformation of leucoanthocyanidin to a colored anthocyanidin before the final glycosylation steps (Hichri et al., 2011). Li et al. (2019) reported that overexpression of *SmANS* enhanced anthocyanin accumulation and resulted in the purple-red phenotype of *S. miltiorrhiza*, indicating the important function of *ANS* in anthocyanin biosynthesis. In *Camellia nitidissima* Chi., Zhou et al. (2017) found that *CnANS* regulated the accumulation of flavonols and anthocyanin, thus contributing to the golden flower of *C. nitidissima*. Moreover, the expression of *CnCHS*, *CnF3H*, and *CnFLS* also play critical roles in the formation of golden flowers (Zhou et al., 2017). Reports in *Arabidopsis thaliana* documented that the anthocyanin content reduced by 70% in the mutant of *At4CL3* compared to the wild type, which confirmed the function of *4CL* in anthocyanin accumulation (Li et al., 2015). Therefore, we speculate that the differential expression of *CjANS* and *Cj4CL* is an important contributor to the petal color diversity in *C. japonica* in this study.

Transcription Factors Involved in Anthocyanin Synthesis in *Camellia japonica* Petals

Except to structural genes, TFs also play important roles in anthocyanin biosynthesis, such as *MYB*, *bHLH*, *WD protein*, and *MADS-box* (Zhang et al., 2014a). In *P. persica*, *PpNAC1* promotes anthocyanin accumulation by activating the transcription of *PpMYB10.1*, while the expression of *PpNAC1* is repressed by *PpSPL1* (Zhou et al., 2015). The upstream regulatory network of *SPL* was explored by Wang et al. (2020), who confirmed that the function of *SPL* in regulating anthocyanin biosynthesis was targeted by *miR156*, *miR858*, and *miR160h*. Thus, anthocyanin biosynthesis in plants involves a complex network that is not only the function of a single structural gene or TF but also the co-functioning of multiple TFs and structural genes (Jian et al., 2019). In this study, a large number of TFs were identified by

analyzing the transcriptome annotation results. Based on the expression level of TFs obtained from the transcriptome data, 19 TFs (detailed in Table1 and **Supplementary Figure 4**) were found to highly associate with the total anthocyanin content ($|PCC| \geq 0.9$), and these TFs may have an essential function in the phenotypic expression of petal color. Interestingly, among these TFs, three *CjMYBs* showed two different expression patterns. The expression level of *CjMYB* (*CSS0015300*) was highest in CK, followed by T1, T2, T3, and T4, which was contrary to the total anthocyanin content trend. Conversely, the expression of two *CjMYBs* (*CSS0014780* and *CSS0045979*) exhibited a similar trend to the total anthocyanin content in the *C. japonica* petals, indicating that *CjMYB* (*CSS0015300*) negatively regulated anthocyanin accumulation, whereas *CjMYBs* (*CSS0014780* and *CSS0045979*) may positively regulate anthocyanin accumulation in *C. japonica*. In plants, MYB often forms protein complexes with bHLH and WD to participate in anthocyanin biosynthesis rather than regulate anthocyanin biosynthesis directly (Feng et al., 2020). Subsequently, two negatively correlated *CjbHLHs* (*MSTRG.36794* and *CSS0025568*) and three positively correlated *CjbHLHs* (*CSS0019162*, *CSS0029144*, and *CSS0018697*) were identified. Whether these *CjbHLH* TFs interact with *CjMYB* to regulate anthocyanin biosynthesis in *C. japonica* remains to be further investigated.

Other Key Candidate Genes Related to Anthocyanin Synthesis in *Camellia japonica* Petals

The formation of anthocyanins depends on glycosylation, hydroxylation, acylation, and methoxylation to maintain stability, and this process is controlled by some transporters and other proteins (Jiang et al., 2020). In this study, key genes involved in the coloring of *C. japonica* petals were identified by WGCNA, and 16 genes exhibiting the highest correlation were selected to construct a directed interaction network diagram. The results showed that these 11 anthocyanins were regulated by one to six genes and only cyanidin-3-*O*-(6″-*O*-p-Coumaroyl) glucoside was negatively regulated. The annotation results of the 16 genes confirmed the regulation of certain genes in anthocyanin synthesis has been reported in other species. Wei et al. (2019) confirmed that *CsGSTF11* was promoted by *CsMYB75* and resulted in anthocyanin hyperaccumulation in purple tea. In this study, two *CjGSTFs* (*CjGSTF11* and *CjGSTF12*) identified by WGCNA may have similar functions in the hyperaccumulation of anthocyanin in *C. japonica* petals. The glycosylation of anthocyanins was catalyzed by *UGTs* (the members of GT family), which is the final step in anthocyanin biosynthesis and leads to diverse anthocyanin molecules in *A. thaliana* (Li et al., 2017). The study of UGTs in *A. thaliana* showed that *UGT79B2* and *UGT79B3* elevate various stress tolerance by modulating anthocyanin accumulation (Li et al., 2017). In the WGCNA, three *UGTs* (*CjUGT73A*, *CjUGT87A*, and *CjUGT71A*) and one *CjGT* were identified to target two, two, one, and four types of anthocyanins, respectively, indicating that these *GTs* are responsible for anthocyanin glycosylation in *C. japonica* petals. Moreover, one zinc-finger protein (*CjGRP2*) and two sugar transporters (*CjERD6L1* and *CjERD6L2*) were identified, and their functional regulation in anthocyanin synthesis in other species has been reported (Jeong et al., 2010; Shi et al., 2018). Taken together, these results indicate that these 16 genes might play important roles in the color diversity in the *C. japonica* petals.

CONCLUSION

In this study, metabolome and transcriptome analyses were used to identify key anthocyanins and genes responsible for *C. japonica* petal coloration. Overall, a total of 187 flavonoid metabolites were detected in *C. japonica* petals, including 25 anthocyanins. Cyanidin-3-*O*-(6″-*O*-malonyl) glucoside was the main anthocyanin component in the *C. japonica* petals, while cyanidin-3-*O*-rutinoside, peonidin-3-*O*-glucoside, cyanidin-3-*O*-glucoside, and pelargonidin-3-*O*-glucoside were responsible for the color intensity of the *C. japonica* petals. Moreover, 37 DEGs (especially *ANS* and *4CL*) in the anthocyanin biosynthesis pathway, 19 differentially expressed TFs (such as *GATA*, *MYB*, *bHLH*, *WRKY*, and *NAC*), and 16 other regulators (including two *GSTFs*, four *GTs*, two sugar transporters, and one zinc-finger protein) were identified as candidate regulators contributing to the color diversity of *C. japonica* petals. The results of this study provide valuable information and new insights for further evaluation on the genetic diversity of *C. japonica*.

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: Genome Sequence Archive in National Genomics Data Center, China National Center for Bioinformation/Beijing Institute of Genomics, Chinese Academy of Sciences; (Wang et al., 2017; CNCB-NGDC Members and Partners, 2021) CRA003840.

AUTHOR CONTRIBUTIONS

FX and MyF designed the research. MyF wrote the manuscript. MyF, XY, JrZ, and JbY collected the experimental materials. XY, LW, and YT accomplished the experiment involved in this manuscript. XyY helped for the data analyzing. SyC, WwZ, and YlL offered the key idea and scientific guidance for this research. FX revised the manuscript critically for important intellectual content. MyF sorted out and analyzed all the materials in this research. All authors listed here contributed and approved the manuscript.

FUNDING

This study was funded by the Special Projects for Technological Innovation in Hubei Province (No. 2019ABA113), the Key

Research and Development Program of Hubei Province (No. 2020BBA043), and the Demonstration Project of Forestry Science and Technology funded by the Central Government of China [The Plan of Forestry Department of Hubei Province, Grant No. (2017) TG08]. The funding bodies did not play any role in the design of the study and collection, analysis, and interpretation of data and in writing the manuscript.

ACKNOWLEDGMENTS

We would like to thank the staff of Wunao Mountain National Forest Park in Macheng for their help during the field experiments. We would like to thank the head of Wunao Mountain National Forest Park in Macheng for our permission to collect all samples in this experiment. We would like to thank Wuhan Metware Biotechnology Co., Ltd. (especially engineer Xiaorui Ma) and Guangzhou Genedenovo Biotechnology Co., Ltd., for assisting in bioinformatics analyzing and sequencing, respectively. We would also like to thank LetPub (www.letpub.com) for its linguistic assistance during the preparation of this manuscript.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2021.685136/full#supplementary-material

Supplementary Figure 1 | Sample correlation analysis and comparative analysis of DEGs. **(A)** Sample relation analysis using FPKM value of all genes. Red represents a higher similarity. **(B)** Venn analysis among CK vs T1, CK vs T2, CK vs T3, and CK vs T4. **(C)** Number of DEGs among all comparison units.

Supplementary Figure 2 | The GO and KEGG enrichment results of overlapping DEGs (634 DEGs). **(A)** GO enrichment results of overlapping DEGs. **(B)** The enrichment results of the top 15 KEGG pathways according to the p value. Red presents significant enrichment.

Supplementary Figure 3 | The identification results of TFs related to anthocyanin accumulation. **(A)** The number and type of TFs and differentially expressed TFs. **(B)** The heatmap analysis of TFs involved in anthocyanin accumulation according to FPKM value. The red represents high expression level, while blue represents low expression level.

Supplementary Figure 4 | The expression level of DEGs in different modules. **(A)** The expression trend of genes in brown module among all samples. **(B)** The expression trend of genes in turquoise module among all samples. **(C)** The expression trend of genes in yellow module among all samples.

Supplementary Figure 5 | qRT-PCR validation of gene expression level in the transcriptome. The color of columns represents different samples. The height of column represents qRT-PCR results of genes ($2^{-\Delta\Delta Ct}$), while the line represents FPKM value of genes.

Supplementary Figure 6 | Correlation analysis of the results between RNA-seq and qRT-PCR. The qRT-PCR ($2^{-\Delta\Delta Ct}$) and RNA-seq results were normalized by $\log_2$.

Supplementary Table 1 | Primers used in qRT-PCR.

Supplementary Table 2 | The classification and quantification results of all detected metabolites.

Supplementary Table 3 | Summary of sequencing data.

Supplementary Table 4 | Detection of sample alignment rate.

Supplementary Table 5 | Evaluation of sample alignment area.

Supplementary Table 6 | Assessment of sample alignment area.

REFERENCES

Alvarez-Suarez, J. M., Giampieri, F., Tulipani, S., Casoli, T., Di Stefano, G., González-Paramás, A. M., et al. (2014). One-month strawberry-rich anthocyanin supplementation ameliorates cardiovascular risk, oxidative stress markers and platelet activation in humans. *J. Nutr. Biochem.* 25, 289–294. doi: 10.1016/j.jnutbio.2013.11.002

An, J., Wang, X., Zhang, X., Bi, S., You, C., and Hao, Y. (2019). *MdBBX22* regulates UV-B-induced anthocyanin biosynthesis through regulating the function of *MdHY5* and is targeted by *MdBT2* for 26S proteasome-mediated degradation. *Plant Biotechnol. J.* 17, 2231–2233. doi: 10.1111/pbi.13196

An, J. P., Qu, F. J., Yao, J. F., Wang, X. N., You, C. X., Wang, X. F., et al. (2017). The bZIP transcription factor *MdHY5* regulates anthocyanin accumulation and nitrate assimilation in apple. *Hortic. Res.* 4:17023. doi: 10.1038/hortres.2017.23

Bai, S., Tao, R., Tang, Y., Yin, L., Ma, Y., Ni, J., et al. (2019). BBX16, a B-box protein, positively regulates light-induced anthocyanin accumulation by activating *MYB10* in red pear. *Plant Biotechnol. J.* 17, 1985–1997. doi: 10.1111/pbi.13114

Boulekbache-Makhlouf, L., Medouni, L., Medouni-Adrar, S., Arkoub, L., and Madani, K. (2013). Effect of solvents extraction on phenolic content and antioxidant activity of the byproduct of eggplant. *Ind. Crop Prod.* 49, 668–674. doi: 10.1016/j.indcrop.2013.06.009

Carocho, M., Barreiro, M. F., Morales, P., and Ferreira, I. C. F. R. (2014). Adding molecules to food, pros and cons: a review on synthetic and natural food additives. *Compr. Rev. Food Sci. Food Saf.* 13, 377–399. doi: 10.1111/1541-4337.12065

Chen, C., Chen, H., Zhang, Y., Thomas, H. R., Frank, M. H., He, Y., et al. (2020). TBtools: an integrative toolkit developed for interactive analyses of big biological data. *Mol. Plant* 13, 1194–1202. doi: 10.1016/j.molp.2020.06.009

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, S., Zhou, Y., Chen, Y., and Gu, J. (2018). fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics* 34, i884–i890. doi: 10.1093/bioinformatics/bty560

Cheng, J., Wei, G., Zhou, H., Gu, C., Vimolmangkang, S., Liao, L., et al. (2014). Unraveling the mechanism underlying the glycosylation and methylation of anthocyanins in peach. *Plant Physiol.* 166, 1044–1058. doi: 10.1104/pp.114.246876

CNCB-NGDC Members and Partners (2021). Database resources of the National Genomics Data Center. China National Center for Bioinformation in 2021. *Nucleic Acids Res.* 8, D18–D28. doi: 10.1093/nar/gkaa1022

de Mejia, E. G., Dia, V. P., West, L., West, M., Singh, V., Wang, Z., et al. (2015). Temperature dependency of shelf and thermal stabilities of anthocyanins from corn distillers' dried grains with solubles in different ethanol extracts and a commercially available beverage. *J. Agric. Food Chem.* 63, 10032–10041. doi: 10.1021/acs.jafc.5b03888

De Rosso, V. V., and Mercadante, A. Z. (2007). The high ascorbic acid content is the main cause of the low stability of anthocyanin extracts from acerola. *Food Chem.* 103, 935–943. doi: 10.1016/j.foodchem.2006.09.047

Du, H., Lai, L., Wang, F., Sun, W., Zhang, L., Li, X., et al. (2018). Characterisation of flower colouration in 30 *Rhododendron* species via anthocyanin and flavonol identification and quantitative traits. *Plant Biol.* 20, 121–129. doi: 10.1111/plb.12649

Ersus Bilek, S., Yılmaz, F. M., and Özkan, G. (2017). The effects of industrial production on black carrot concentrate quality and encapsulation of anthocyanins in whey protein hydrogels. *Food Bioprod. Process.* 102, 72–80. doi: 10.1016/j.fbp.2016.12.001

Fan, R., Sun, Q., Zeng, J., and Zhang, X. (2020). Contribution of anthocyanin pathways to fruit flesh coloration in pitayas. *BMC Plant Biol.* 20:361. doi: 10.1186/s12870-020-02566-2

Fang, H., Dong, Y., Yue, X., Chen, X., He, N., Hu, J., et al. (2019a). *MdCOL4* interaction mediates crosstalk between UV-B and high temperature to control fruit coloration in apple. *Plant Cell Physiol.* 60, 1055–1066. doi: 10.1093/pcp/pcz023

Fang, H., Dong, Y., Yue, X., Hu, J., Jiang, S., Xu, H., et al. (2019b). The B-box zinc finger protein MdBBX20 integrates anthocyanin accumulation in response to ultraviolet radiation and low temperature. *Plant Cell Environ.* 42, 2090–2104. doi: 10.1111/pce.13552

Feng, C., Ding, D., Feng, C., and Kang, M. (2020). The identification of an R2R3-MYB transcription factor involved in regulating anthocyanin biosynthesis in *Primulina swinglei* flowers. *Gene* 752:144788. doi: 10.1016/j.gene.2020.144788

Gu, Z., Zhu, J., Hao, Q., Yuan, Y., Duan, Y., Men, S., et al. (2019). A novel R2R3-MYB transcription factor contributes to petal blotch formation by regulating organ-specific expression of *PsCHS* in tree peony (*Paeonia suffruticosa*). *Plant Cell Physiol.* 60, 599–611. doi: 10.1093/pcp/pcy232

Hichri, I., Barrieu, F., Bogs, J., Kappel, C., Delrot, S., and Lauvergeat, V. (2011). Recent advances in the transcriptional regulation of the flavonoid biosynthetic pathway. *J. Exp. Bot.* 62, 2465–2483. doi: 10.1093/jxb/erq442

Holton, T. A., and Cornish, E. C. (1995). Genetics and biochemistry of anthocyanin biosynthesis. *Plant Cell* 7, 1071–1083. doi: 10.1105/tpc.7.7.1071

Horiuchi, R., Nishizaki, Y., Okawa, N., Ogino, A., and Sasaki, N. (2020). Identification of the biosynthetic pathway for anthocyanin triglucoside, the precursor of polyacylated anthocyanin, in red cabbage. *J. Agric. Food Chem.* 68, 9750–9758. doi: 10.1021/acs.jafc.0c03480

Huang, G., Zeng, Y., Wei, L., Yao, Y., Dai, J., Liu, G., et al. (2020). Comparative transcriptome analysis of mulberry reveals anthocyanin biosynthesis mechanisms in black (*Morus atropurpurea* Roxb.) and white (*Morus alba* L.) fruit genotypes. *BMC Plant Biol.* 20:279. doi: 10.1186/s12870-020-02486-1

Jeong, S., Das, P. K., Jeoung, S. C., Song, J., Lee, H. K., Kim, Y., et al. (2010). Ethylene suppression of sugar-induced anthocyanin pigmentation in *Arabidopsis*. *Plant Physiol.* 154, 1514–1531. doi: 10.1104/pp.110.161869

Jian, W., Cao, H., Yuan, S., Liu, Y., Lu, J., Lu, W., et al. (2019). *SlMYB75*, an MYB-type transcription factor, promotes anthocyanin accumulation and enhances volatile aroma production in tomato fruits. *Hortic. Res.* 6:22. doi: 10.1038/s41438-018-0098-y

Jiang, T., Zhang, M., Wen, C., Xie, X., Tian, W., Wen, S., et al. (2020). Integrated metabolomic and transcriptomic analysis of the anthocyanin regulatory networks in *Salvia miltiorrhiza* Bge. flowers. *BMC Plant Biol.* 20:349. doi: 10.1186/s12870-020-02553-7

Kelebek, H., and Selli, S. (2011). Characterization of phenolic compounds in strawberry fruits by RP-HPLC-DAD and investigation of their antioxidant capacity. *J. Liq. Chromatogr. Relat. Technol.* 34, 2495–2504. doi: 10.1080/10826076.2011.591029

Khoo, H. E., Azlan, A., Tang, S. T., and Lim, S. M. (2017). Anthocyanidins and anthocyanins: colored pigments as food, pharmaceutical ingredients, and the potential health benefits. *Food Nutr. Res.* 61:1361779. doi: 10.1080/16546628.2017.1361779

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

Kitts, D. D., and Tomiuk, S. (2013). Studies on mitigating lipid oxidation reactions in a value-added dairy product using a standardized cranberry extract. *Agriculture* 3, 236–252. doi: 10.3390/agriculture3020236

Lang, X., Li, N., Li, L., and Zhang, S. (2019). Integrated metabolome and transcriptome analysis uncovers the role of anthocyanin metabolism in *Michelia maudiae*. *Int. J. Genomics* 2019:4393905. doi: 10.1155/2019/4393905

Langmead, B., and Salzberg, S. L. (2012). Fast gapped-read alignment with Bowtie 2. *Nat. Methods* 9, 357–359. doi: 10.1038/nmeth.1923

Li, H., Liu, J., Pei, T., Bai, Z., Han, R., and Liang, Z. (2019). Overexpression of *SmANS* enhances anthocyanin accumulation and alters phenolic acids content in *Salvia miltiorrhiza* and *Salvia miltiorrhiza* Bge f.alba Plantlets. *Int. J. Mol. Sci.* 20:2225. doi: 10.3390/ijms20092225

Li, P., Li, Y. J., Zhang, F. J., Zhang, G. Z., Jiang, X. Y., Yu, H. M., et al. (2017). The *Arabidopsis* UDP-glycosyltransferases *UGT79B2* and *UGT79B3*, contribute to cold, salt and drought stress tolerance via modulating anthocyanin accumulation. *Plant J.* 89, 85–103. doi: 10.1111/tpj.13324

Li, Y., Im Kim, J., Pysh, L., and Chapple, C. (2015). Four isoforms of *Arabidopsis* 4-Coumarate: CoA Ligase have overlapping yet distinct roles in phenylpropanoid metabolism. *Plant Physiol.* 169, 2409–2421. doi: 10.1104/pp.15.00838

Lu, W., Chen, J., Ren, X., Yuan, J., Han, X., Mao, L., et al. (2018). One novel strawberry MADS-box transcription factor *FaMADS1a* acts as a negative regulator in fruit ripening. *Sci. Hortic.* 227, 124–131. doi: 10.1016/j.scienta.2017.09.042

Mehrtens, F., Kranz, H., Bednarek, P., and Weisshaar, B. (2005). The *Arabidopsis* transcription factor *MYB12* is a flavonol-specific regulator of phenylpropanoid biosynthesis. *Plant Physiol.* 138, 1083–1096. doi: 10.1104/pp.104.058032

Mourtzinos, I., Prodromidis, P., Grigorakis, S., Makris, D. P., Biliaderis, C. G., and Moschakis, T. (2018). Natural food colourants derived from onion wastes: application in a yoghurt product. *Electrophoresis* 39, 1975–1983. doi: 10.1002/elps.201800073

Nakatsuka, T., Nishihara, M., Mishiba, K., and Yamamura, S. (2005). Two different mutations are involved in the formation of white-flowered gentian plants. *Plant Sci.* 169, 949–958. doi: 10.1016/j.plantsci.2005.06.013

Ongkowijoyo, P., Luna-Vital, D. A., and Gonzalez de Mejia, E. (2018). Extraction techniques and analysis of anthocyanins from food sources by mass spectrometry: an update. *Food Chem.* 250, 113–126. doi: 10.1016/j.foodchem.2018.01.055

Pan, L., Li, J., Yin, H., Fan, Z., and Li, X. (2020). Integrated physiological and transcriptomic analyses reveal a regulatory network of anthocyanin metabolism contributing to the ornamental value in a novel hybrid cultivar of *Camellia japonica*. *Plants* 9:1724. doi: 10.3390/plants9121724

Pertea, M., Kim, D., Pertea, G. M., Leek, J. T., and Salzberg, S. L. (2016). Transcript-level expression analysis of RNA-seq experiments with HISAT, StringTie and Ballgown. *Nat. Protoc.* 11, 1650–1667. doi: 10.1038/nprot.2016.095

Pertea, M., Pertea, G. M., Antonescu, C. M., Chang, T. C., Mendell, J. T., and Salzberg, S. L. (2015). StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. *Nat. Biotechnol.* 33, 290–295. doi: 10.1038/nbt.3122

Pineda-Vadillo, C., Nau, F., Guerin-Dubiard, C., Jardin, J., Lechevalier, V., Sanz-Buenhombre, M., et al. (2017). The food matrix affects the anthocyanin profile of fortified egg and dairy matrices during processing and in vitro digestion. *Food Chem.* 214, 486–496. doi: 10.1016/j.foodchem.2016.07.049

Schaefer, H. M., Schaefer, V., and Levey, D. J. (2004). How plant-animal interactions signal new insights in communication. *Trends Ecol. Evol.* 19, 577–584. doi: 10.1016/j.tree.2004.08.003

Schmittgen, T. D., and Livak, K. J. (2008). Analyzing real-time PCR data by the comparative C(T) method. *Nat. Protoc.* 3, 1101–1108. doi: 10.1038/nprot.2008.73

Shahid, M., Shahid ul, I., and Mohammad, F. (2013). Recent advancements in natural dye applications: a review. *J. Clean. Prod.* 53, 310–331. doi: 10.1016/j.jclepro.2013.03.031

Shi, H., Liu, G., Wei, Y., and Chan, Z. (2018). The zinc-finger transcription factor *ZAT6* is essential for hydrogen peroxide induction of anthocyanin synthesis in *Arabidopsis*. *Plant Mol. Biol.* 97, 165–176. doi: 10.1007/s11103-018-0730-0

Smeriglio, A., Barreca, D., Bellocco, E., and Trombetta, D. (2016). Chemistry, pharmacology and health benefits of anthocyanins. *Phytother. Res.* 30, 1265–1286. doi: 10.1002/ptr.5642

Sun, Y., Fan, Z., Li, X., Li, J., and Yin, H. (2014). The *APETALA1* and *FRUITFUL* homologs in *Camellia japonica* and their roles in double flower domestication. *Mol. Breed.* 33, 821–834. doi: 10.1007/s11032-013-9995-9

Tang, B., Li, L., Hu, Z., Chen, Y., Tan, T., Jia, Y., et al. (2020). Anthocyanin accumulation and transcriptional regulation of anthocyanin biosynthesis in purple pepper. *J. Agric. Food Chem.* 68, 12152–12163. doi: 10.1021/acs.jafc.0c02460

Todaro, A., Cimino, F., Rapisarda, P., Catalano, A. E., Barbagallo, R. N., and Spagna, G. (2009). Recovery of anthocyanins from eggplant peel. *Food Chem.* 114, 434–439. doi: 10.1016/j.foodchem.2008.09.102

Wang, Y., Liu, W., Wang, X., Yang, R., Wu, Z., Wang, H., et al. (2020). *MiR156* regulates anthocyanin biosynthesis through targets and other microRNAs in poplar. *Hortic. Res.* 7:118. doi: 10.1038/s41438-020-00341-w

Wang, Y., Song, F., Zhu, J., Zhang, S., Yang, Y., Chen, T., et al. (2017). GSA: genome sequence archive. *Genom. Proteom. Bioinf.* 15, 14–18. doi: 10.1016/j.gpb.2017.01.001

Wei, K., Wang, L., Zhang, Y., Ruan, L., Li, H., Wu, L., et al. (2019). A coupled role for *CsMYB75* and *CsGSTF1* in anthocyanin hyperaccumulation in purple tea. *Plant J.* 97, 825–840. doi: 10.1111/tpj.14161

Wishart, D. S., Jewison, T., Guo, A. C., Wilson, M., Knox, C., Liu, Y., et al. (2013). HMDB 3.0–The human metabolome database in 2013. *Nucleic Acids Res.* 41, D801–D807. doi: 10.1093/nar/gks1065

Xia, E. H., Li, F. D., Tong, W., Li, P. H., Wu, Q., Zhao, H. J., et al. (2019). Tea plant information archive: a comprehensive genomics and bioinformatics platform for tea plant. *Plant Biotechnol. J.* 17, 1938–1953. doi: 10.1111/pbi.13111

Xue, L., Wang, J., Zhao, J., Zheng, Y., Wang, H., Wu, X., et al. (2019). Study on cyanidin metabolism in petals of pink-flowered strawberry based on transcriptome sequencing and metabolite analysis. *BMC Plant Biol.* 19:423. doi: 10.1186/s12870-019-2048-8

Yang, B., He, S., Liu, Y., Liu, B., Ju, Y., Kang, D., et al. (2020). Transcriptomics integrated with metabolomics reveals the effect of regulated deficit irrigation on anthocyanin biosynthesis in *Cabernet Sauvignon* grape berries. *Food Chem.* 314:126170. doi: 10.1016/j.foodchem.2020.126170

Zhang, Q., Wang, L., Liu, Z., Zhao, Z., Zhao, J., Wang, Z., et al. (2020). Transcriptome and metabolome profiling unveil the mechanisms of *Ziziphus jujuba* Mill. peel coloration. *Food Chem.* 312:125903. doi: 10.1016/j.foodchem.2019.125903

Zhang, Y., Butelli, E., and Martin, C. (2014a). Engineering anthocyanin biosynthesis in plants. *Curr. Opin. Plant Biol.* 19, 81–90. doi: 10.1016/j.pbi.2014.05.011

Zhang, Y., Chen, G., Dong, T., Pan, Y., Zhao, Z., Tian, S., et al. (2014b). Anthocyanin accumulation and transcriptional regulation of anthocyanin biosynthesis in purple bok choy (*Brassica rapa* var. chinensis). *J. Agric. Food Chem.* 62, 12366–12376. doi: 10.1021/jf503453e

Zhang, Y., Cheng, Y., Xu, S., Ma, H., Han, J., and Zhang, Y. (2020). Tree peony variegated flowers show a small insertion in the *F3'H* gene of the acyanic flower parts. *BMC Plant Biol.* 20:211. doi: 10.1186/s12870-020-02428-x

Zhang, Y., Cheng, Y., Ya, H., Xu, S., and Han, J. (2015). Transcriptome sequencing of purple petal spot region in tree peony reveals differentially expressed anthocyanin structural genes. *Front Plant Sci.* 6:964. doi: 10.3389/fpls.2015.00964

Zhang, Y., Hu, Z., Chu, G., Huang, C., Tian, S., Zhao, Z., et al. (2014c). Anthocyanin accumulation and molecular analysis of anthocyanin biosynthesis-associated genes in eggplant (*Solanum melongena* L.). *J. Agric. Food Chem.* 62, 2906–2912. doi: 10.1021/jf404574c

Zhang, Y., Zhou, T., Dai, Z., Dai, X., Li, W., Cao, M., et al. (2019). Comparative transcriptomics provides insight into floral color polymorphism in a orchid population. *Int. J. Mol. Sci.* 21:247. doi: 10.3390/ijms21010247

Zhou, C., Mei, X., Rothenberg, D. O. N., Yang, Z., Zhang, W., Wan, S., et al. (2020). Metabolome and transcriptome analysis reveals putative genes involved in anthocyanin accumulation and coloration in white and pink tea (*Camellia sinensis*) flower. *Molecules* 25:190. doi: 10.3390/molecules25010190

Zhou, H., Kui, L. W., Wang, H., Gu, C., Dare, A. P., Espley, R. V., et al. (2015). Molecular genetics of blood-fleshed peach reveals activation of anthocyanin biosynthesis by *NAC* transcription factors. *Plant J.* 82, 105–121. doi: 10.1111/tpj.12792

Zhou, X. W., Li, J. Y., Zhu, Y. L., Ni, S., Chen, J. L., Feng, X. J., et al. (2017). De novo assembly of the *Camellia nitidissima* transcriptome reveals key genes of flower pigment biosynthesis. *Front. Plant Sci.* 8:1545. doi: 10.3389/fpls.2017.01545

Zhu, Z. J., Schultz, A. W., Wang, J., Johnson, C. H., Yannone, S. M., Patti, G. J., et al. (2013). Liquid chromatography quadrupole time-of-flight mass spectrometry characterization of metabolites guided by the METLIN database. *Nat. Protoc.* 8, 451–460. doi: 10.1038/nprot.2013.004

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.

frontiers in Plant Science

ORIGINAL RESEARCH
published: 14 July 2021
doi: 10.3389/fpls.2021.700958

Terpene Synthase-b and *Terpene Synthase-e/f* Genes Produce Monoterpenes for *Phalaenopsis bellina* Floral Scent

Hsin Huang[1†], Yi-Wei Kuo[1†], Yu-Chen Chuang[1†], Ya-Ping Yang[1], Li-Min Huang[1], Mei-Fen Jeng[2], Wen-Huei Chen[2] and Hong-Hwa Chen[1,2,3]*

[1] Department of Life Sciences, National Cheng Kung University, Tainan, Taiwan, [2] Orchid Research and Development Center, National Cheng Kung University, Tainan, Taiwan, [3] Institute of Tropical Plant and Microbial Sciences, National Cheng Kung University, Tainan, Taiwan

Orchids are the most species-rich plants and highly interactive with pollinators via visual or olfactory cues. Biosynthesis and emission of volatile organic compounds (VOCs) to the atmosphere facilitate the olfactory cues and ensure successful pollination. *Phalaenopsis bellina* is a scented orchid with monoterpenes as major VOCs, comprising linalool, geraniol, and their derivatives. Comparative transcriptomics analysis identified four *terpene synthase-b* (*TPS-b*) genes and two *TPS-e/f* genes with differential gene expression between scented and scentless *Phalaenopsis* species. Here, we confirmed their differential expression between scented and scentless *Phalaenopsis* orchids and excluded one *TPS-b* candidate. We analyzed the temporal and spatial expression and functionally characterized these *TPSs*. Both *TPS-b* and *TPS-e/f* genes showed an increased expression on blooming day or 3 days post-anthesis (D + 3) before the optimal emission of floral scent on D + 5, with especially high expression of *PbTPS5* and *PbTPS10*. The *TPS-b* genes are expressed exclusively in reproductive organs, whereas the *TPS-e/f* genes are expressed in both reproductive and vegetative organs. *In planta* functional characterization of both *PbTPS5* and *PbTPS10* in tobacco and scentless *Phalaenopsis* plants did not produce terpenoids. Further ectopic expression in scented *Phalaenopsis* cultivar *P.* I-Hsin Venus showed that linalool was the main product, with *PbTPS10* displaying 3-fold higher activity than *PbTPS5.* On *in vitro* enzyme assay with purified recombinant TPS-b proteins ectopically expressed in *Escherichia coli*, geraniol was the product catalyzed by *PbTPS5* and *PbTPS9*. *PbTPS3* was a linalool/(β)-cis-ocimene synthase and *PbTPS4* a linalool synthase. In conclusion, both *TPS-b* and *TPS-e/f* enzymes orchestrated floral monoterpene biosynthesis in *P. bellina*.

Keywords: *Phalaenopsis*, monoterpene, floral scent, terpene synthase, geraniol, linalool, orchid, recombinant protein

OPEN ACCESS

Edited by:
Xiumin Fu,
South China Botanical Garden, Chinese Academy of Sciences, China

Reviewed by:
Zhenming Yu,
South China Botanical Garden, Chinese Academy of Sciences, China
Ming Sun,
Beijing Forestry University, China

****Correspondence:***
Hong-Hwa Chen
hhchen@mail.ncku.edu.tw

† 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: *27 April 2021*
Accepted: *21 June 2021*
Published: *14 July 2021*

Citation:
Huang H, Kuo Y-W, Chuang Y-C, Yang Y-P, Huang L-M, Jeng M-F, Chen W-H and Chen H-H (2021) Terpene Synthase-b and Terpene Synthase-e/f Genes Produce Monoterpenes for Phalaenopsis bellina Floral Scent. Front. Plant Sci. 12:700958. doi: 10.3389/fpls.2021.700958

INTRODUCTION

Orchidaceae is one of the most species-rich families in angiosperms and highly interactive with their pollinators. Interaction between pollinators and orchids has played a major role in orchid evolution (Schiestl et al., 2000; Sanguinetti et al., 2012). In Orchidaceae, both visual and olfactory captures are crucial for successful pollination. For olfactory capture, orchids emit floral volatile organic

compounds (VOCs), including terpenoids, phenylpropanoids, and benzoids, into the atmosphere. Floral VOCs have been characterized in several orchids, including monoterpenes of linalool and geraniol in *Phalaenopsis bellina* (Hsiao et al., 2006). Emission of the floral scent is developmentally regulated: It is undetectable on blooming day or with flower buds, starts to be detected on 3 days post-anthesis (D + 3), and is optimized on D + 5. It decreases as the flower becomes senescent (Hsiao et al., 2006).

Terpenes represent the largest group of plant secondary metabolites. Terpenoids have many diverse structures, with various carbon skeletons and a large assortment of functional groups. The biosynthesis of a terpene starts from the formation of the basic C5 units, isopentenyl diphosphate (IPP), and its allylic isomer, dimethylallyl diphosphate (DMADP). IPP is synthesized via two independent pathways: the mevalonate (MVA) pathway in the cytosol and methylerythritol phosphate (MEP) pathway in the plastids. The action of various short-chain prenyltransferases produces direct precursors for terpene synthases (TPSs): monoterpene synthase (MTPS) uses geranyl diphosphate (GDP) as a substrate for producing monoterpenes (C10), sesquiterpene synthase (STPS) uses farnesyl diphosphate (FDP) as a substrate for the biosynthesis of sesquiterpenes (C15), and diterpene synthase (DTPS) uses geranylgeranyl diphosphate (GGDP) as a substrate for producing diterpenes (C20) (Spitzer-Rimon et al., 2010; Colquhoun et al., 2011). Both MTPS and DTPS in the MEP pathway are present in the plastids for monoterpene and diterpene biosynthesis, whereas STPS in the MVA pathway is present in the cytoplasm for sesquiterpene biosynthesis (Lichtenthaler et al., 1997).

TPS genes are a large family classified into seven major subfamilies, namely, *TPS-a* to *TPS-h* (Chen et al., 2011). The *TPS-c* subfamily is the original subfamily and contains copalyl diphosphate synthases (CPSs), CPS/kaurene synthases (CPS/KSs), and DTPSs. *TPS-a* and *TPS-b* subfamilies are the angiosperm-specific subfamilies, and they contain STPSs and MTPSs, respectively (Bohlmann et al., 1998; Chen et al., 2011). The *TPS-d* subfamily is for gymnosperms MTPS and DTPS. The *TPS-g* subfamily lacks the RRX_8W motif for producing acyclic monoterpene and sesquiterpene. The *TPS-h* subfamily is identified as DTPS in *Selaginella moellendorffii* (Chen et al., 2011). The *TPS-e/f* subfamily contains KSs, MTPSs, STPSs, and DTPSs. Diverse functions have been reported for *TPS-e/f* enzymes, including linalool synthase in *Clarkia breweri* (Dudareva et al., 1996), geranyllinalool synthase in *Arabidopsis thaliana*, and farnesene synthase in kiwifruit (Nieuwenhuizen et al., 2009). The *TPS-f* subfamily is probably dicot-specific, whereas the *TPS-e* subfamily is not (Chen et al., 2011).

Phalaenopsis bellina is a native *Phalaenopsis* species in Borneo, Sarawak, and a common breeding parent for scented cultivars with a monoterpene phenotype. It has linalool, geraniol, and their derivatives as the major floral VOCs (Hsiao et al., 2006). The combination of bioinformatics and genomics identified PbGDPS as the key enzyme in orchid floral scent biosynthesis based on its high enrichment in the scented *P. bellina* but not scentless *Phalaenopsis equestris* and *Phalaenopsis aphrodite* subsp. *formosana* (hereafter *Phalaenopsis aphrodite*) (Hsiao et al., 2008; Chuang et al., 2018a). PbGDPS exhibits a dual prenyltransferase activity, producing GDP and FDP as the products, but lacks the $DD(X)_{2-4}D$ domain (Hsiao et al., 2008). The flower-specific expression profile of *PbGDPS* is concomitant with the emission pattern of monoterpenes on D + 5 (Hsiao et al., 2006). *PbGDPS* is the first enzyme with substantially differential expression between *P. bellina* and *Phalaenopsis aphrodite*, followed by several *TPS* genes in the monoterpene biosynthesis pathway (Chuang et al., 2018b). Comparative transcriptomics revealed the upregulated expression of *TPS-e/f* as well as *TPS-b* genes in *P. bellina* as compared with the scentless *Phalaenopsis aphrodite* (**Supplementary Figure 1B**).

Here, we confirm the differential gene expression of these *TPS* genes between *P. bellina* and *Phalaenopsis aphrodite*, illustrate the temporal and spatial gene expression, and provide a functional characterization of the TPSs in *TPS-b* and *TPS-e/f* subfamilies to further understand floral monoterpene biosynthesis in *Phalaenopsis* orchids. Functional characterization was revealed by both *in planta* and *in vitro* enzyme activity assays with transient overexpression and purified recombinant proteins ectopically overexpressed in *Escherichia coli*, respectively. These results suggest that both *TPS-b* and *TPS-e/f* enzymes contribute to monoterpene biosynthesis in the *P. bellina* floral scent.

MATERIALS AND METHODS

Plant Materials

Phalaenopsis bellina was from the Ming-Hui Orchids Nursery (Changhua, Taiwan) and Quan-Ya Orchid Nursery (Tainan, Taiwan) (**Supplementary Figure 2A**), *P.* I-Hsin Venus (hereafter Venus) was from I-Hsin Biotechnology Corp. (Chiayi, Taiwan) (**Supplementary Figure 2B**), and *Phalaenopsis aphrodite* and *P.* Sogo Yukidian "V3" (hereafter V3) were from Taiwan Sugar Corp. (Tainan, Taiwan). All *Phalaenopsis* plants were grown in the greenhouse at the National Cheng Kung University under natural light and controlled temperature from 27 to 30°C with 80% humidity. Venus, a commercial scented hybrid (**Supplementary Figure 2B**), was selected because it has mild levels of floral linalool emission. The genetic background of Venus is as follows: *Phalaenopsis amboinensis* (25%), *Phalaenopsis equestris* (25%), *Phalaenopsis venosa* (18.7%), *Phalaenopsis violacea* (15.6%), *Phalaenopsis amabilis* (10.4%), *Phalaenopsis lueddemanniana* (3.1%), and *Phalaenopsis aphrodite* (2.1%) via ORCHIDEYA.CA[1].

Nicotiana tabacum seeds were obtained from Dr. Ching-Chun Chang's laboratory (Department of Biotechnology and Bioindustry Sciences, NCKU) and germinated directly in soil and grown under a 16-h day/8-h night photoperiod with temperatures set at 27°C. Humidity was adjusted to 70%.

Phylogenetic Analysis

For TPS phylogenetic analysis, previously studied TPS sequences were downloaded from NCBI isolated from other plant species, and the full-length sequences were analyzed with the putative

[1]http://www.orchideya.ca/

TPSs isolated from *P. bellina* and *Phalaenopsis aphrodite* transcriptomes. Information on TPSs used for phylogenetic analysis is in **Supplementary Tables 1**, **2**. The alignment was performed by using ClustalW of the Molecular Evolutionary Genetics Analysis (MEGA7) program. The phylogenetic tree was constructed by using the neighbor-joining method with the Jones–Taylor–Thornton model and pairwise deletion with 1,000 bootstrap replicates.

Quantitative Real-Time PCR and Gene Expression Clustering

Total RNA was extracted from seven stages of *P. bellina* flowers, starting from D − 1 to D + 17, as described in the study by Chuang et al. (2018b). After DNA contamination was removed by DNase I (New England Biolabs, Beverly, MA, United States), reverse transcription of DNA to cDNA involved using SuperScript III (Invitrogen, Carlsbad, CA, United States). Primers were designed by using Primer Express 3.0 (Applied Biosystems, Foster City, CA, United States). Quantitative Reverse transcription-PCR (RT-PCR) (qRT-PCR) involved using a StepOne Real Time PCR System and SYBR Green kit (Applied Biosystems, Foster City, CA, United States) as described in the study by Chuang et al. (2018b). Several reference genes widely used in *Phalaenopsis* orchids were identified by a local BLASTN search (Chen et al., 2005; Hsiao et al., 2008; Hsieh et al., 2013; Hsu et al., 2015). *Actin1* had equivalent levels in the RNA-seq analysis of the *P. bellina* transcriptome and microarray data for *Phalaenopsis aphrodite* (Chuang et al., 2018b), so *PbActin1* in *P. bellina* was selected as an internal calibrator for further analysis. The expression of all genes was normalized by the reference gene *PbActin1*. Calculation of mean and error bars was based on three replicates. The temporal gene expression experiments for both *TPS-e/f* and *TPS-b* were performed in different years with *P. bellina* from different orchid nurseries. *P. bellina* from various orchid nurseries have different scent emission profiles: Some last longer, and others are shorter but with the same volatile compounds. The floral scent emission for *P. bellina* lasted longer for *TPS-b* gene expression analysis, so we analyzed *TPS-b* gene expression from bud to D + 17, whereas those used for *TPS-e/f* gene expression analysis has a shorter scent emission profile, so it was analyzed from 5 days before anthesis (D − 5) to D + 10.

Transient Expression in *Nicotiana tabacum*

The modified pCAMBIA1304 vector containing *PeMYB2*, which contributes to the red color formation in *Phalaenopsis* spp. flowers (Hsu et al., 2015), was used as a binary vector for the construction of transient expression in *N. tabacum* and *Phalaenopsis* orchids. *TPS* genes were constructed into the modified binary vector, and then the recombinant vectors containing inserted *TPS* genes (dubbed pCAMBIA1304-*TPS*) were transformed into *Agrobacterium tumefaciens* strain GV3101 by electroporation. The transformed *Agrobacteria* were spread on LB agar plates containing rifampicin (10 μg/ml), gentamycin (25 μg/ml), and kanamycin (50 μg/ml). The plates were incubated for 2 days at 28°C, and colonies were tested for insertion by colony PCR with ProTaq DNA Polymerase (Protech Technology Enterprise, Taiwan). A transformed single colony was used to inoculate a 5-ml culture of LB medium plus appropriate antibiotics, and cells were grown for 24 h at 28°C and 220 rpm. This pre-culture was used to inoculate a 100-ml culture and grown for about 4 h at 28°C to OD_{600} = 0.8–1.0. *Agrobacteria* were harvested by centrifugation at 5,000 × *g* and 4°C for 15 min, then resuspended in infiltration medium (Murashige and Skoog medium, 1 mM acetosyringone) to OD_{600} = 0.4, and incubated for 2–4 h at room temperature. Diluted *Agrobacteria* was injected into the abaxial side of the three youngest *N. tabacum* leaves that were > 1 cm by using a 1-ml syringe. Transformed plants were maintained in a growth chamber at 28°C under long-day conditions (16 h/8 h for light/dark), and gas chromatography–mass spectrometry (GC-MS) analysis was performed at 5–7 days post-infiltration.

Transient Overexpression in *Phalaenopsis* Orchids

The expression plasmid, pCAMBIA1304-*TPS* as described above, was transformed into *Agrobacterium tumefaciens* strain *EHA105* by electroporation. The growth of agrobacteria was as described above. The transformed *Agrobacteria* was injected into the adaxial side of sepals and petals at different floral stages (flower bud or just bloomed flower) by using a 0.5-ml insulin syringe. Transformed plants were maintained at 25°C, and GC-MS analysis was performed at 5 dpi. Because the constructs containing *PeMYB2* gene activated the anthocyanin pathway, red color formation in transformed flowers was considered a positive transformation in *Phalaenopsis* orchids.

Expression and Purification of Recombinant Proteins in *Escherichia coli*

For ectopic overexpression in *E. coli*, the open reading frames of PbTPSs were amplified and cloned into the expression vector pET-28b (Novagen, Madison, WI, United States), which has a His-tag on both the N- and C-termini of the inserted fragment. The construct was introduced into *E. coli* C41 (DE3) pLysS strain (Invitrogen, Carlsbad, CA, United States). Sequence analysis and protein structure modeling confirmed that no mutant occurred near the protein active site on DNA amplification. A single colony of the constructs was incubated in 5 ml LB medium with 50 μl/ml kanamycin overnight at 37°C. Then, the culture was transferred to 750 ml LB medium with 50 μl/ml kanamycin, incubated at 37°C to OD_{600} = 0.4–0.6, and induced with 0.4 mM isopropyl β-D-1-thiogalactopyranoside (IPTG) and grown under 16°C for 22 h. Cells were pelleted by centrifugation and suspended in extraction buffer (0.3 M NaCl, 50 mM HEPES, pH 7.4), then disrupted by sonication (SONICS, VC750, Newtown, CT, United States). Cell extracts were filtered through a 0.45-μm filter. Recombinant proteins were purified by using HiTrap TALON crude (GE Healthcare, Abingdon, United Kingdom) with different concentrations of imidazole solution (0.3 M NaCl, 50 mM HEPES, 5 mM/150 mM imidazole, pH 7.4). Protein expression was verified on Coomassie Blue-stained SDS–PAGE gel and Western blot analysis with 1/10,000 diluted

mouse anti-histidine monoclonal antibody (Roche Diagnostics, Penzberg, Germany).

Enzyme Activity Assay of Purified Recombinant Proteins

The affinity-purified recombinant proteins were concentrated and incubated with assay buffer (HEPES, pH7.4, 0.1 M KCl, 10 mM $MgCl_2$, 10 mM DTT, 10% glycerol) by using a 30 MWCO PES Vivaspin 50-ml concentrator (Sartorius Stedium Biotech, Goettingen, Germany). For functional characterization, recombinant protein was added to a single vial and diluted with assay buffer to a total volume of 500 μl, then incubated with GDP (2.5 μM) for 1 h at 30°C.

Volatile Collection and GC-MS Analysis

The floral volatiles of the *Phalaenopsis* orchids were collected by using a scent-extracting apparatus (Chuang et al., 2017). The scent emission pattern of *Phalaenopsis* orchids was measured from a single flower, with three biological replicates. The volatiles collected were eluted by hexane and identified by gas chromatography/high-resolution mass spectrometry (GC/HRMS) at the NCKU Instrument Center, as described in the study by Hsiao et al. (2006). As a negative control, metabolites originating from the scent-extracting apparatus were analyzed for the background. The volatiles were collected by using a dynamic headspace sampling system and solid-phase microextraction (Aharoni et al., 2003) in a glass container with 200 mg Tenax-TA resin (60/80 mesh; Supelco, Bellefont, PA, United States) packed into a plastic tube. The headspace sampling apparatus inserted a silicone tube to connect to the solid phase extraction columns and another silicone tube connected the column with air pumps at 350 mmHg collecting. The collected compounds were analyzed by GC-MS (QP2010, SHIMADXU, Shimadzu Co., Tokyo). An INNOWAX column (60 cm, 0.3 mm, 0.25 μm phase thickness) was used, and the oven was programmed from 40 to 230°C (held for 5 min) at 5°C/min increments. The pressure of the helium inlet was set at 75.2 kPa with 34.6 cm/s linear velocity (split flow 8.3 μl/min). The injector temperature was kept at 240°C with the injected volume set to 1 μl and the electron energy to 70 eV. Mass spectra and reconstructed chromatograms were obtained by automatic scanning of the samples in the mass range m/z 20–500 Da. Peaks on mass chromatograms with characteristic fragments were checked for homogeneity. The identities of all compounds were determined by comparing retention times and mass fragmentation patterns with the NIST98 database (US Environmental Protection Agency, 1998) and NIST05 database (SHIMADXU, Shimadzu Co., Tokyo). For quantitative analysis, 10 μg/ml ethyl myristate was used as an internal standard.

Protoplast Isolation

Protoplasts of floral buds were isolated from *Phalaenopsis aphrodite* as described in the study by Shen et al. (2017). Briefly, 10 g of 2-cm flower buds was sterilized in 70% alcohol for 1 min, followed by three washes in sterilized distilled water. The floral bud was cut into 1-mm^2 pieces and immersed in sterilized water for 1 h. Then, floral buds were transferred to 10 ml enzyme solution (1 ml of 0.2M MES, 0.4 g cellulose RS, 0.2 g macerozyme R-10, 6 ml of 1M sucrose, 100 μl of 2M KCl, 50 μl of 2M $CaCl_2$, 10% BSA) for incubation for 4 h on a shaker (30 rpm) in darkness at room temperature. After incubation, the enzyme–protoplast mixture was passed through steel screens of 100- and 50-μm pore sizes to remove undigested materials. The sample was centrifuged at 700 rpm for 5–10 min. The protoplasts on the surface were collected, resuspended in the wash buffer (cell and protoplast washing salt containing 0.6 M mannitol), and centrifuged at 700 rpm for 5 min; then the supernatant was discarded. After three washes, the pellet was resuspended in 0.5 ml of W5 solution (150 mM NaCl, 125 mM $CaCl_2$, 5 mM KCl, 2 mM MES, pH 5.7, 5 mM glucose). Then, the supernatant was spun, removed, and resuspended in 500 μl of MMg solution (4 mM MES, pH 5.7, 0.6M mannitol, 15 mM $MgCl_2$) before polyethylene glycol (PEG) 4000 treatment.

Transient Expression of Green Fluorescent Protein Fusion Proteins

For cellular localization experiments, green fluorescent protein (GFP) was constructed either upstream or downstream from the analyzed genes, such as GFP-PbTPS3 or PbTPS3-GFP, respectively. The Gateway Cloning System (Invitrogen, Carlsbad, CA, United States) was used for generating p2FGW7-fused protein transformation constructs consisting of the open reading frame of PbTPS3 or PbTPS4 cDNA under the control of the cauliflower mosaic virus (CaMV) 35S promoter. The plasmids were transformed into protoplasts of *Phalaenopsis aphrodite* by PEG 4000 treatment. An empty vector that expresses GFP under the control of the CaMV 35S promoter served as a negative control. GFP fluorescence was observed under a Carl Zeiss LSM780 confocal laser scanner microscope (CLSM-780, Zeiss, Germany) with GFP excited at 488 nm with a krypton/argon laser and chlorophyll autofluorescence excited at 514 nm with a HeNe-Laser. The emissions were collected through a 506- to 530-nm band pass filter (for GFP) and a 680- to 750-nm band pass filter (for chloroplasts). To strengthen the image visibility and resolution of the original images, all images were enhanced to 50% sharpness and 30% contrast by using the built-in chart format of PowerPoint.

RESULTS

Phylogenetic Analysis of Upregulated *TPSs* in *Phalaenopsis bellina* During Floral Development

Comparative transcriptomic analysis of scented *P. bellina* and scentless *Phalaenopsis aphrodite* revealed the upregulation of several *TPS* genes in *P. bellina* (Chuang et al., 2018b; **Supplementary Figure 1B**). Among them, *PbTPS5-1*, *PbTPS5-2*, *PbTPS9-1*, *PbTPS9-2-1*, *PbTPS9-2-2*, *PbTPS10-1*, *PbTPS10-2*, and *PbTPS7* are in the *TPS-b* subfamily, and *PbTPS4*, *PbTPS3-1*, and *PbTPS3-2* are in the *TPS-e/f* subfamily. Alternative splicing had occurred and produced different transcripts from the same gene, such as *TPS5-2*, which has an intron retention between exons 5 and 6 that caused a stop codon and early termination. A similar situation occurred for *TPS10-1*, in that an intron

retention between exons 4 and 5 caused a stop codon and early termination (**Supplementary Figure 3**). So, we focused on the major transcripts for phylogenetic analysis (**Supplementary Figure 1A**). These were six genes: *PbTPS3*, *PbTPS4*, *PbTPS5*, *PbTPS7*, *PbTPS9*, and *PbTPS10*. Further phylogenetic analysis with known *TPS* genes of other plants showed that *PbTPS5*, *PbTPS7*, *PbTPS9*, and *PbTPS10* in the *TPS-b* subfamily (red dots, **Figure 1**) and *PbTPS3* and *PbTPS4* in the *TPS-e/f* subfamily (blue dots, **Figure 1** and **Supplementary Figure 4**).

Differential Gene Expression of *PbTPSs* Between *Phalaenopsis bellina* and *Phalaenopsis aphrodite*

To further confirm the differential expression of these six *PbTPS* genes between scented and scentless *Phalaenopsis* orchids, qRT-PCR was performed. We examined the expression of four *TPS-b* genes (*PbTPS5*, *PbTPS7*, *PbTPS9*, and *PbTPS10*) and two *TPS-e/f* genes (*PbTPS3* and *PbTPS4*) at the flower bud stage and D + 5 (**Figure 2**). Previously, we found that the emission of monoterpenes in *P. bellina* peaked at the full-bloom stages (D + 4 and D + 5) but was absent in floral buds (Chuang et al., 2018b). All *PbTPS* genes in *P. bellina* showed higher expression at D + 5 than at the bud stage. Except for *PbTPS7*, the qRT-PCR results were consistent with transcriptomic analysis showing an enhanced expression in the scented orchids. Expression was higher for *PbTPS3*, *PbTPS4*, *PbTPS5*, *PbTPS9*, and *PbTPS10* in *P. bellina* than in *Phalaenopsis aphrodite* in floral buds and at D + 5 (**Figure 2**). These results suggested that *PbTPS3*, *PbTPS4*, *PbTPS5*, *PbTPS9*, and *PbTPS10* were highly associated with floral scent monoterpene biosynthesis; *PbTPS7* was removed from the following analysis.

Temporal and Spatial Gene Expression of the *PbTPS-b* Subfamily

We examined the temporal expression of three *TPS-b* genes from the flower bud stage to D + 17 of *P. bellina* by qRT-PCR (**Figure 3**). Both *PbTPS5* and *PbTPS10* showed much higher expression than *PbTPS9* in *P. bellina* during floral development (**Figure 3A**). *PbTPS5* and *PbTPS10* expressed highly on blooming day (Dd) and D + 3, respectively (**Figure 3A**). *PbTPS9* showed a biphasic pattern, with expression high on D + 5, decreased on D + 7, and then high again on D + 13 (**Figure 3A**).

The three *TPS-b* genes were exclusively expressed in reproductive organs, with differential profiles. *PbTPS5*, *PbTPS9*, and *PbTPS10* expressed in sepal, petal, and lip and to a lesser extent in column (**Figure 3B**). Especially, *PbTPS5* expressed highly in petal, whereas *PbTPS10* expressed highly in sepal.

Temporal and Spatial Expression Profiles of the *PbTPS-e/f* Subfamily

We analyzed the temporal expression of two *PbTPS-e/f* genes, *PbTPS3* and *PbTPS4*, by qPCR from 5 days before anthesis (D − 5) to D + 10. *PbTPS3* showed increasingly high expression from Dd to D + 5, then reduced expression on D + 7, whereas *PbTPS4* showed high expression on D + 3 and D + 5, then diminished expression (**Figure 3C**). The expression was much lower for *PbTPS4* than for *PbTPS3* in flowers (**Figure 3C**). *PbTPS3* expressed mainly in the leaf, and in the flower, it expressed higher in the lip than in the sepal, petal, or column (**Figure 3D**). *PbTPS4* also showed high expression in the lip and column and very low expression in the sepal and petal; it was expressed in the root and leaf but at a lower level than in the floral organs (**Figure 3D**).

Transient Overexpression of *PbTPS5* and *PbTPS10* in *Nicotiana tabacum*

The temporal expression profiles of *PbTPS5* and *PbTPS10* were concomitant with the emission of monoterpenes such as linalool and geraniol, so we studied their enzyme functions by ectopic overexpression in *N. tabacum*. Several compounds were produced when overexpressing *PbGDPS* to provide GDP as the substrate. A monoterpene myrcene derivative, tetrahydromyrcenol, was produced when transiently overexpressing *PbTPS10* or *PbGDPS* + *PbTPS10* in *N. tabacum* (**Supplementary Figure 5**). However, neither linalool nor geraniol was produced in *N. tabacum* overexpressing *PbTPS5* or *PbTPS10*. *N. tabacum* may have no other orchid-related genes required for monoterpene biosynthesis and hinder the production of linalool or geraniol.

Transient Overexpression of *PbTPS5* and *PbTPS10* in Scentless *Phalaenopsis*

We attempted the transient overexpression of *PbTPS5* and *PbTPS10* in the scentless species *Phalaenopsis aphrodite* and a scentless cultivar V3. Overexpression of *PbTPS5* and *PbTPS10* with or without *PbGDPS* in *Phalaenopsis aphrodite* did not induce the terpene production in the emitted volatiles or the internal non-volatiles (**Supplementary Figures 6A,C**), even though monoterpenes of α-terpineol and α-cedrene were detected in the non-volatile compounds in the overexpression of *PbGDPS* in *Phalaenopsis aphrodite* This observation suggests that other enzymes or factors might be needed to produce volatile monoterpenes in *Phalaenopsis aphrodite*. Further overexpression of *PbTPS5* and *PbTPS10* was performed in V3, and the production of tetrahydromyrcenol and sesquiterpenoids was enhanced, but neither linalool nor geraniol was detected in the volatiles and non-volatiles of V3 (**Supplementary Figures 6B,D**). Previously, we showed that the scentless *Phalaenopsis* orchids lack *PbbZIP4* and *PbbHLH4* transcription factors for floral volatile biosynthesis (Chuang et al., 2018a,b). Hence, some other factors (such as transcription factors) may be required and are missing in the scentless orchid flowers. Alternatively, the presence of repressors may hinder the monoterpene biosynthesis in the scentless *Phalaenopsis* orchids, even though the key enzymes are provided.

Transient Overexpression of *PbTPS5* and *PbTPS10* in a Scented *Phalaenopsis* Cultivar

We then overexpressed *PbTPS5* and *PbTPS10* in a scented cultivar, Venus, an offspring of *P. bellina* that has gone through

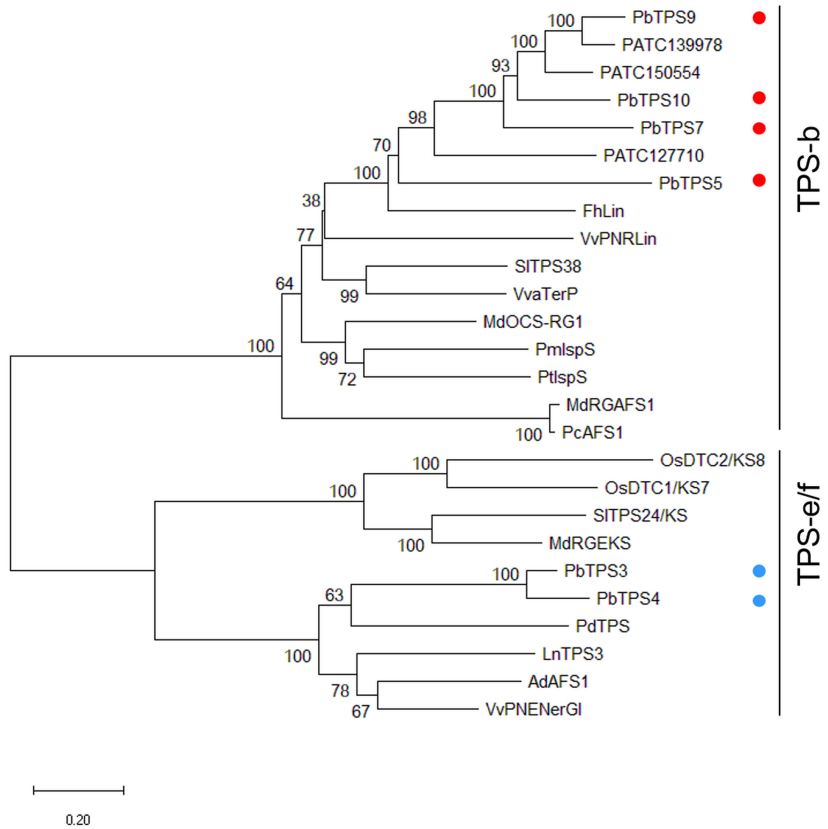


FIGURE 1 | Phylogenetic analysis of *PbTPS3*, *PbTPS4*, *PbTPS5*, *PbTPS7*, *PbTPS9*, *PbTPS10*, and other plant TPSs. The coding sequence was used to construct a phylogenetic tree using the nearest neighbor-joining method with the Jones–Taylor–Thornton model and pairwise deletion with 1,000 bootstrap replicates by MEGA 7.0. The numbers at each node represent the bootstrap values. The MTPS of *Phalaenopsis* orchid are marked with red dots. *PbTPS5*, *PbTPS7*, *PbTPS9*, and *PbTPS10* are grouped in the *TPS-b* subfamily and labeled in red dots. *PbTPS3* and *PbTPS4* are grouped in the *TPS-e/f* subfamily and labeled in blue dots. Information on *TPS* genes used for phylogenetic analysis is in **Supplementary Table 1**.

more than 10 generations of crossing. In contrast to *P. bellina* producing one or two flowers and with strong scent, Venus is floriferous (**Supplementary Figure 2B**) but with reduced floral scent. Before the overexpression, we analyzed the endogenous gene expression and floral scent emission of Venus. We first detected the scent compounds of Venus and examined the gene expression profiles of endogenous *PbGDPS*, *PbTPS5*, and *PbTPS10*. Venus produced VOCs of linalool, octanoic acid, and benzaldehyde in flowers, with linalool as the major floral scent, yet no geraniol was detected (**Figure 4B**). The floral scent emission followed a similar pattern as that of *P. bellina* (i.e., the strong scent was detected from D + 3). Concomitantly, *PbGDPS* and *PbTPS5* significantly expressed from Dd and D + 3, respectively (**Figure 4A**), yet with little or no expression of *PbTPS10* during floral development from flower bud to D + 13 (**Figure 4A**). This result implied that *PbTPS5* alone can account for linalool biosynthesis.

Cauliflower mosaic virus 35S promoter was used to drive the overexpression of both *PbTPS5* and *PbTPS10*. The emission of linalool was enhanced with transient overexpression of *PbTPS5* and *PbTPS10* in Venus floral tepals (peak 2, **Figure 4C**). Thus, both *PbTPS5* and *PbTPS10* caused a 7.7- and 29.0-fold increase, respectively, of linalool production as compared with GUS. In addition, emission of tetrahydromyrcenol was detected overexpression of *PbTPS5* and *PbTPS10*, respectively (peak 1, **Figure 4C**). However, geraniol was not produced on overexpressing *PbTPS5* or *PbTPS10* in Venus.

Functional Characterization With Purified *TPS-b* Recombinant Proteins Ectopically Overexpressed in *Escherichia coli*

To further confirm the enzyme activity, *in vitro* assay was performed with purified recombinant proteins ectopically expressed in *E. coli*. *PbTPS5* was cloned into PET-28b with 6x His at the N- and C-terminus and transformed into *E. coli*. C41 (DE3). The affinity-purified PbTPS5 protein was assayed in a buffer containing substrate GDP, and the products were analyzed by GC-MS. Geraniol was detected as the main product catalyzed by PbTPS5 (**Figure 5A**).

The enzyme activity of PbTPS9 was examined with the purified recombinant proteins as described above. The product of PbTPS9 was geraniol (**Figure 5B**).

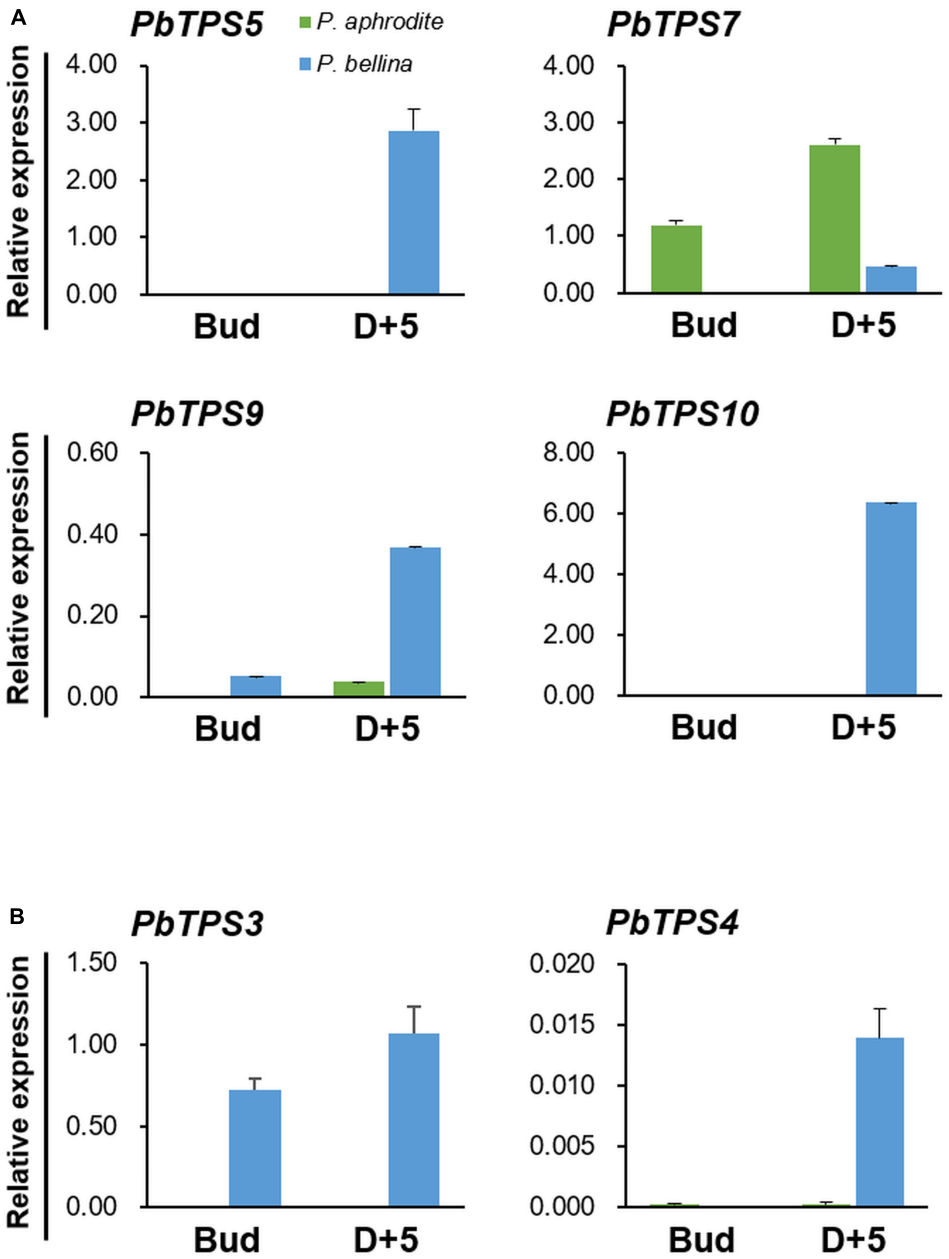


FIGURE 2 | Differential gene expression of *PbTPSs* between *P. bellina* and *Phalaenopsis aphrodite*. Panels **(A,B)** are the gene expression of *TPS-b* and *TPS-e/f* members, respectively. The expression of *PbTPSs* was analyzed at the floral bud stage (Bud) and 5 days after anthesis (D + 5). Green and blue column represent the expression of *PbTPSs* in *Phalaenopsis aphrodite* and *P. bellina*, respectively. Data were obtained from three independent experiments. The expression of all genes was normalized to the reference gene, *PbActin1* (Chuang et al., 2018b). Calculation of mean and standard error was based on triplicate repeats.

Functional Characterization of Purified *TPS-e/f* Recombinant Proteins Ectopically Overexpressed in *Escherichia coli*

To functionally characterize the two *TPS-e/f* genes PbTPS3 and PbTPS4, we analyzed the affinity-purified recombinant proteins by using three different prenyl diphosphates as substrates, namely, GDP (C10), FDP (C15), and GGDP (C20). The reaction products were analyzed by GC-MS. PbTPS3 was a dual enzyme as a linalool/(β)-cis-ocimene synthase (**Figure 6A**). The authentic standard of linalool was used for comparing retention time. PbTPS4 was a linalool synthase and used GDP as the substrate (**Figure 6B**). Neither FDP nor GGDP could be catalyzed by PbTPS3 and PbTPS4 to produce sesquiterpene or diterpene (**Supplementary Figure 7**). Overall, these data indicate that PbTPS3 and PbTPS4 are the MTPS that exclusively produced linalool and (β)-cis-ocimene, or linalool only, respectively.

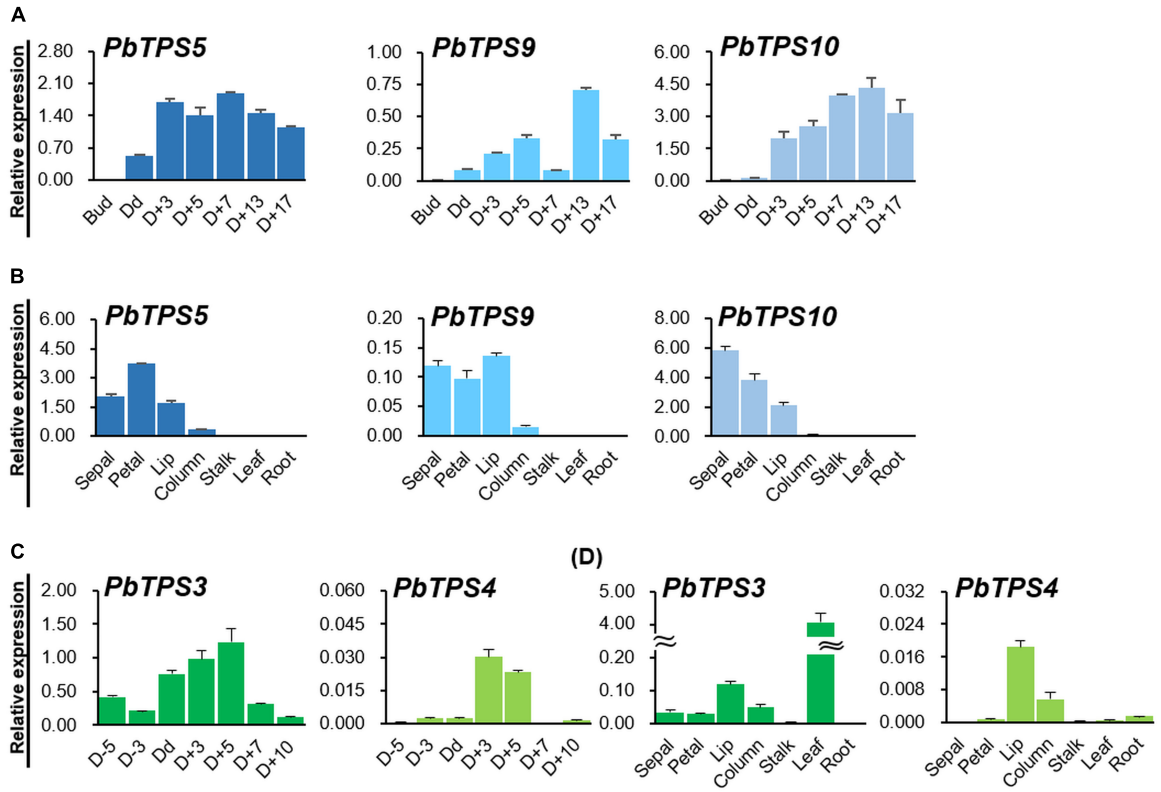


FIGURE 3 | Temporal and spatial gene expression of *TPS-b* and *TPS-e/f* members *in P. bellina*. **(A)** Temporal and **(B)** spatial expression profiles of *PbTPS5*, *PbTPS9*, and *PbTPS10*. **(C)** Temporal and **(D)** spatial expression of *PbTPS3* and *PbTPS4*. Temporal expression of *PbTPS5*, *PbTPS9*, and *PbTPS10* was analyzed at floral bud stage (Bud), day of anthesis (Dd), and 3–17 days after anthesis (D + 3 to D17). Temporal expression of *PbTPS3* and *PbTPS4* was analyzed from 3 to 5 days before anthesis (D-5 and D-3) and day of anthesis (Dd) to 10 days after anthesis (D + 10) in *P. bellina*. Spatial gene expression of *PbTPS* genes was determined in various tissues, including sepal, petal, lip, column, stalk, leaf, and root. Data were obtained from three independent experiments. The expression of all genes was normalized to the reference gene, *PbActin1* (Chuang et al., 2018b). Calculation of mean and standard error was based on triplicate repeats.

Subcellular Localization of PbTPS3 and PbTPS4

Monoterpene synthases are mainly present in the plastids. However, both PbTPS3 and PbTPS4 lack typical transit peptides for chloroplast localization (**Supplementary Figure 8**). We then performed the prediction of the subcellular localizations according to their amino acid sequences with various software packages. Unexpectedly, differential prediction results were obtained. For example, ChloroP 1.1 predicted that both PbTPS3 and PbTPS4 were not localized in the chloroplasts. TargetP suggested that they were localized in other sites and mitochondria, respectively. pSORT II indicated the possible localization of PbTPS3 in the cytosol (69.6%) and nucleus (26.1%) and possible localization of PbTPS4 in the nucleus (34.8%) and cytosol and mitochondria (21.7%). Wolf PSORT indicated that PbTPS3 and PbTPS4 possibly localized in the nucleus and chloroplast, respectively.

Therefore, subcellular localization of PbTPS3 and PbTPS4 was determined with GFP added at the N- or C-terminus of the full-length cDNAs (**Figure 7**). These constructs were transformed into protoplasts isolated from mature flower buds of *Phalaenopsis aphrodite*, and the GFP signal and autofluorescence emitted from chloroplasts were examined by confocal laser scanning microscopy. The GFP fluorescence of both PbTPS3 and PbTPS4 fusion proteins was exclusively overlaid with the autofluorescence signal of chloroplasts, which suggests that they are localized exclusively in the chloroplast (**Figure 7**). In contrast, the controls of no DNA added or empty vector showed no or very low GFP signal, respectively (**Figure 7**). These results confirmed that both PbTPS3 and PbTPS4 were localized in the plastids even without a typical transit peptide signal.

DISCUSSION

Both *TPS-b* and *TPS-e/f* Subfamilies Are Involved in the Biosynthesis of Monoterpenes in *Phalaenopsis bellina* Flower

In addition to the three *TPS* genes in the *TPS-b* subfamily with an enhanced expression in *P. bellina* as compared with *Phalaenopsis aphrodite*, two *TPS* genes in the *TPS-e/f* subfamily were also enriched. In *P. bellina*, differential gene expression of *PbTPS* genes is involved in floral monoterpene biosynthesis in various

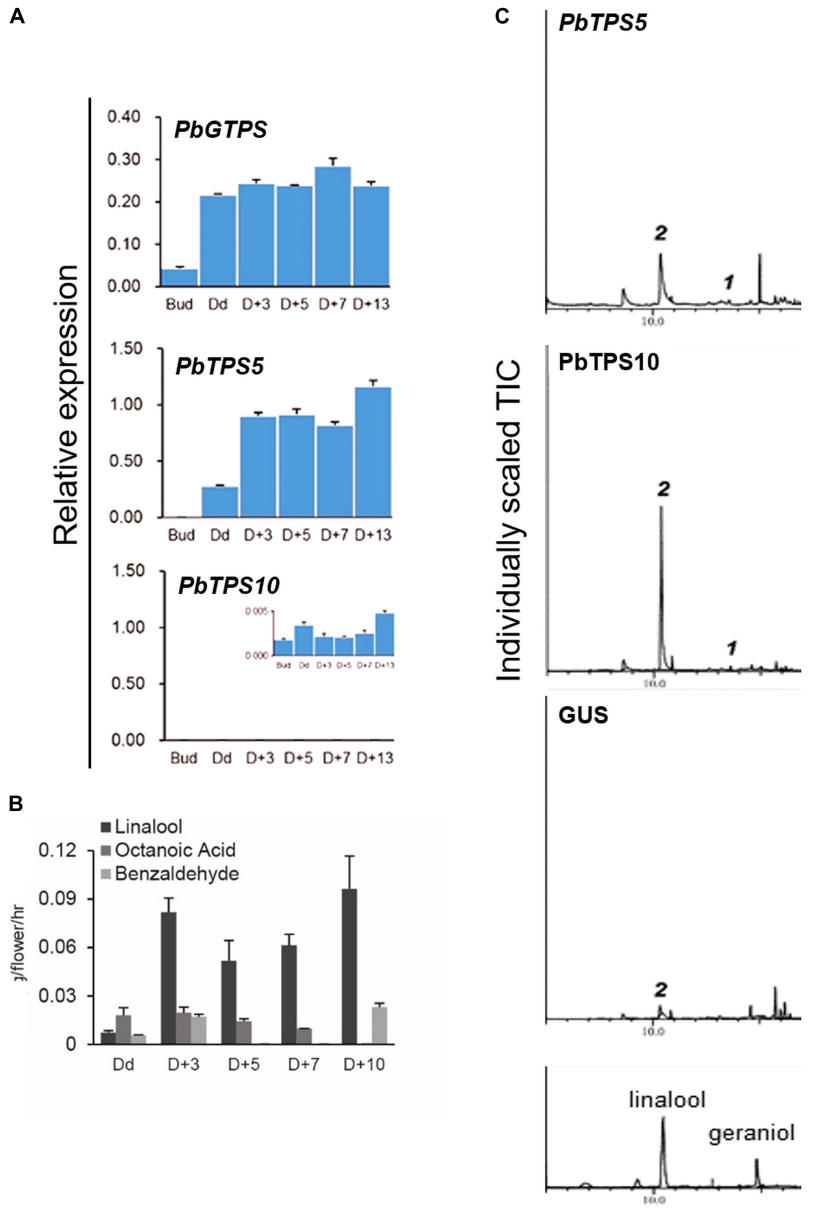


FIGURE 4 | (A) qRT-PCR analysis of gene expression of *PbGDPS*, *PbTPS5*, and *PbTPS10* in flowers. The inset for *PbTPS10* is the enlargement of the original scale. **(B)** Floral volatile organic compounds (VOCs) emitted from *P.* I-Hsin Venus. **(C)** Functional characterization of *PbTPS5* and *PbTPS10* transiently overexpressed in Venus. Plants overexpressing GUS were a negative control.

parts of a single flower. For example, *PbTPS5*, *PbTPS9*, and *PbTPS10* were involved in monoterpene biosynthesis in tepal (sepal and petal) and *PbTPS3*, *PbTPS4*, *PbTPS5*, *PbTPS9,* and *PbTPS10* in lip (**Figure 8A**). In addition, multiple *TPS* genes are involved in the same enzymatic function, and one *TPS* gene may catalyze more than one product of monoterpenes. For example, PbTPS3, PbTPS4, PbTPS5, and PbTPS10 produce linalool, and PbTPS5 and PbTPS9 produce geraniol. In addition, PbTPS3 produces (β)-cis-ocimene, with *PbTPS3* expressed in both flower and leaf. PbTPS3 has a dual enzyme activity of linalool synthase/(β)-cis-ocimene synthase (**Figure 8B**).

Both PbTPS3 and PbTPS4 contain a DDXXD and NSE/DTE motif for GDP binding, and they lack the RRX8W motif for cyclic monoterpene synthesis (**Supplementary Figure 8**). PbTPS3 is the dual function enzyme for producing acyclic monoterpenes, including linalool and (β)-cis-ocimene. It was expressed in flower and to a much higher extent in leaf (**Figure 3D**). However, GC-MS analysis of leaf VOCs revealed only 1% (β)-cis-ocimene but no linalool (**Supplementary Figure 9**). The heterologous expression of geraniol synthase from *Ocimum basilicum* in *E. coli* and plant systems revealed that heterologous expression affects the amount of geraniol production in the leaf tissue or *E. coli*

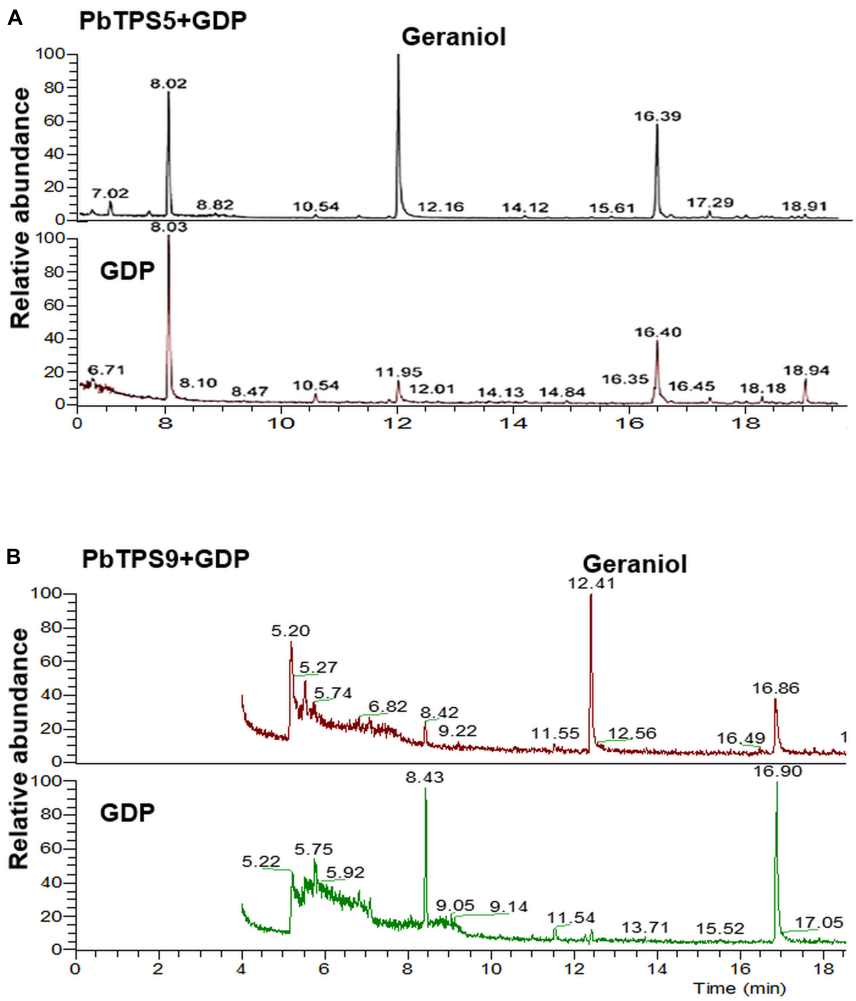


FIGURE 5 | Functional characterization of purified recombinant protein of **(A)** PbTPS5 + GDP and **(B)** PbTP9 + GDP. GC-MS analysis of authentic linalool standard incubated with enzyme assay buffer. As negative controls, reactions were performed with only GDP added without any recombinant proteins.

(Fischer et al., 2013). Therefore, the enzymatic functions of *TPS* genes depend on the enzyme amino acid sequence or substrate and also the conditions of an enzymatic reaction, especially when comparing plant and bacterial systems (Brillada et al., 2013).

Functional Characterization of *PbTPS5* and *PbTPS10*

PbTPS5 showed the enzyme activity of linalool synthase when overexpressed in Venus, and Venus itself produced only linalool and expressed *PbTPS5* slightly (**Figure 3**). Negative regulators of *PbTPS10* may be present in the Venus flowers and prevent its gene expression and production of linalool. Previously, differences in biochemical conditions between plant and microbial expression systems produce different products (Kollner et al., 2004). These differences between heterologous systems could result from the protein conformational change in the binding pocket due to different environments (e.g., pH, ion concentrations) or different posttranslational modifications (Fischer et al., 2013). For monoterpene synthesis, GDP is the precursor, and the conversion of GDP into monoterpenols involves the formation of a carbocation intermediate (Ilc et al., 2016). Any subtle differences in protein conformation altered by the environment (chemicals or other proteins) could affect the profile of carbocation intermediate, thus producing different products (Fischer et al., 2013). In addition, the presence of other enzymes that convert one monoterpenol into another monoterpenol cannot be ruled out because of the many enzymes in plants. The other factor that can be taken account in affecting the production of monoterpenes is intracellular pH. Changes in intracellular pH can cause the chemical transformation of monoterpenols. Previous study found that monoterpenols tended to interconvert under acidic condition but are stable at neutral pH (Fischer et al., 2011). Previously, we detected the pH value of various colored *Phalaenopsis* orchids and found that it increases from red-purple, white, purple-violet to violet blue color from pH 4.77 to 5.54 (Liang et al., 2020). Although the white flowers of *Phalaenopsis aphrodite* and V3 have similar pH values (5.05

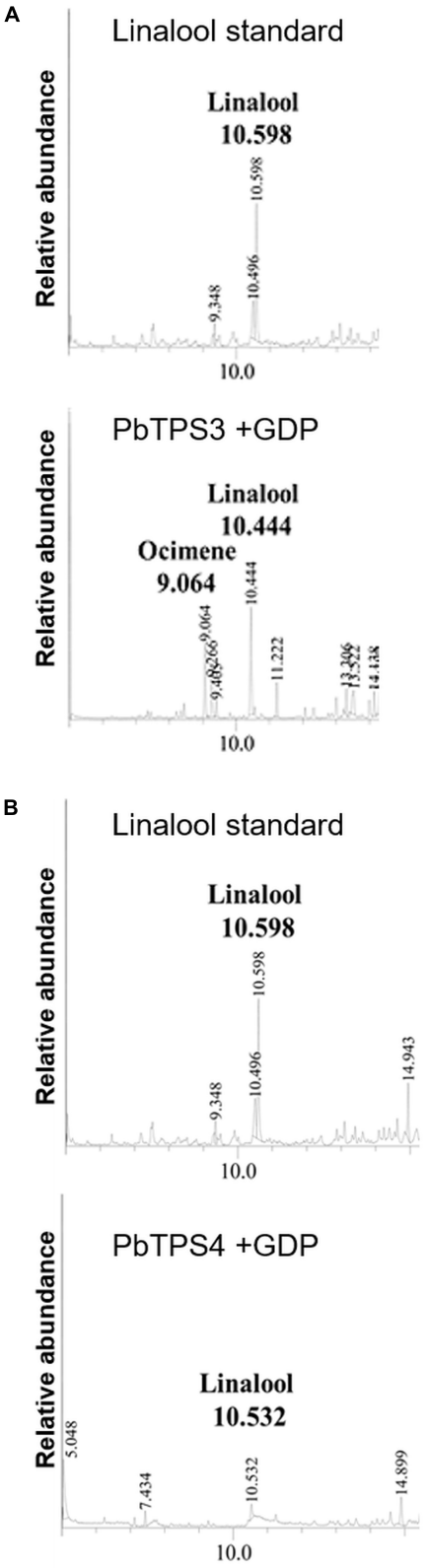


FIGURE 6 | Functional characterization of purified recombinant protein of **(A)** PbTPS3 + GDP and **(B)** PbTPS4 + GDP. GC-MS analysis of authentic linalool standard incubated with enzyme assay buffer.

and 5.13, respectively), different compounds were produced by overexpressing TPS genes in these two orchids, which suggests that the pH value was not the major effect on the content of volatiles in *Phalaenopsis* orchids (**Supplementary Figure 5**).

However, *in vitro* enzyme activity assay revealed that geraniol was the main product catalyzed by purified recombinant *PbTPS5*. With overexpression of *PbTPS5* and *PbTPS10* in Venus, linalool was the only product for both enzymes, with *PbTPS10* displaying high ability for linalool production. In addition, molecular modeling and docking were performed by using SwissDock program with both *PbTPS5* and *PbTPS10* as targets and GDP as a ligand. The ligand was difficult to dock into the active site of *PbTPS5* as compared with *PbTPS10*, which implies that the binding pocket was easier to access with *PbTPS10* than with *PbTPS5*. This finding explains why the activity of *PbTPS10* was better than that of *PbTPS5* in monoterpene synthesis. Within the flower of *P. bellina*, *PbTPS10* mainly expressed in the sepal and slightly in the petal, whereas *PbTPS5 is* mainly expressed in the petal but less in the sepal. When combining the gene expression and enzyme activity data, the monoterpene biosynthesis in different floral organs of *P. bellina* could be driven by different genes and for different scent compounds. Recent reports showed that many TPS genes can produce more than one terpenoid, which could be related to the size of the active cavity. Thus, the amino acid sequences of *TPS* genes are among the factors that affect the TPS functional property (Pazouki and Niinemets, 2016). Also, previous *TPS* studies indicated that the outcome of *in planta* experiments could be affected by the plants chosen; for instance, when geraniol synthase from *O. basilicum* was expressed in grapevine, citronellol and nerol were produced (Arendt et al., 2016). The products of *TPS* genes may vary in different host plants because the function of MTPS is based on not only the amino acid sequence but also the cellular background (Fischer et al., 2013).

Differential Positive and Negative Gene Regulation of *PbTPS5* and *PbTPS10*

Even though the coding sequences of PbTPS5 and PbTPS10 are similar, they showed distinct temporal and spatial expression profiles. For example, the temporal expression of *PbTPS5* started on the blooming date, whereas that of *PbTPS10* started on D + 3. Spatially, *PbTPS5* expressed highly in the petal, and *PbTPS10* expressed highly in the sepal, and both expressed similarly in the lip. In addition, these two genes revealed distinct profiles of transcriptional regulation (Chuang et al., 2018b). Both *PbbZIP4* and *PbNAC1* positively regulated the promoter of *PbTPS5* but did not affect the promoter of *PbTPS10* (Chuang et al., 2018b). Conversely, both *PbbHLH4* and *PbERF9* highly positively regulated the promoter of *PbTPS10* but did not affect that of *PbTPS5*. In contrast, *PbbHLH2* and *PbbZIP26* negatively regulated *PbTPS5*, and *PbbHLH5*, *PbbZIP26*, and *PbMYB22* negatively regulated *PbTPS10* (Chuang et al., 2018b). Even though both *PbTPS5* and *PbTPS10* were negatively regulated by *PbbZIP26*, only *PbTPS10* expression was inhibited in Venus, which suggests that both bHLH5 and MYB22 may function as negative regulators to prevent the expression of *PbTPS10*.

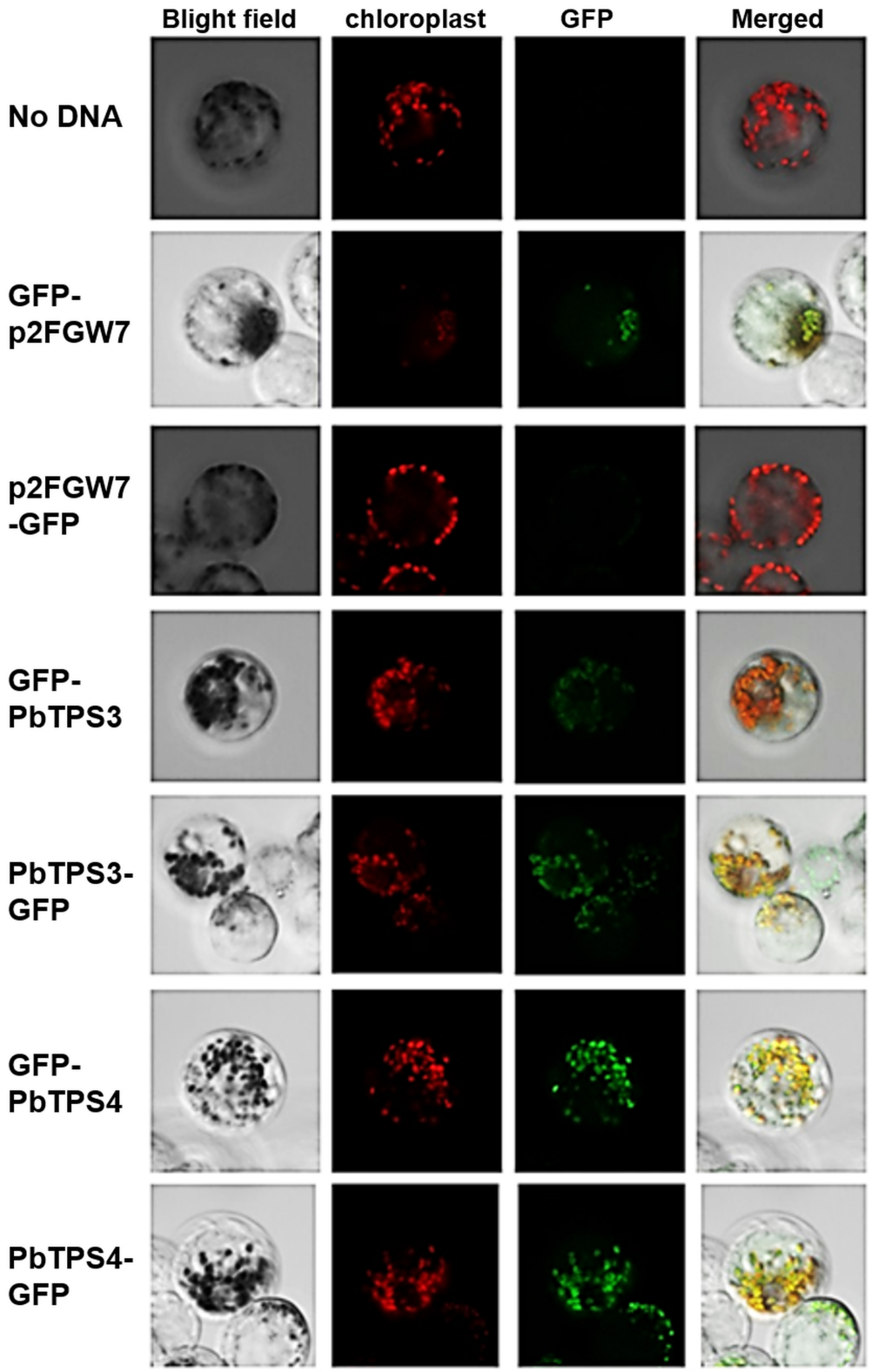


FIGURE 7 | Subcellular localization of *PbTPS3* and *PbTPS4*. Protoplasts of *Phalaenopsis aphrodite* tepals were transformed with GFP fusion proteins (GFP-PbTPS3, PbTPS3-GFP, GFP-PbTPS4, and PbTPS4-GFP). Images were taken by confocal laser scanning microscopy. Bright field, the image in bright field; Chloroplast with autofluorescence image false-colored in red; GFP, GFP fluorescence image false-colored in green; Merge, merge of bright field, chloroplast and GFP fluorescence images.

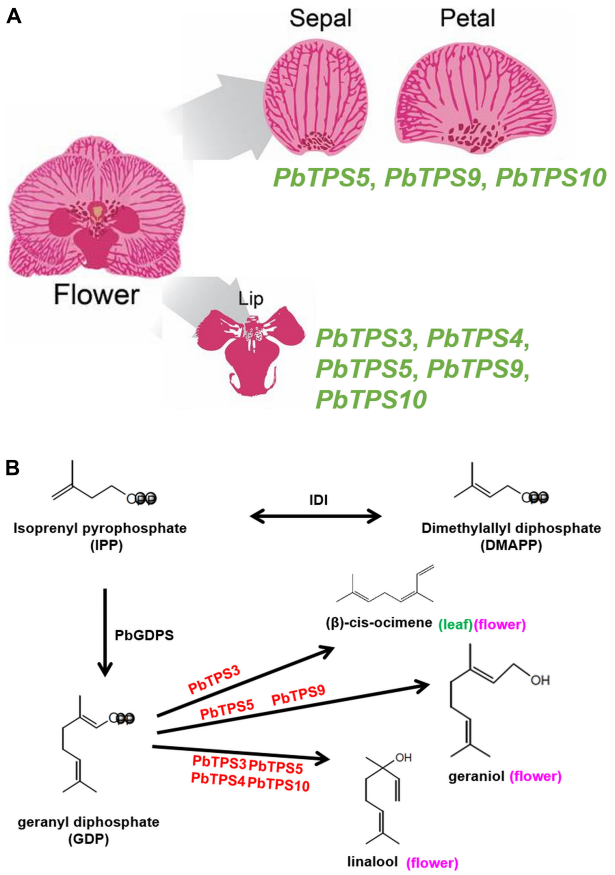


FIGURE 8 | (A) The biosynthesis pathway of linalool, geraniol, and (β)-*cis*-ocimene in flower or (β)-*cis*-ocimene in the leaf of *P. bellina* from GDP. More than one enzyme can use GDP as substrate for producing linalool (*PbTPS3*, *PbTPS4*, *PbTPS5*, and *PbTPS10*), geraniol (*PbTPS5* and *PbTPS9*), or (β)-*cis*-ocimene (*PbTPS3*). In addition, one enzyme may produce more than one product using GDP as the substrate, such as *PbTPS3* for both linalool and (β)-*cis*-ocimene biosynthesis. **(B)** Various *PbTPS* genes involved in monoterpene biosynthesis in a single flower.

MYB transcription factor has a negative regulation role in floral scent biosynthesis (Reddy et al., 2017). Further assessments will be needed to elucidate the negative regulation of floral scent biosynthesis.

Functional Characterization of *PbTPS3* and *PbTPS4*

The purified recombinant *PbTPS3* had dual enzyme functions as a linalool/(β)-cis-ocimene synthase, whereas that of the purified recombinant *PbTPS4* produced only linalool. Both *PbTPS3* and *PbTPS4* were expressed in reproductive organs, including sepal, petal, and lip. In addition, *PbTPS3* was expressed in the column at a similar level as in the sepal, whereas *PbTPS4* was expressed much lower in the column. So, the major floral scent of *P. bellina* in the flower tepal (sepal and petal) is dominated by linalool and geraniol, whereas the lip has the floral scent of (β)-cis-ocimene in addition to linalool, and the column has the floral scent of (β)-cis-ocimene. Emission of (β)-cis-ocimene is 15.8 ng/flower/h, accounting for 2% of the floral scent of *P. bellina* (Chuang et al., 2018b).

PbTPS3 and *PbTPS4* Are New Members of the *TPS-e/f* Subfamily

TPS enzymes in the *TPS-e/f* subfamily have dual functions or play different roles. For example, the (β)-cis-ocimene/farnesene synthase of *Actinidia deliciosa* has dual functions, and it produces the (β)-cis-ocimene and farnesene mixture for pollinator attraction (Nieuwenhuizen et al., 2009). The foregoing observations indicate that the *TPS* genes in the *TPS-e/f* subfamily have various ecological roles in plants, including floral scent biosynthesis and plant defense. For example, linalool synthase is involved in linalool biosynthesis in *Clarkia concinna* and *C. breweri*. The expression of linalool synthase in the stigma of *C. breweri* leads to the emission of linalool and linalool oxides to serve as pollinator attractants (Dudareva et al., 1996). In addition, the geranyllinalool synthase of *A. thaliana* can produce the diterpene product of geranyllinalool for plant defense (Herde et al., 2008). Thus, the (β)-cis-ocimene/linalool synthase *PbTPS3* and linalool synthase *PbTPS4* in *P. bellina* are new members in the *TPS-e/f* subfamily.

Both linalool and (β)-cis-ocimene have a defense role against insect herbivores in many plants such as cotton, soybean, and rice (Pare and Tumlinson, 1999). The promoter of *PbTPS3* contains three and 10 binding domains of the plant defense transcription factors WBOXATNPR1 and WRKY71OS, respectively, and the promoter of *PbTPS4* contains three binding domains for WRKY71OS (**Supplementary Figure 10**). Thus, both *PbTPS3* and *PbTPS4* may play a role in defense in *P. bellina*. *PbTPS3* may play different roles in different organs of *P. bellina*. In flower, *PbTPS3* plays a role in linalool synthesis to attract pollinators but plays a defense role in the leaf.

The Evolution and Diversity of the *TPS-e/f* Subfamily

Linalool synthases have been identified in two *TPS* subfamilies, namely, *TPS-b* and *TPS-e/f*, so linalool synthase functionality has arisen multiple times independently during evolution (Nieuwenhuizen et al., 2009). The *TPS-e/f* subfamily is unique in that the members contain a conserved conifer diterpene internal sequence (CDIS) domain. Thus, the CDIS domain exists in a common DTPS ancestor but was lost during the evolution of MTPS and STPS (Brodmann et al., 2008). The *TPS-f* subfamily is probably dicot-specific, whereas the *TPS-e* subfamily is not (Chen et al., 2011). The *TPS-e/f* subfamily has the DDXXD conserved motif and lacks the DXDD conserved motif. From multiple alignment of amino acid sequences of the TPS-e/f subfamily in *P. equestris*, *P. bellina*, and other plants, *PbTPS3* and *PbTPS4* are cladded in the *TPS-e/f* subfamily (**Figure 1**; **Supplementary Figures 4**, **8**). Our study indicates that *PbTPS3* and *PbTPS4* from orchids belong to the *TPS-e/f* subfamily and enlarge the content of this subfamily.

The subcellular localization analysis indicated that both *PbTPS3* and *PbTPS4* GFP fusion proteins are located in the plastids. The temporal expression patterns of *PbTPS3* and *PbTPS4* were concomitant with the volatile emission pattern of *P. bellina*. Taken together, *PbTPS3* and *PbTPS4* play different ecological roles of (β)-cis-ocimene/linalool and linalool biosynthesis in *P. bellina*.

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 in the article/ **Supplementary Material**.

AUTHOR CONTRIBUTIONS

HH, Y-WK, and Y-CC performed the experiments of quantitative real-time RT-PCR, did functional characterization, and analyzed the data. Y-PY performed the original screening. L-MH performed the phylogenetic analysis. M-FJ assisted the sequence analysis and molecular docking. W-HC provided the suggestions for plant materials. H-HC conceived research plans and composed the article with assistances of all the authors, completed the writing, and served as the corresponding author for communication. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by a grant from the Ministry of Science and Technology, Taiwan (MOST 107-2313-B-006-003-MY3), to H-HC.

ACKNOWLEDGMENTS

We thank Sheng-Yang Wang (Department of Forestry, National Chung Hsing University) and Fang-Hua Chu (Department of Forestry, National Taiwan University) for assistance with *in vitro* enzyme activity assay.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2021.700958/full#supplementary-material

REFERENCES

Aharoni, A., Giri, A. P., Deuerlein, S., Griepink, F., De Kogel, W. J., Verstappen, F. W., et al. (2003). Terpenoid metabolism in wild-type and transgenic *Arabidopsis* plants. *Plant Cell* 15, 2866–2884. doi: 10.1105/tpc.016253

Arendt, P., Pollier, J., Callewaert, N., and Goossens, A. (2016). Synthetic biology for production of natural and new-to-nature terpenoids in photosynthetic organisms. *Plant J.* 87, 16–37. doi: 10.1111/tpj.13138

Bohlmann, J., Meyer-Gauen, G., and Croteau, R. (1998). Plant terpenoid synthases: molecular biology and phylogenetic analysis. *Proc. Natl. Acad. Sci. U. S. A.* 95, 4126–4133. doi: 10.1073/pnas.95.8.4126

Brillada, C., Nishihara, M., Shimoda, T., Garms, S., Boland, W., Maffei, M. E., et al. (2013). Metabolic engineering of the C16 homoterpene TMTT in *Lotus japonicus* through overexpression of (E,E)-geranyllinalool synthase attracts generalist and specialist predators in different manners. *New Phytol.* 200, 1200–1211. doi: 10.1111/nph.12442

Brodmann, J., Twele, R., Francke, W., Holzler, G., Zhang, Q. H., and Ayasse, M. (2008). Orchids mimic green-leaf volatiles to attract prey-hunting wasps for pollination. *Curr. Biol.* 18, 740–744. doi: 10.1016/j.cub.2008.04.040

Chen, F., Tholl, D., Bohlmann, J., and Pichersky, E. (2011). The family of terpene synthases in plants: a mid-size family of genes for specialized metabolism that is highly diversified throughout the kingdom. *Plant J.* 66, 212–229. doi: 10.1111/j.1365-313X.2011.04520.x

Chen, Y. H., Tsai, Y. J., Huang, J. Z., and Chen, F. C. (2005). Transcription analysis of peloric mutants of *Phalaenopsis* orchids derived from tissue culture. *Cell Res.* 15, 639–657. doi: 10.1038/sj.cr.7290334

Chuang, Y. C., Hung, Y. C., Hsu, C. Y., Yeh, C. M., Mitsuda, N., Ohme-Takagi, M., et al. (2018a). A dual repeat cis-element determines expression of GERANYL DIPHOSPHATE SYNTHASE for monoterpene production in *Phalaenopsis* orchids. *Front. Plant Sci.* 9:765. doi: 10.3389/fpls.2018.00765

Chuang, Y. C., Hung, Y. C., Tsai, W. C., Chen, W. H., and Chen, H. H. (2018b). PbbHLH4 regulates floral monoterpene biosynthesis in *Phalaenopsis* orchids. *J. Exp. Bot.* 69, 4363–4377. doi: 10.1093/jxb/ery246

Chuang, Y. C., Lee, M. C., Chang, Y. L., Chen, W. H., and Chen, H. H. (2017). Diurnal regulation of the floral scent emission by light and circadian rhythm in the *Phalaenopsis* orchids. *Bot. Stud.* 58:50. doi: 10.1186/s40529-017-0204-8

Colquhoun, T. A., Kim, J. Y., Wedde, A. E., Levin, L. A., Schmitt, K. C., Schuurink, R. C., et al. (2011). PhMYB4 fine-tunes the floral volatile signature of *Petunia x hybrida* through PhC4H. *J. Exp. Bot.* 62, 1133–1143. doi: 10.1093/jxb/erq342

Dudareva, N., Cseke, L., Blanc, V. M., and Pichersky, E. (1996). Evolution of floral scent in *Clarkia*: novel patterns of S-linalool synthase gene expression in the C-breweri flower. *Plant Cell* 8, 1137–1148. doi: 10.1105/tpc.8.7.1137

Fischer, M. J., Meyer, S., Claudel, P., Bergdoll, M., and Karst, F. (2011). Metabolic engineering of monoterpene synthesis in yeast. *Biotechnol. Bioeng.* 108, 1883–1892. doi: 10.1002/bit.23129

Fischer, M. J. C., Meyer, S., Claudel, P., Perrin, M., Ginglinger, J. F., Gertz, C., et al. (2013). Specificity of *Ocimum basilicum* geraniol synthase modified by its expression in different heterologous systems. *J. Biotechnol.* 163, 24–29. doi: 10.1016/j.jbiotec.2012.10.012

Herde, M., Gartner, K., Kollner, T. G., Fode, B., Boland, W., Gershenzon, J., et al. (2008). Identification and regulation of TPS04/GES, an *Arabidopsis* geranyllinalool synthase catalyzing the first step in the formation of the insect-induced volatile C(16)-homoterpene TMTT. *Plant Cell* 20, 1152–1168. doi: 10.1105/tpc.106.049478

Hsiao, Y. Y., Jeng, M. F., Tsai, W. C., Chuang, Y. C., Li, C. Y., Wu, T. S., et al. (2008). A novel homodimeric geranyl diphosphate synthase from the orchid *Phalaenopsis bellina* lacking a DD(X)(2-4)D motif. *Plant J.* 55, 719–733. doi: 10.1111/j.1365-313X.2008.03547.x

Hsiao, Y. Y., Tsai, W. C., Kuoh, C. S., Huang, T. H., Wang, H. C., Wu, T. S., et al. (2006). Comparison of transcripts in *Phalaenopsis* bellina and *Phalaenopsis equestris* (Orchidaceae) flowers to deduce monoterpene biosynthesis pathway. *BMC Plant Biol.* 6:14. doi: 10.1186/1471-2229-6-14

Hsieh, M. H., Pan, Z. J., Lai, P. H., Lu, H. C., Yeh, H. H., Hsu, C. C., et al. (2013). Virus-induced gene silencing unravels multiple transcription factors involved in floral growth and development in *Phalaenopsis* orchids. *J. Exp. Bot.* 64, 3869–3884. doi: 10.1093/jxb/ert218

Hsu, C. C., Chen, Y. Y., Tsai, W. C., Chen, W. H., and Chen, H. H. (2015). Three R2R3-MYB Transcription factors regulate distinct floral pigmentation

patterning in *Phalaenopsis* spp. *Plant Physiol.* 168, 175–191. doi: 10.1104/pp.114.254599

Ilc, T., Parage, C., Boachon, B., Navrot, N., and Werck-Reichhart, D. (2016). Monoterpenol oxidative metabolism: role in plant adaptation and potential applications. *Front. Plant Sci.* 7:509. doi: 10.3389/fpls.2016.00509

Kollner, T. G., Schnee, C., Gershenzon, J., and Degenhardt, J. (2004). The sesquiterpene hydrocarbons of maize (Zea mays) form five groups with distinct developmental and organ-specific distribution. *Phytochemistry* 65, 1895–1902. doi: 10.1016/j.phytochem.2004.05.021

Liang, C. Y., Rengasamy, K. P., Huang, L. M., Hsu, C. C., Jeng, M. F., Chen, W. H., et al. (2020). Assessment of violet-blue color formation in *Phalaenopsis* orchids. *BMC Plant Biol.* 20:212. doi: 10.1186/s12870-020-02402-7

Lichtenthaler, H. K., Schwender, J., Disch, A., and Rohmer, M. (1997). Biosynthesis of isoprenoids in higher plant chloroplasts proceeds via a mevalonate-independent pathway. *FEBS Lett.* 400, 271–274. doi: 10.1016/s0014-5793(96)01404-4

Nieuwenhuizen, N. J., Wang, M. Y., Matich, A. J., Green, S. A., Chen, X. Y., Yauk, Y. K., et al. (2009). Two terpene synthases are responsible for the major sesquiterpenes emitted from the flowers of kiwifruit (*Actinidia deliciosa*). *J. Exp. Bot.* 60, 3203–3219. doi: 10.1093/jxb/erp162

Pare, P. W., and Tumlinson, J. H. (1999). Plant volatiles as a defense against insect herbivores. *Plant Physiol.* 121, 325–331. doi: 10.1104/pp.121.2.325

Pazouki, L., and Niinemets, U. (2016). Multi-Substrate Terpene Synthases: their occurrence and physiological significance. *Front. Plant Sci.* 7:1019. doi: 10.3389/fpls.2016.01019

Reddy, V. A., Wang, Q., Dhar, N., Kumar, N., Venkatesh, P. N., Rajan, C., et al. (2017). Spearmint R2R3-MYB transcription factor MsMYB negatively regulates monoterpene production and suppresses the expression of geranyl diphosphate synthase large subunit (MsGPPS.LSU). *Plant Biotechnol. J.* 15, 1105–1119. doi: 10.1111/pbi.12701

Sanguinetti, A., Buzatto, C. R., Pedron, M., Davies, K. L., Ferreira, P. M., Maldonado, S., et al. (2012). Floral features, pollination biology and breeding system of *Chloraea membranacea* Lindl. (Orchidaceae: Chloraeinae). *Ann. Bot.* 110, 1607–1621. doi: 10.1093/aob/mcs221

Schiestl, F. P., Ayasse, M., Paulus, H. F., Lofstedt, C., Hansson, B. S., Ibarra, F., et al. (2000). Sex pheromone mimicry in the early spider orchid (*ophrys sphegodes*): patterns of hydrocarbons as the key mechanism for pollination by sexual deception. *J. Comp. Physiol. A.* 186, 567–574. doi: 10.1007/s003590000112

Shen, Y., Meng, D., Mcgrouther, K., Zhang, J., and Cheng, L. (2017). Efficient isolation of *Magnolia* protoplasts and the application to subcellular localization of *MdeHSF1*. *Plant Methods* 13:44. doi: 10.1186/s13007-017-0193-3

Spitzer-Rimon, B., Marhevka, E., Barkai, O., Marton, I., Edelbaum, O., Masci, T., et al. (2010). *EOBII*, a gene encoding a flower-specific regulator of phenylpropanoid volatiles' biosynthesis in Petunia. *Plant Cell* 22, 1961–1976. doi: 10.1105/tpc.109.067280

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.

frontiers in Plant Science

ORIGINAL RESEARCH
published: 03 August 2021
doi: 10.3389/fpls.2021.710826

Auxin-Responsive R2R3-MYB Transcription Factors HcMYB1 and HcMYB2 Activate Volatile Biosynthesis in *Hedychium coronarium* Flowers

Yanguo Ke[1,2†], Farhat Abbas[1†], Yiwei Zhou[1], Rangcai Yu[3] and Yanping Fan[1,4]*

[1] The Research Center for Ornamental Plants, College of Forestry and Landscape Architecture, South China Agricultural University, Guangzhou, China, [2] College of Economics and Management, Kunming University, Kunming, China, [3] College of Life Sciences, South China Agricultural University, Guangzhou, China, [4] Guangdong Key Laboratory for Innovative Development and Utilization of Forest Plant Germplasm, South China Agricultural University, Guangzhou, China

OPEN ACCESS

Edited by:
Renata Rivera-Madrid, Scientific Research Center of Yucatán (CICY), Mexico

Reviewed by:
Qian Shen, Shanghai Jiao Tong University, China
Yang Qu, University of New Brunswick, Canada

***Correspondence:**
Yanping Fan
fanyanping@scau.edu.cn

[†] 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: *17 May 2021*
Accepted: *13 July 2021*
Published: *03 August 2021*

Citation:
Ke Y, Abbas F, Zhou Y, Yu R and Fan Y (2021) Auxin-Responsive R2R3-MYB Transcription Factors HcMYB1 and HcMYB2 Activate Volatile Biosynthesis in Hedychium coronarium Flowers. Front. Plant Sci. 12:710826. doi: 10.3389/fpls.2021.710826

Auxin, an important plant hormone, induces the biosynthesis of various secondary metabolites by modulating the expression of auxin-responsive genes. In the ornamental plant *Hedychium coronarium*, linalool and methyl benzoate are biosynthesized by the terpene synthase (TPS) HcTPS5 and the benzoic/salicylic acid methyltransferase (BSMT) HcBSMT2, respectively. However, the transcriptional regulation of this process remains unclear. Here, we identified and functionally characterized the R2R3-MYB transcription factors HcMYB1 and HcMYB2 in regulating the biosynthesis of these floral aroma compounds. *HcMYB1* and *HcMYB2* are specifically expressed in flowers, their expression is correlated with the emission of volatile compounds in flowers, and is induced by auxin. Moreover, HcMYB1 and HcMYB2 interact with the *HcBSMT2* promoter region. HcMYB2 activates the expression of the linalool synthase gene *HcTPS5*. In flowers with *HcMYB1* or *HcMYB2* silenced, the levels of floral scent compounds were significantly reduced, and *HcBSMT2* and *HcTPS5* were downregulated compared with the wild type. Moreover, HcMYB1 form protein-protein interaction with key scent-related HcIAA4 protein to regulate floral aroma production. Taken together, these results indicate that HcMYB1 and HcMYB2 play crucial roles in regulating the formation of scent compounds in *Hedychium coronarium (H. coronarium)* flowers in response to auxin signaling.

Keywords: *Hedychium coronarium*, floral scent, auxin, MYB transcription factors, biosynthesis

Abbreviations: AbA, aureobasidin A; ABA, abscisic acid; Aux/IAA, auxin/indole-3-acetic acid; BALD, benzaldehyde dehydrogenase; BiFC, bimolecular fluorescence complementation; BSMT, salicylic acid/benzoic acid methyltransferase; cDNA, complementary DNA; DMAPP, dimethylallyl pyrophosphate; FPP, farnesyl diphosphate; FPPS, FPP synthase; GC-MS, gas chromatography-mass spectrometer; GFP, green fluorescent protein; IAA, indole-3-acetic acid; NLS, nuclear localization signal; OD, optical density; PAL, phenylalanine ammonia lyase; SD, synthetically defined medium; TPS, terpene synthase; Y2H, yeast two-hybrid; Y1H, yeast one-hybrid.

INTRODUCTION

Floral scent compounds are among the most important secondary metabolites in plants and comprise three major groups based on their origins: terpenoids, phenylpropanoids/benzenoids, and fatty acid derivatives (Dudareva et al., 2013; Muhlemann et al., 2014; Abbas et al., 2017). These compounds are attractive not only to humans but also to pollinators that facilitate fertilization, and thus they play key roles in plant evolution and the plant lifecycle (Raguso, 2009). The anti-herbivore or antimicrobial activity of volatiles released from flowers protects the vulnerable reproductive parts of the plant against pathogen attack (Dudareva et al., 2006). Floral scent also is an important trait that increases the aesthetic values of ornamental plants to humans (Dudareva et al., 2006; Pichersky and Dudareva, 2007), and scented compounds derived from flowers are widely used as flavorings, in cosmetics and perfumes, and as medicinal substances (Gershenzon and Dudareva, 2007; Muhlemann et al., 2014). However, notwithstanding the importance of floral scents to both plant biology and industry, little is known about the transcriptional regulation of this process.

MYB transcription factors are important regulators of the biosynthesis of plant secondary metabolites, such as phenylpropanoids (Liu L. et al., 2015; Zhou and Memelink, 2016; Ramya et al., 2017). MYB proteins possess two regions: a conserved MYB DNA-binding domain at the N-terminus and a diverse modulator region at the C-terminus that is responsible for their regulatory activity. MYB transcription factors (TFs) are classified into four subunits/groups based on the number of adjacent repeats in the DNA-binding domains: the R2R3-, 1R-, 3R-, and 4R-MYBs (Dubos et al., 2008). Most MYB TFs involved in regulating secondary metabolite biosynthesis in flowers belong to the R2R3-MYB family (Ramya et al., 2017). To date, however, only a few MYB TFs identified from floral scent model species, i.e., snapdragon (*Antirrhinum majus*) and petunia (*Petunia* spp.) have been shown to regulate the expression of structural genes related to volatile biosynthetic pathways. The R2R3-MYB TFs ODORANT1 (ODO1) and EMISSION OF BENZENOID II (EOBII) regulate volatile biosynthesis genes in the benzenoid/phenylpropanoid pathway in petunia (Verdonk et al., 2005; Spitzer-Rimon et al., 2012). In snapdragon, the MYBs AmMYB305 and AmMYB340 are involved in regulating the volatile phenylpropanoid/benzenoid metabolic pathway (Uimari and Strommer, 1997; Shin et al., 2002). In addition, the R2R3-MYB TF FaEOBII regulates the production of the volatile eugenol in ripe strawberry (*Fragaria* × *ananassa*) receptacles by activating the expression of *CINNAMYL ALCOHOL DEHYDROGENASE* (*FaCAD1*) and *EUGENOL SYNTHASE* (*FaEGS2*) (Medina-Puche et al., 2015). In the *Cymbidium* orchid cultivar "Sael Bit," *CsMYB1* is highly expressed in floral organs and is involved in regulating the biosynthesis of floral volatiles such as polyacrylate and 2-methyl butyraldehyde in petals (Ramya et al., 2019).

Floral and fruit volatiles are also regulated by the essential plant hormone auxin, which induces the biosynthesis of numerous secondary metabolites by regulating the expression of auxin-responsive genes (Zhou and Memelink, 2016). Treatment with exogenous auxin increases the emission of the volatile compound linalool in apple (*Malus domestica*) scions (Kim et al., 2011) and modifies the quantity of fruit flavor compounds. In strawberry, auxin treatment enhances the aggregation of phenolic volatiles such as 2-phenylethanol, phenylacetaldehyde, and methyl benzoate, and inhibits the production of benzyl cyanide, 2-isobutylthiazole, 1-hexanol, and 1-nitro-2-phenylethane (Wu et al., 2018). Exogenous auxin treatment also modifies the expression of several key genes associated with the biosynthetic pathways of scent volatiles, including *PHENYL ALDEHYDE REDUCTASE 1* (*SlPAR1*), *SlPAR2*, and *SlSAMT1*, in tomato (*Solanum lycopersicum*) (Wu et al., 2018). In grapefruit (*Citrus* × *paradisi*), auxin treatment influences sugar accumulation in various ways, as well as the accumulation of volatile compounds and the expression of aroma-related genes (Jia et al., 2017). Many MYB TF genes respond to auxin signalings, such as the *Arabidopsis thaliana* genes *AtMYB44*, *AtMYB77*, and *AtMYB108* (Shin et al., 2007; Dubos et al., 2010) and ten R2R3-MYB genes in cassava (*Manihot esculenta*) (Liao et al., 2016). Nevertheless, how auxin is involved in regulating the phenylpropanoid and terpenoid biosynthetic pathway via MYB TFs was not known.

Hedychium coronarium is a perennial herb of the Zingiberaceae family that is cultivated as a cut flower, garden plant, and medicinal plant and for aromatic oil production. At blooming, *H. coronarium* flowers emit large amounts of volatile compounds, including the monoterpenes linalool and (*E*,Z)-β-ocimene, and benzenoids such as methyl benzoate (Fan et al., 2003, 2007; Li and Fan, 2011; Lan et al., 2013; Yue et al., 2015). We previously identified several structural genes in the *H. coronarium* volatile biosynthetic pathway, including genes encoding terpene synthases (TPSs) and benzoic/salicylic acid methyltransferase (BSMT). A total of 12 *HcBSMT* and 62 *HcTPS* genes were found in *H. coronarium*. *HcBSMT2* specifically expressed in flowers, its expression level was enormously high among all *HcBSMT* genes and correlated with flower development (**Supplementary Figure 6**). Likewise, the expression values of *HcTPS3* and *HcTPS5* were tremendously high among all the *HcTPSs* and specifically expressed in *H. coronarium* flowers (**Supplementary Figure 7**). Moreover, their expression pattern positively correlated with flower development as well as with the emission of monoterpenes, and their encoded enzymes localize to plastids (Yue et al., 2015). Our functional characterization indicated that HcTPS3 functions in (*E*)-β-ocimene production, HcTPS5 functions in linalool production, and HcBSMT2 functions in methyl benzoate production (Yue et al., 2015, 2021). Analysis of previously generated RNA-seq data showed that six *HcMYB* family members were clustered in a group involved in regulating secondary metabolism. The expression levels of these MYB family members were analyzed in different tissues (flowers, bracts, leaves, and rhizomes). The RNA-sequence data showed that among six *HcMYBs*, *HcMYB1*, and *HcMYB2* were highly flower-specific and the abundance of their transcripts correspond with the flower development as well as with the emission of floral volatile contents. The relative transcript abundance of six *HcMYB*

family members has been provided in **Supplementary Figure 3**. However, the hormone-responsive transcriptional regulation of these genes has not been elucidated. Furthermore, we comprehensively analyzed *Aux/IAA* genes in *H. coronarium* genome. The genome−wide analysis and characterization of *Aux/IAA* genes reveal the potential role of *HcIAA2* and *HcIAA4* in floral aroma production in *H. coronarium* (Ke et al., 2019).

In the current study, we functionally characterized two R2R3-MYB TF genes (*HcMYB1* and *HcMYB2*) that are expressed in an auxin-responsive manner specifically in flowers. These MYB TFs regulate phenylpropanoid/benzenoid and terpenoid biosynthesis specifically in *H. coronarium* flowers by activating *HcBSMT2* and *HcTPS5* expression. Furthermore, the interaction of HcMYB1 with key scent-related auxin protein (HcIAA4) was revealed via yeast two-hybrid (Y2H) assay and bimolecular fluorescence complementation (BiFC) assays. These findings shed light on the mechanism underlying the emission of floral scent compounds in *H. coronarium*.

RESULTS

Identification and Characterization of Scent-Related R2R3-MYB Family Members

In a previous transcriptomic analysis, we identified a clade of genes whose expression rose throughout flower development and with increasing floral scent emissions (Yue et al., 2015). Among these genes, *HcMYB1* and *HcMYB2* are specifically expressed in flowers. The full-length complementary DNA (cDNA) sequences of *HcMYB1* and *HcMYB2* contain open reading frames (ORFs) of 618 and 693 bp, encoding polypeptides of 205 and 230 amino acid residues with molecular masses of 23.26 and 24.68 kDa, respectively. Analysis of the predicted protein sequences of *HcMYB1* and *HcMYB2* revealed the presence of 2R and 3R repeat signatures at the N-termini: these features of R2R3 DNA-binding MYB proteins (**Figure 1A**) are essential for their interactions with regulatory sequences in the promoters of their target genes (Kranz et al., 1998; Dubos et al., 2010; Medina-Puche et al., 2014).

We performed a phylogenetic analysis of HcMYB1 and HcMYB2 compared to R2R3-MYBs involved in secondary metabolism in other plant species. HcMYB1 and HcMYB2 clustered into different groups (**Figure 1B**). HcMYB1 belongs to Group III and shares high amino acid homology with AtMYB77 and AtMYB44 (Aharoni et al., 2001; Shin et al., 2007; Jaradat et al., 2013). In contrast, HcMYB2 belongs to Group I, whose members include AmMYB305, AmMYB340 (*A. majus*), AtMYB24 (*A. thaliana*), AtMYB21 (*A. thaliana*), PhEOBII (*Petunia* × *hybrida*), FaEOBII (*Fragaria* × *ananassa*) NlMYB305 (*Nicotiana langsdorffii*), and PsMYB26 (*Pisum sativum*) (Uimari and Strommer, 1997; Shin et al., 2002; Li et al., 2006b; Liu et al., 2009; Spitzer-Rimon et al., 2010; Medina-Puche et al., 2015). The two proteins clustered into different groups and may have different functions and/or operate through different pathways to take part in floral volatile production. Furthermore, to interrogate the evolutionary relationship of six HcMYB family members with Arabidopsis MYBs, a phylogenetic tree was built. The phylogenetic analysis revealed that all MYB proteins can be clustered into five different groups (G A–G F). HcMYB1 was clustered into group G B, HcMYB2/6/5 belongs to group G D, while HcMYB3/4 was grouped into G A (**Supplementary Figure 1**).

HcMYB1 and *HcMYB2* Are Expressed During Flower Development and in Response to Auxin

The accumulation of floral volatiles increases as flower development proceeds (Yue et al., 2015; Abbas et al., 2019).

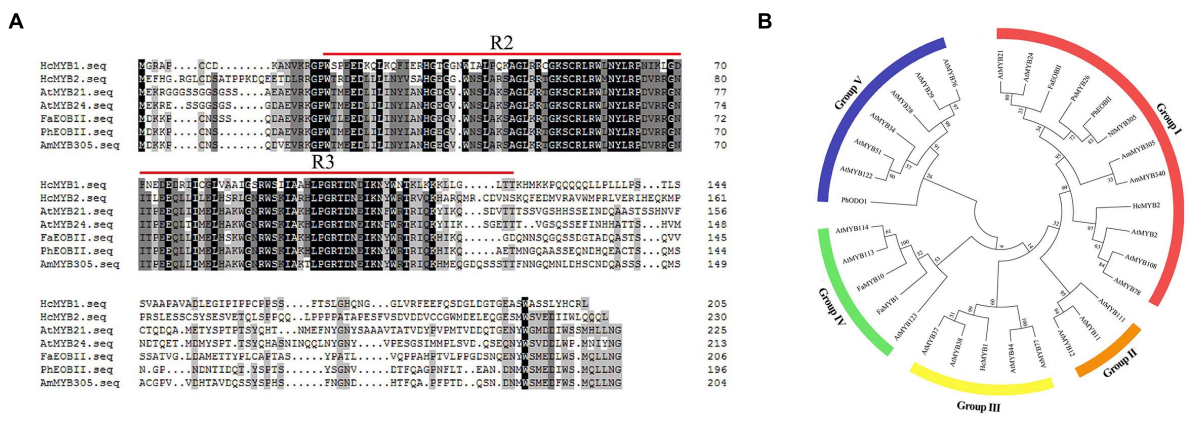


FIGURE 1 | Multiple sequence alignment and phylogenetic analysis of R2R3-MYB proteins. **(A)** Multiple sequence alignment of HcMYB1 and HcMYB2 with other related MYBs from different plants. The conserved R2 and R3 motifs are underlined. Multiple sequence alignment was performed using Clustal Ω and edited with GeneDoc. **(B)** Phylogenetic analysis of HcMYB1 and HcMYB2 together with selected R2R3-MYB proteins from different plants. Amino acid sequence alignment was performed using Clustal Ω, and the tree was constructed using the NJ method in MEGA X. The bootstrap values were set to 1000. The accession numbers for the protein sequences are listed in **Supplementary Table 3**. The sequence info of HcMYB1 and HcMYB2 is given in **Supplementary Material**.

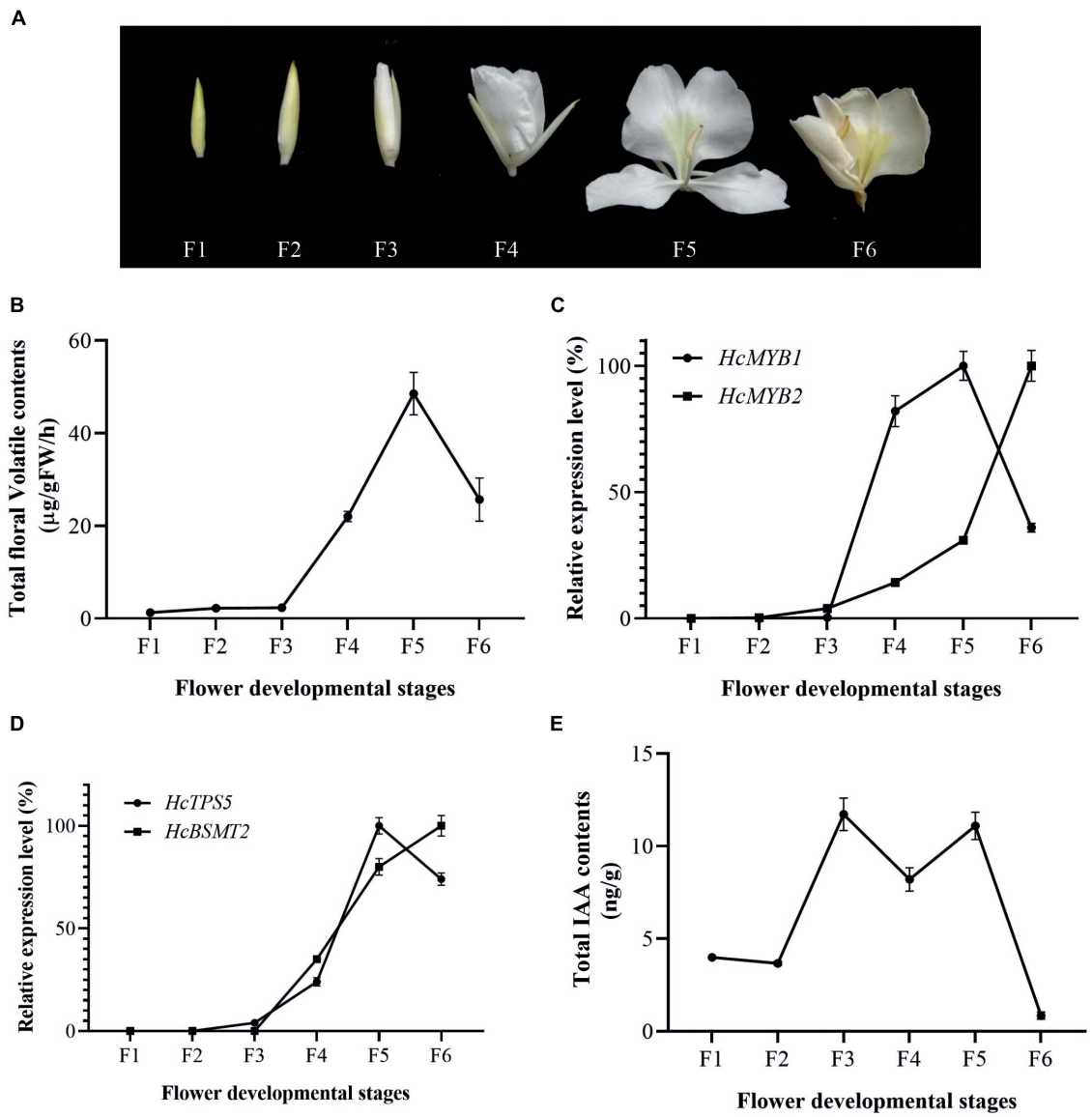


FIGURE 2 | Changes in gene expression, IAA, and volatiles over the course of flower development **(A)** Pictorial view of different stages in floral development in *H. coronarium*. **(B)** Total floral volatile contents during flower development (F1–F6). **(C)** Expression analysis of *HcMYB1* and *HcMYB2* during different stages of flower development. **(D)** Relative expression levels of *HcTPS5* and *HcBSMT2*. **(E)** Endogenous auxin contents during flower development. Bars represent the means ± SD (n = 3–5).

To study this process, we divided the flower development process into six stages (**Figure 2A**). The emission of floral volatiles was low during the bud period (F1 and F2), substantially increased beginning at the initial flowering stage (F3), peaked during the full-bloom stage (F4 and F5), and declined at the senescence stage (F6) (**Figure 2B**).

To examine the relationship between *HcMYB1* and *HcMYB2* and key volatile biosynthesis genes (*HcTPS5* and *HcBSMT2*) involved in floral volatile contents, we measured the expression levels of these genes. *HcMYB1* and *HcMYB2* transcript levels were low during early flower development and substantially increased thereafter. *HcMYB1* expression peaked at the full-bloom stage (F4–F5) and decreased at the senescence stage (F6), whereas *HcMYB2* was most strongly expressed at F6 (**Figure 2C**). We detected similar expression patterns for *HcTPS5* and *HcBSMT2* during flower development (**Figure 2D**). Moreover, the expression levels of *HcMYB1* and *HcMYB2* were positively correlated with the emissions of floral volatiles; this correlation was highly significant for *HcMYB1* (**Supplementary Figure 2**). Moreover, HcMYB1 showed a highly significant correlation with the emission of linalool contents (**Supplementary Figure 3**). These results suggest that these genes play important roles in floral scent formation in *H. coronarium*.

Auxin plays a crucial role throughout flower development (Krizek, 2011; Ke et al., 2018). We, therefore measured total auxin levels in *H. coronarium* during flower development (F1–F6). The total auxin contents were low during F1 and F2, peaked at F3–F5, dropped slightly at F4, and declined further at F6 (**Figure 2E**). The emission of total floral volatiles was correlated with indole-3-acetic acid (IAA) contents, suggesting that auxin might play a crucial role in the biosynthesis of these compounds. Under IAA treatment, the contents of the major floral volatiles ocimene, linalool, and methyl benzoate increased by 16, 17, and 20%, respectively, compared to those in control flowers not treated with IAA (CK) (**Figure 3A**). Moreover, the expression of key structural volatile biosynthesis genes (*HcTPS1*, *HcTPS3*, *HcTPS5*, *HcTPS8*, *HcPAL*, and *HcBSMT2*) was upregulated by this treatment (**Figure 3B**).

To characterize the expression levels of *HcMYB1* and *HcMYB2* in response to IAA treatment, we performed qRT-PCR analysis. *HcMYB1* and *HcMYB2* transcript levels strongly increased after IAA treatment, reaching their highest levels at 12 h after treatment. In contrast, a rapid increase in *HcMYB2* and *HcTPS5* expression was observed at 2 h after treatment (**Figure 3C**). Volatile biosynthesis genes (*HcTPS5* and *HcBSMT2*) were also upregulated at 12 h after IAA treatment (**Figure 3D**). The results suggest that the biosynthesis of floral volatiles is spatially and temporally regulated by HcMYB1 and HcMYB2, which are strongly associated with auxin-induced volatile emissions in *H. coronarium*.

We also examined the effect of p-chlorophenoxyisobutyric acid (PCIB) (inhibit auxin action) on the floral volatile compounds (**Figure 4**). In contrast to auxin, the emission of floral volatile compounds decreases. Under PCIB treatment, the contents of eucalyptol, allo-ocimene, β-ocimene, methyl benzoate, and linalool were decreased by 57, 81, 89, 100, and 42%, respectively, compared to those in control flowers not treated with PCIB (**Figure 4A**). As expected, similar to IAA, the volatile contents of caryophyllene do not change significantly. We perform qRT-PCR analysis to characterize the expression level of key genes under PCIB treatment. The expression level of key volatile biosynthesis genes *HcTPS1*,

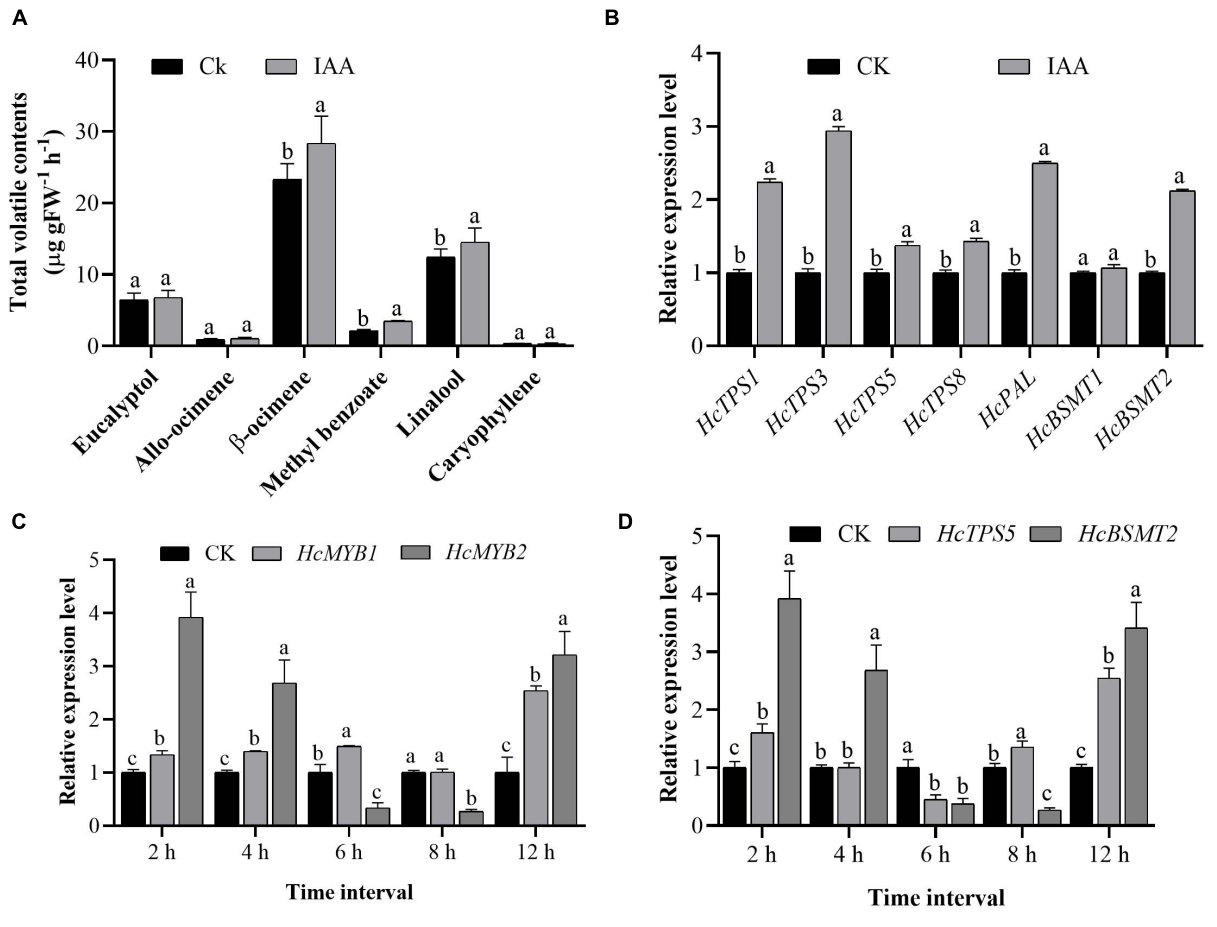


FIGURE 3 | Exogenous auxin treatment influences the expression of *HcMYB1*, *HcMYB2*, and key volatile biosynthesis genes. **(A)** The emission of major floral volatile compounds in *H. coronarium* after auxin treatment. **(B)** Relative expression levels of key volatile biosynthesis genes after auxin treatment. **(C,D)** Expression patterns of *HcMYB1/2* and key structural genes (*HcTPS5* and *HcBSMT2*) in response to auxin treatment. Bars represent the means ± SD (n = 3–5). Different letters on bars indicate statistically significant differences ($P < 0.05$).

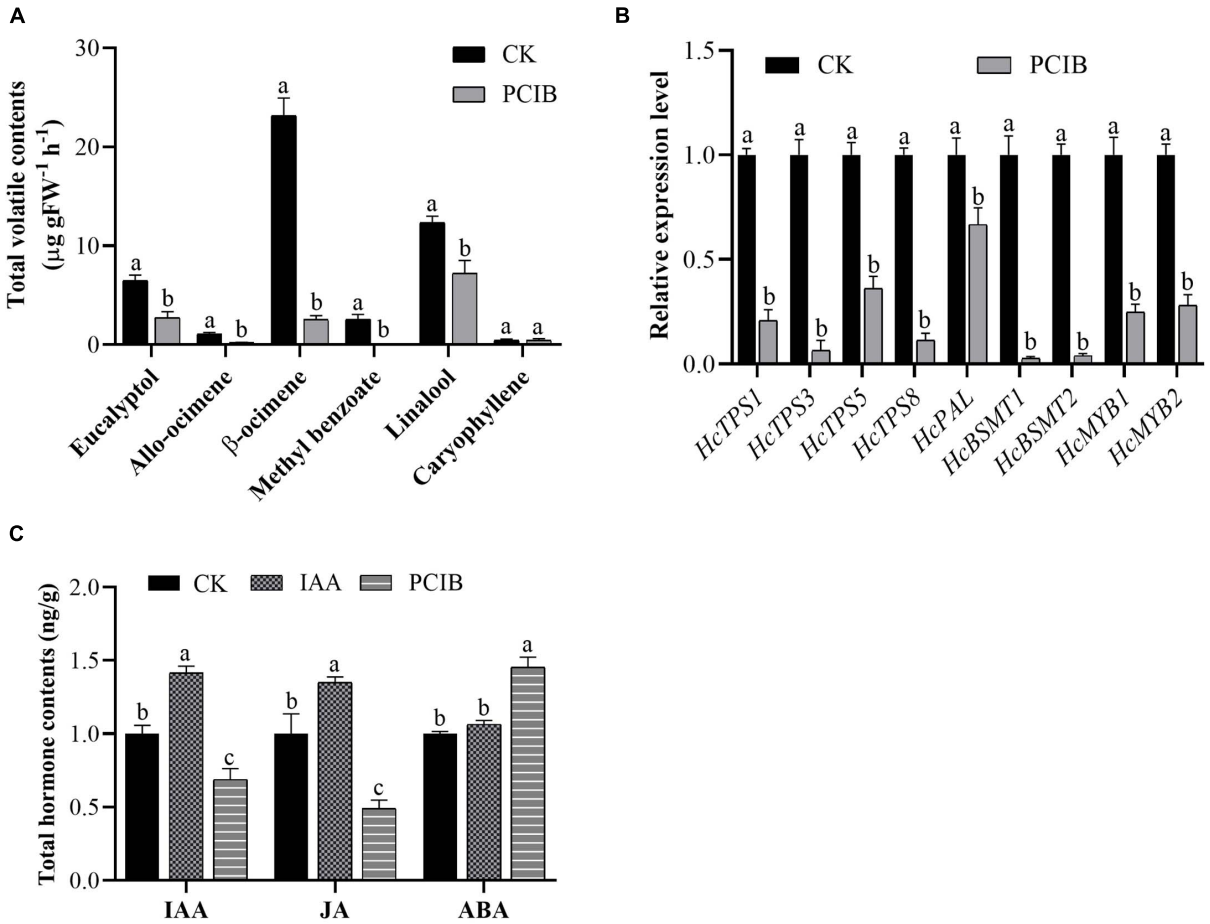


FIGURE 4 | Marked decrease in scent emission, decreased expression level of key structural volatile synthesis genes, and total IAA, JA, and ABA contents under PCIB treatment. **(A)** Headspace analyses of individual scent compounds under PCIB treatment using GC–MS. **(B)** Relative expression level of key structural genes under PCIB treatment. GAPDH was used as an endogenous control. Bars represent the means ± SD (n = 3–5). **(C)** Total pool contents of IAA, JA, and ABA in response to auxin and PCIB treatment. Each time point represents the average of three to four independent experiments, with SD indicated by vertical lines. Graphs represent the average of three independent experiments. Different letters on bars indicate statistically significant differences ($P < 0.05$).

HcTPS3, *HcTPS5*, *HcTPS8*, *HcPAL*, *HcBSMT1*, *HcBSMT2*, *HcMYB1*, and *HcMYB2* were downregulated by 79, 93, 64, 89, 33, 97, 96, 75, and 72%, respectively, relative to control (**Figure 4B**). The data endorse the aforementioned findings that auxin plays a crucial role in the biosynthesis of floral volatile compounds.

Furthermore, we quantify the total hormone contents under IAA and PCIB treatment using ultra-performance liquid chromatography-tandem mass spectrometer (UPLC-MS/MS). The data showed that under IAA treatment, the total IAA and jasmonic acid (JA) contents were increased by 42 and 35% compared to control. Likewise, the total IAA and JA contents were decreased by 31 and 51%, respectively, while abscisic acid (ABA) contents were increased by 45% under PCIB treatment (**Figure 4C**). However, the ABA contents do not change significantly under IAA treatment, suggesting that auxin might play a key role in the biosynthesis of floral volatile compounds via crosstalk with the abovementioned hormones.

HcMYB1 and HcMYB2 Localize to the Nucleus and Exhibit Transactivation Activity

Most MYB TFs specifically localize to the nucleus (Zou et al., 2008; Zhu et al., 2015; Zhou et al., 2017). However, some MYB TFs localize to both the nucleus and cytoplasm (Li et al., 2006a). The nuclear localization prediction server WoLF PSORT[1] predicted that HcMYB1 and HcMYB2 localize to the nucleus. To assess this prediction, we generated HcMYB1-GFP and HcMYB2-GFP constructs in which these genes were driven by the CaMV *35S* promoter and used them to transform *Arabidopsis* protoplasts. In HcMYB1-GFP- and HcMYB2-GFP-transformed protoplasts, observed green fluorescent protein (GFP) signals specifically in the nuclei, whereas control (GFP) protoplasts showed a ubiquitous distribution of GFP throughout the protoplasts (**Figure 5A**). We included nuclear localization

[1]https://wolfpsort.hgc.jp/

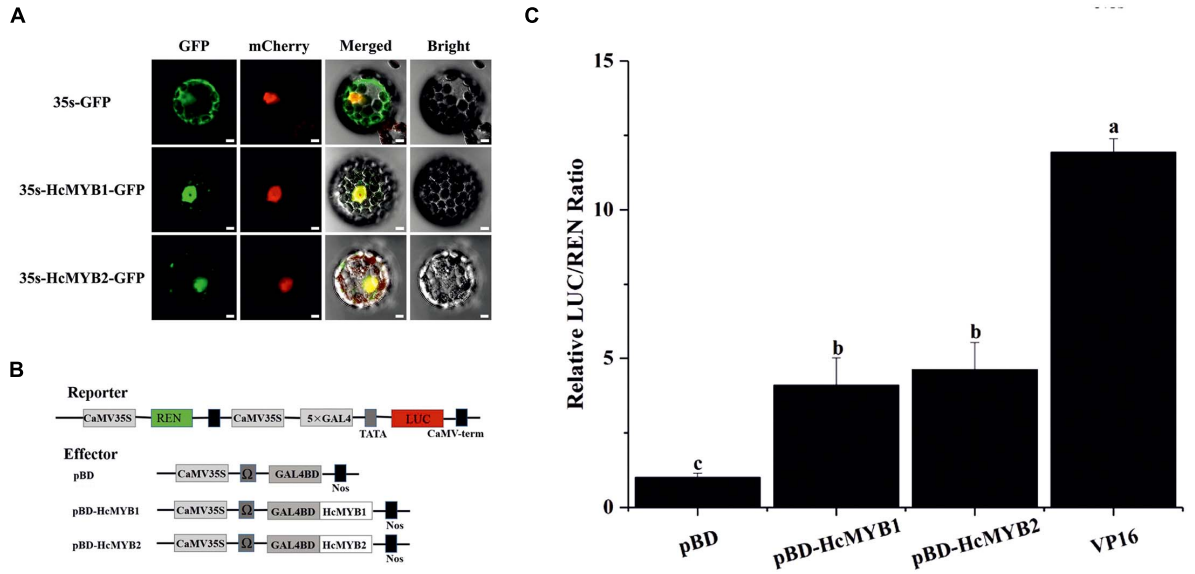


FIGURE 5 | HcMYB1 and HcMYB2 show nuclear localization in *Arabidopsis* protoplasts. **(A)** The images were obtained using a confocal laser scanning microscope and processed in adobe Photoshop. GFP: green fluorescent protein. mCherry: nuclear marker. Merged: combined GFP and mCherry signals. Bright: bright field. Bars: 10 μm. **(B)** Diagram of the reporter and effector plasmids. **(C)** Transcriptional activity of HcMYB1 and HcMYB2 together with DNA via GAL4 BD (pBD-HcMYB1 and pBD-HcMYB2) transiently co-expressed in *N. benthamiana* cells. The activity is presented as the ratio of LUC to REN luminescence. The VP16 and pBD represent the positive and negative control, respectively. The ratio of LUC/REN of the empty pBD vector is used as the calibrator (set as 1). Data are represented as means ± standard deviation of three replicates. Compared with the pBD, significant differences at the level of $P < 0.05$ analyzed by the student's *t*-test, are indicated by different letters above the bar.

signal (NLS)-mCherry in each transformation as a marker for nuclear localization. These results demonstrate that HcMYB1 and HcMYB2 are nucleus-localized proteins, which is in keeping with their expected roles as transcription factors.

Transcription factors regulate their target genes via transactivation activity. To investigate the transactivation activities of HcMYB1 and HcMYB2, we performed transient expression analysis in *Nicotiana benthamiana* leaves. We fused five copies of the GAL4 DNA-binding element (GAL4BD) and the minimal TATA region (5′-TATAAA-3′) of the *35S* promoter to the firefly luciferase (*LUC*) reporter; the Renilla luciferase (*REN*) reporter gene driven by the *35S* promoter as the reporter vector. The LUC/REN ratio from the reporter vector was used as an internal control. We constructed effector plasmids harboring the ORFs of *HcMYB1* and *HcMYB2* (**Figure 5B**). Unlike the GAL4BD negative control (empty vector, pBD), HcMYB1 and HcMYB2 activated the *LUC* reporter gene. The LUC/REN ratios of HcMYB1, HcMYB2, and GAL4BD-VP16 were 3. 1-, 3. 7-, and 11.1-fold higher, respectively, compared to the negative control (**Figure 5C**). These results indicate that HcMYB1 and HcMYB2 function as transcriptional activators.

Virus-Induced Gene Silencing of *HcMYB1* and *HcMYB2* in Flowers Modifies the Emission Levels of Volatiles

To investigate the potential involvement of *HcMYB1* and *HcMYB2* in floral scent formation, we suppressed their expression through virus-induced gene silencing (VIGS) in flowers (Renner et al., 2009; Yuan et al., 2011). We confirmed that this led to significant decreases in *HcMYB1* and *HcMYB2* transcript levels compared to those in unsilenced control flowers (**Figures 6A,B**). The contents of the volatiles methyl benzoate and linalool in flowers decreased by approximately 57 and 21%, respectively, in response to *HcMYB1* silencing, whereas the eucalyptol and ocimene contents did not change significantly (**Figure 6C**). In *HcMYB2*-silenced flowers, the contents of methyl benzoate, linalool, ocimene, and eucalyptol decreased by 68, 37, 18, and 17%, respectively, compared to the control (**Figure 6C**). We also analyzed the expression levels of key volatile biosynthesis genes (*HcTPS3*, *HcTPS5*, and *HcBSMT2*) in *H. coronarium*. In *HcMYB1*-silenced flowers, *HcTPS5* and *HcBSMT2* were significantly downregulated, whereas *HcTPS3* did not exhibit any significant changes in expression, compared to the control. Furthermore, in *HcMYB2*-silenced flowers, *HcTPS3*, *HcTPS5*, and *HcBSMT2* were all significantly downregulated compared to the control (**Figure 6D**). These results indicate that HcMYB1 and HcMYB2 play important and overlapping roles in the formation of floral volatiles in *H. coronarium*.

HcMYB1 and HcMYB2 Activate Structural Genes Involved in the Volatile Biosynthetic Pathway

MYB TFs transcriptionally regulate several genes by binding to the MEB [(T)(T)TGAC(C/T)] sequences in their promoters (Rushton et al., 2010). *In silico cis*-element analysis revealed

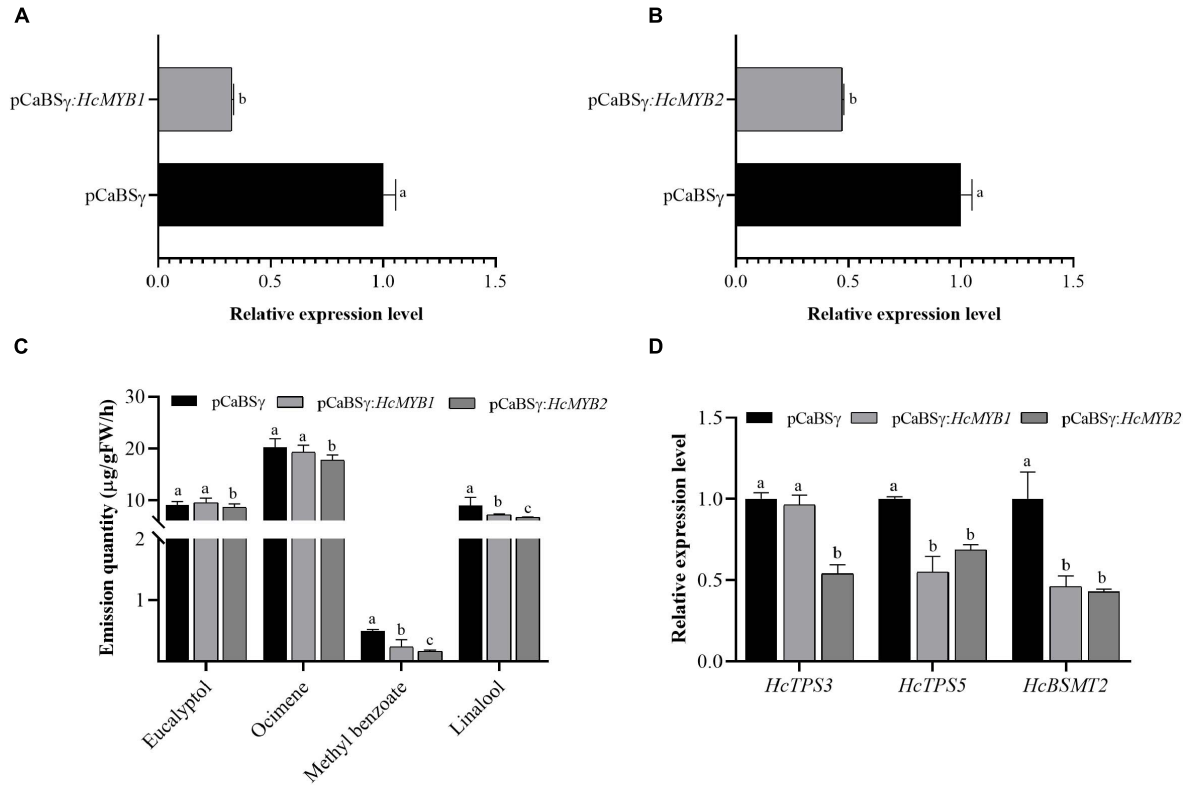


FIGURE 6 | *HcMYB1* and *HcMYB2* gene silencing alter *HcTPS3*, *HcTPS5*, and *HcBSMT2* expression and volatile emission in flowers. **(A,B)** Relative expression levels are presented as fold change values. **(C)** Emission levels of floral volatiles. **(D)** Expression levels of *HcTPS3*, *HcTPS5*, and *HcBSMT2*. Data represent the mean ± SD (n = 3–5). Different letters indicate significant differences ($P < 0.05$).

the presence of MYB-binding motifs in *HcBSMTs* and *HcTPSs* sequences. MYB-core binding motifs were present in ten out of twelve *HcBSMTs*. The number of MYB-binding motifs varies from one to thirteen. Interestingly, the number of MYB-bindings motifs in *HcBSMT2* was highest compared to other *HcBSMTs* (**Supplementary Table 1**). Similarly, MYB-core binding motifs were found in sixty out of sixty-two *HcTPSs*. The promoter sequence analysis of *HcBSMT2* (1131 bp) and *HcTPS5* (1555 bp) revealed the presence of MYB-core binding motifs in their sequences. There were 13 and 5 copies of MYB-binding motifs in the sequences of *HcBSMT2* and *HcTPS5*, respectively, suggesting that HcMYB1 and HcMYB2 might target these genes.

To determine whether HcMYB1 and/or HcMYB2 bind to the promoters of *HcBSMT2* and *HcTPS5*, we performed a yeast one-hybrid (Y1H) assay. Bait strains co-expressing HcMYB1 and HcMYB2 and harboring *proHcBSMT2* grew well in SD-Leu medium containing the antibiotic aureobasidin A (AbA), whereas bait strains harboring *proHcTPS5* grew well only when they expressed HcMYB2 (**Figures 7A,B**). These results indicate that HcMYB1 and HcMYB2 bind to the *HcBSMT2* promoter, while HcMYB2 binds to the *HcTPS5* promoter.

To confirm the binding ability of HcMYB1 and HcMYB2 to the *HcBSMT2* and *HcTPS5* promoters, we performed an electrophoretic mobility shift assay (EMSA) using GST-HcMYB1 and GST-HcMYB2 in *Escherichia coli*. The probes used for *proHcTPS5* and *proHcBSMT2* were 49 bp, which start from (+)1168 to (+)1216 and (+)291 to (+)339, respectively. The sequences of the probes are listed in **Supplementary Table 2**. Purified recombinant GST-HcMYB1 and GST-HcMYB2 fusion proteins bound to biotin-labeled probes derived from the *HcBSMT2* promoter, leading to a mobility shift, whereas no mobility shift occurred in the presence of GST alone (**Figure 7C**). Next, we performed a competition assay, which showed that adding a 100-fold amount of unlabeled probe molecules (as compared to the labeled molecules) to the binding reaction reduced the intensity of the protein-DNA complex signal, and adding 500-fold unlabeled probes prevented any protein-DNA complex from being detected (**Figure 7C**). We also observed binding between HcMYB2 and a biotin-labeled probe from the *HcTPS5* promoter (**Figure 7D**). These results endorse the aforementioned data that HcMYB1 binds to the *HcBSMT2* promoter and HcMYB2 binds to the *HcBSMT2* and *HcTPS5* promoters.

To test the ability of HcMYB1 and HcMYB2 to activate the *HcBSMT2* and *HcTPS5* promoters, we performed a dual-luciferase assay. We individually cloned the promoter regions of *HcBSMT2* and *HcTPS5* into reporter plasmids and the ORFs of *HcMYB1* and *HcMYB2* into effector plasmids (**Figure 7E**). HcMYB1 and HcMYB2 significantly enhanced *HcBSMT2*

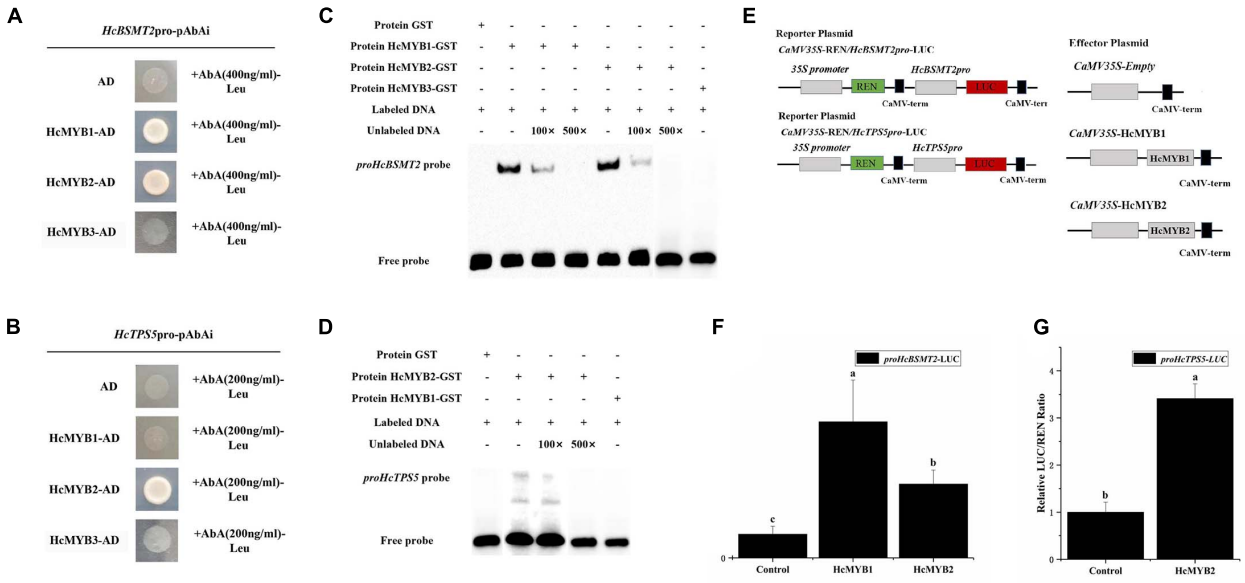


FIGURE 7 | HcMYB1 and HcMYB2 interact with the *HcBSMT2* and *HcTPS5* promoters. **(A)** Interactions of HcMYB1 and HcMYB2 with the *HcBSMT2* promoter. **(B)** Interactions of HcMYB1 and HcMYB2 with the *HcTPS5* promoter. HcMYB3 and empty vector were used as a control. **(C)** Binding activity of HcMYB1 and HcMYB2 with the *proHcBSMT2* probe. **(D)** Binding activity of HcMYB2 with the *proHcTPS5* probe. **(E)** Diagram of the reporter and effector plasmids. **(F)** HcMYB1 and HcMYB2 transactivate the *HcBSMT2* promoter. **(G)** HcMYB2 transactivates the *HcTPS5* promoter. The promoter activity is presented as the ratio of LUC to REN luminescence. Different letters on bars indicate statistically significant differences ($P < 0.05$).

promoter activity (by 4.8-fold and 2.2-fold, respectively) compared to the control (**Figure 7D**). Meanwhile, HcMYB2 significantly enhanced *HcTPS5* promoter activity (by 2.4-fold) compared to the control (**Figure 7F**). Therefore, HcMYB1 and HcMYB2 activate the *HcBSMT2* promoter and HcMYB2 activates the *HcTPS5* promoter, indicating that these TFs have different target genes in *N. benthamiana* leaves. These findings indicate that both HcMYB1 and HcMYB2 are transcriptional activators of volatile biosynthesis genes in flowers.

HcMYB1 Interacts With the Auxin-Responsive Protein HcIAA4 by Y2H and BiFC Assays

MYB proteins interact with many other proteins involved in hormone signal transduction, such as the jasmonic acid (JA)-responsive repressor proteins of the JASMONATE ZIM-DOMAIN (JAZ) family and the ABA signal receptor protein PYRABACTIN RESISTANCE LIKE (PYL) (Qi et al., 2014; Zhao et al., 2014). In a Y2H assay, HcMYB1 interacted with the auxin-responsive protein HcIAA4, whereas HcMYB2 did not (**Figure 8A**). To verify the interaction between HcMYB1 and HcIAA4, we performed a BiFC assay. Expressing the N-terminal half of YFP fused to HcMYB1 (HcMYB1-YFP^N) and the C-terminal half of YFP fused to HcIAA4 (HcIAA4-YFP^C) in *N. benthamiana* leaves resulted in fluorescence. Moreover, the reciprocal experiment with the C-terminal half of YFP fused to HcMYB1 (HcMYB1-YFP^C) and the N-terminal half of YFP fused to HcIAA4 (HcIAA4-YFP^N) also resulted in fluorescence and the control combinations of YFP^C + HcMYB1-YFP^N and HcIAA4-YFP^C + YFP^N did not result in fluorescence (**Figure 8B**). To elucidate the functional significance of the interaction between HcMYB1 and HcIAA4, we co-transformed *N. benthamiana* leaves with the same amounts of effectors carrying HcMYB1 and/or HcIAA4 in combination with the *HcBSMT2pro*-LUC reporter constructs. The effect of HcMYB1 on *HcBSMT2pro* expression was repressed in the presence of HcIAA4 (**Figure 8C**). These results demonstrate that HcMYB1 directly activates *HcBSMT2* expression, which is modulated by its interacting partner HcIAA4.

DISCUSSION

The transcriptional regulatory network governing floral scent emission has not been thoroughly elucidated. To date, only a few TFs that regulate the expression of scent-related genes have been identified (Katiyar et al., 2012; Abbas et al., 2021b). R2R3-MYB TFs are key regulators of the phenylpropanoid and terpenoids biosynthetic pathway in plants (Du et al., 2009; Zhu et al., 2015; Yang et al., 2020). MYB TFs in the same subgroup have similar functions (Zhu et al., 2015). Here, we used previously characterized R2R3-MYB proteins involved in secondary metabolism to construct a phylogenetic tree with HcMYB1 and HcMYB2 from *H. coronarium* (**Figure 1B**). HcMYB1 was clustered in Group III with AtMYB77, which modulates auxin signal transduction (Shin et al., 2007), and AtMYB44, a stress-responsive protein involved in senescence and ABA signaling (Jaradat et al., 2013), suggesting that HcMYB1 might play a key role in hormone signaling. Meanwhile, HcMYB2

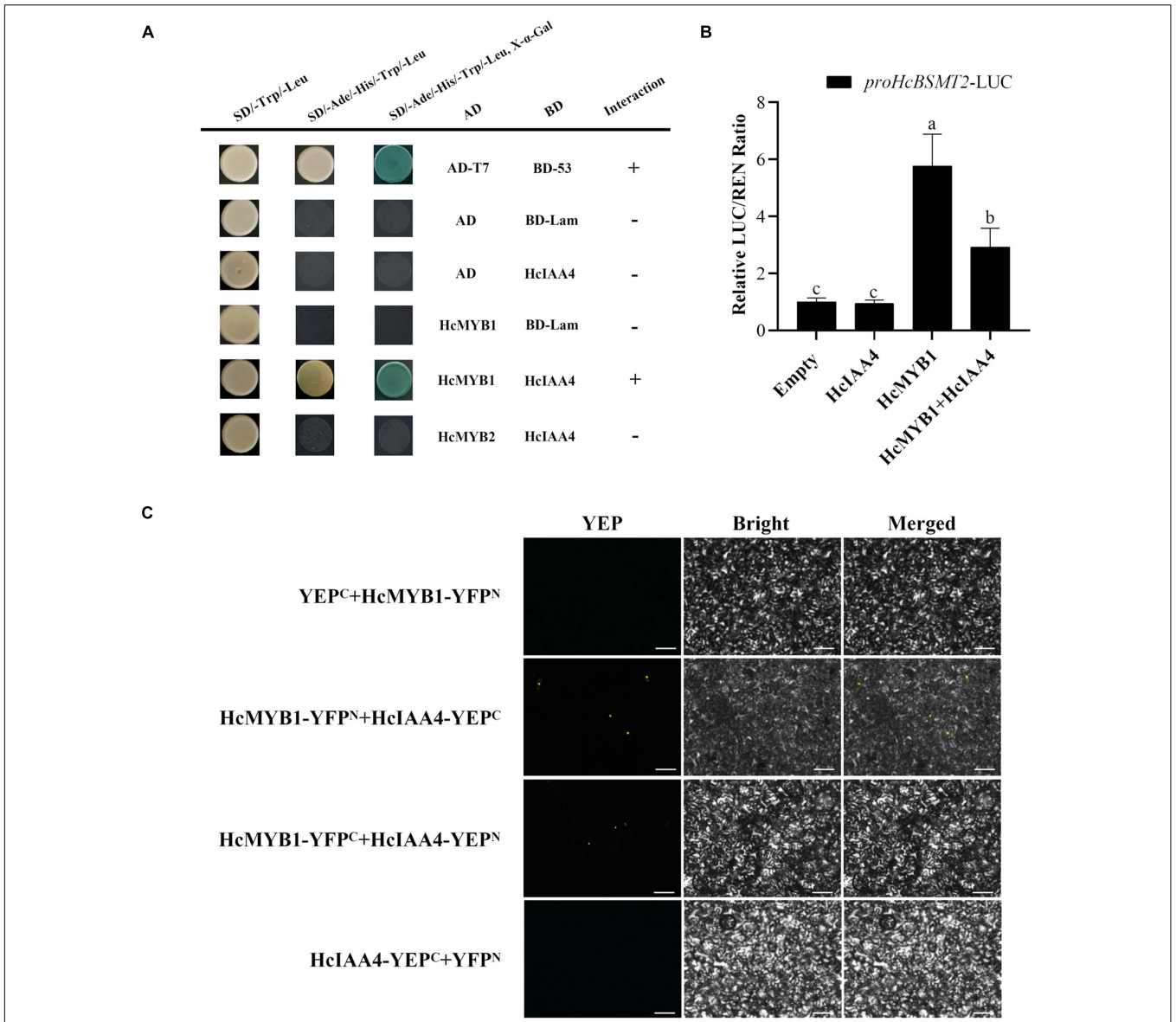


FIGURE 8 | HcMYB1 interacts with HcIAA4. **(A)** Analysis of the interactions of HcIAA4, HcMYB1, and HcMYB2 via yeast two-hybrid assay. Yeast cells were transformed with DBD–HcARF4 + AD–HcMYB1, DBD–HcARF4 + AD–HcMYB2, DBD + AD–HcMYB1, DBD–HcIAA4 + pGADT7-T, pGBKT7-53 + pGADT7-T, DBD–HcIAA4 + pGADT7-T or pGBKT7-Lamin + pGADT7-T. The vectors were transformed into yeast strain Y2HGold and then transformants were screened by plating on SD/-Trp/-His/-Ade + X-α-gal medium. BD-Lam was used as a negative control. ADT7-BD53 was used as a positive control. Three independent replicates were performed. **(B)** BiFC assay between HcMYB1 and HcIAA4 in *N. benthamiana* leaves. Fusion constructs YFP^C + HcMYB1-YFP^N, HcMYB1-YFP^N + HcIAA4-YFP^C, HcMYB1-YFP^C + HcIAA4-YFP^N, and HcIAA4-YFP^C + YFP^N were co-infiltrated into *N. benthamiana* leaves and observed for fluorescence complementation. Bars, 5 μm. **(C)** The activation ability of HcMYB1 is modulated by HcIAA4. Different letters on bars indicate statistically significant differences ($P < 0.05$).

was classified into Group I along with AmMYB305, AmMYB340, AtMYB24, AtMYB21, PhEOBII, and FaEOBII (**Figure 1**). These R2R3-MYB TFs regulate the metabolic pathway of the volatile compound phenylpropanoid (Uimari and Strommer, 1997; Shin et al., 2002; Li et al., 2006b; Liu et al., 2009; Spitzer-Rimon et al., 2010; Medina-Puche et al., 2015), suggesting that HcMYB2 might be involved in regulating the floral volatile metabolic pathway in *H. coronarium*.

HcMYB1 and *HcMYB2* Are Expressed During Specific Stages of Floral Development and Are Correlated With Volatile Production

The production and emission of fragrance compounds by flowers are strictly regulated during the floral lifespan and often peak when the flower is in full bloom and pollinators

are active (Dudareva et al., 2013; Muhlemann et al., 2014; Abbas et al., 2017). Methyl benzoate and linalool are the main phenylpropanoids and terpenoid volatiles that contribute to flower scent in *H. coronarium* (Fan et al., 2003, 2007; Báez et al., 2011; Yue et al., 2015; Abbas et al., 2021a). We observed that the expression of *HcMYB1* changed during flower development, with the highest expression level detected at the full bloom stage (F5) (**Figure 2B**. A similar expression pattern was detected in lilac (*Syringa oblata*) during different stages of flower development, as two R2R3-MYB TF genes were upregulated at the full-bloom stage compared to the bud stage (Zhu et al., 2015). Similarly, the expression level of *FhMYB5* gradually increased during the flower developmental stages, resembling anthocyanin biosynthesis pattern and function in the flavonoid pathway in *Freesia hybrida* (*F. hybrida*) (Li et al., 2019). In *Rosa hybrida*, mRNA levels of a putative scent-related gene (*RhMYB1*) were developmentally regulated peaking at full bloom stage similar with other rose scent-related genes, such as phenylacetaldehyde synthase *RhPAAS*, the sesquiterpene synthase *RhGDS*, the alcohol acetyltransferase *RhAAT* and the orcinol O-methyltransferases (*OOMT*) (Lavid et al., 2002; Yan et al., 2011).

Interestingly, the expression patterns of *HcMYB1* and *HcMYB2* during development were similar to those of *HcBSMT2* and *HcTPS5*, which are responsible for the formation of the volatiles methyl benzoate and linalool, respectively (**Figure 2D**). Both compounds reach their highest levels in flowers at the full bloom stage (Yue et al., 2015). Similarly, *HcBSMT2* and *HcTPS5* expression levels were highest in flowers during the periods when the largest amounts of volatiles were released (**Supplementary Figure 5**). Thus, the expression of TF genes (*HcMYB1* and *HcMYB2*) and structural genes (*HcBSMT2* and *HcTPS5*) was associated with flower development and the production of high levels of volatiles (**Supplementary Figure 1**). In *F. hybrida* and *A. thaliana*, FhMYB21L1 and FhMYB21L2, TF genes were synchronously expressed with *FhTPS1* and could activate its expression significantly (Yang et al., 2020). These results suggest that HcMYB1 and HcMYB2 regulate volatile production during flower development, which is similar to the roles of R2R3-MYBs PhEOBII and FaEOBII in petunia and strawberry, respectively (Van Moerkercke et al., 2011; Medina-Puche et al., 2015).

HcMYB1 and *HcMYB2* Expression Is Regulated by Auxin Like That of Other Scent-Related Genes in *H. coronarium*

The volatile biosynthesis pathway, and particularly the emission of methyl benzoate and linalool, is induced by IAA, suggesting that IAA regulates the expression of transcription factors or key enzymes involved in this pathway at the protein or transcript level. In the current study, we demonstrated that auxin induces the expression of both TF genes (*HcMYB1* and *HcMYB2*) and key biosynthesis genes [phenylalanine ammonia lyase (*PAL*), *BSMT*, and *TPS*] (**Figures 3B,C**). IAA treatment also upregulated *HcBSMT2* and *HcTPS5* expression, especially at 12 h (**Figure 3D**). In *A. thaliana*, expression of *AtTPS21* and *AtTPS11* was induced by the phytohormones, and both inductions require AtMYC2 (Hong et al., 2012). R2R3-MYBs such as *AtMYB77* and *AtMYB44* are involved in the response to auxin signaling in *Arabidopsis* (Shin et al., 2007; Yamaguchi et al., 2013). Similarly, *FaMYB10* and *FaEOBII* are regulated by auxin in strawberry (Perkins-Veazie, 1995; Chai et al., 2011; Jia et al., 2011). On the other hand, auxin contents decrease and *FaMYB10* and *FaEOBII* expression increases during fruit development in strawberry (Medina-Puche et al., 2014, 2015). During flower development in *H. coronarium*, auxin contents increased and *HcMYB1* and *HcMYB2* expression increased (**Figures 2C,E**). The differences in auxin response patterns between the *HcMYB* and *FaMYB* genes may be due to the evolutionary distance between *H. coronarium* (Zingiberaceae) and *F.* × *ananassa* (Rosacea). To validate the function of auxin, the flowers were treated with PCIB which is widely used to inhibit auxin action (Oono et al., 2003). The data showed that in contrast to auxin, the emission of main floral volatiles and expression level of aforementioned key structural volatile synthesis genes significantly downregulated (**Figures 4A,B**). The following data endorse the abovementioned findings that auxin plays an essential role in floral scents. Furthermore, relative to the control, total IAA and JA contents significantly upregulated and downregulated under auxin and PCIB treatment, respectively (**Figure 4C**). Several studies showed that MYB TF respond to various phytohormones. Under JA treatment, the transcript abundance of *Pinus taeda PtMYB14* and *PtMYB13* rapidly increased by 14-fold and 2-fold, respectively, while *Picea glauca PgMYB14* and *PgMYB15* transcripts increased 4-fold and 2-fold. Furthermore, the characterization of the aforementioned TF genes reveals *PtMYB14* as a putative regulator of an isoprenoid and flavonoid-oriented response in conifers (Bedon et al., 2010). In Apples, MdMYB9 and MdMYB11 were involved in the regulation of the JA-induced biosynthesis of anthocyanin and proanthocyanidin (An et al., 2014). The regulatory patterns of *MYBs* are dependent on developmental stage, tissue type, and environmental conditions. Much remains to be learned about the mechanistic basis of the responses of MYB TFs to auxin signaling molecules during volatile formation.

HcMYB1 and HcMYB2 Activate Key Structural Genes Involved in Volatile Biosynthesis in *H. coronarium* Flowers

The structural genes *HcBSMT2* and *HcTPS5* are essential for the formation of methyl benzoate and linalool, respectively, in *H. coronarium* flowers (Yue et al., 2015). Notably, we detected MYB-binding elements in the promoters of *HcBSMT2* and *HcTPS5* (**Supplementary Table 1**). This result is supported by the finding that HcMYB1 transactivates the *HcBSMT2* promoter and that HcMYB2 transactivates the *HcBSMT2* and *HcTPS5* promoters (**Figure 7**). In certain plants, floral scent biosynthesis is dependent on transcriptional regulation, and TFs control volatile emissions (Colquhoun et al., 2011; Muhlemann et al., 2012). ODO1 was the first R2R3-type MYB transcription factor shown to regulate the benzenoid biosynthesis pathway in petunia, followed by the R2R3-MYB TFs EOBI and EOBII.

ODO1 strongly influences the floral scent pathway by regulating the transcript levels of many key genes (*PAL*, *CM*, *DAHPS*, *SAMS*, and *EPSPS*). Meanwhile, *ODO1* is directly regulated by EOBII. Moreover, EOBI directly binds to and activates the promoters of *ODO1*, *IGS*, and *PAL* to regulate scent production (Verdonk et al., 2005; Spitzer-Rimon et al., 2010, 2012). In *F. hybrida*, FhMYB5 and FhbHLH mainly contribute to the regulation of anthocyanin and proanthocyanidin via activating the expression of biosynthetic genes (*FhCHS*, *FhCHI*, *FhF3H*, *FhF3′H*, *FhF3′5′H*, and *FhDFR*) involved in the flavonoid pathway (Li et al., 2019). In spearmint, MsMYB negatively regulates monoterpene production and suppresses the expression of geranyl diphosphate synthase (Reddy et al., 2017). Likewise, several R2R3-MYB transcription factor have been identified which are potentially involved in the regulation of flavonoid biosynthesis via controlling the expression of structural genes (Cao et al., 2021; Zhang et al., 2021). Similarly, *Lilium* hybrid *ODO1* (*LhODO1*) regulates fragrance biosynthesis via regulating the expression of structural genes involved in the shikimate and benzenoid/phenylpropanoid pathway (Yoshida et al., 2018).

In *H. coronarium*, HcMYB1 and HcMYB2 directly activate the methyl benzoate biosynthesis gene *HcBSMT2*, whereas HcMYB2 activates the linalool biosynthesis gene *HcTPS5*. Therefore, HcMYB2 activates two different groups of volatile biosynthesis genes, suggesting it plays a dual role in controlling both the phenylpropanoid and terpenoid pathways. Several R2R3-MYB TFs regulate the biosynthesis of one or more units of phenylpropanoid-derived compounds, such as MdMYB3, AtMYB4, and AtMYB12 (Aharoni et al., 2001; Deluc et al., 2006, 2008; Luo et al., 2008; Rommens et al., 2008; Liu J. et al., 2015). Here, we examined the effects of HcMYB1 and HcMYB2 on floral scent via gene silencing (**Figure 6**). Linalool and methyl benzoate levels significantly decreased in flowers when *HcMYB1* or *HcMYB2* was silenced, confirming the direct connection between the functional activity of these two TFs and volatile biosynthesis. The silencing of *HcMYB1* or *HcMYB2* also led to the downregulation of key structural scent-related genes (*HcTPS3*, *HcTPS5*, and *HcBSMT2*) from the terpenoid and phenylpropanoid pathways (**Figure 6D**). Similar results were obtained in petunia, where the silencing of R2R3-MYB (*ODO1*) led to the downregulation of several scent-related genes (Spitzer-Rimon et al., 2012). In addition, overexpressing *PAP1* from *Arabidopsis* modulated the accumulation of terpenoid and phenylpropanoid scent compounds in rose flowers (Zvi et al., 2012). Nevertheless, before this study, little was known about the transcriptional regulatory mechanism underlying scent compound biosynthesis in non-model fragrance plants such as *H. coronarium*.

HcIAA4 Interacts With and Modulates the Transcriptional Activity of HcMYB1

MYB TFs form complexes by interacting with other proteins, such as MYB-helix-loop-helix (bHLH)-WD40 proteins involved in regulating anthocyanin biosynthesis (Zhou and Memelink, 2016). MYB TFs also interact with other proteins involved in hormone signaling pathways, such as JAZ and PYL, which are crucial components of the JA and ABA signal-transduction pathways, respectively (Qi et al., 2014; Zhao et al., 2014). However, little is known about the interactions of MYB TFs with proteins in the auxin-signaling pathway. In *Arabidopsis*, the auxin signaling pathway repressor Aux/IAA (AtIAA29) interacts with the TF WRKY57 to mediate leaf senescence (Jiang et al., 2014). It was also observed that both HcMYB1 and HcIAA4 showed high protein expression in flowers (**Supplementary Figure 4**). In the current study, we uncovered an interaction between HcMYB1 and the auxin-responsive protein HcIAA4 via Y2H and BiFC assays (**Figures 8A,B**). We also demonstrated that HcIAA4 represses the activity of HcMYB1 (**Figure 7C**). Similarly, in *Arabidopsis*, AtJAZ proteins interact with MYBs such as MYB75, thereby decreasing their transcriptional activity (Qi et al., 2014). In *A. thaliana* and *F. hybrida*, MYB21 interacts with MYC2 to form MYB–bHLH complex to regulate the expression of TPS genes and floral scent emission in flowers (Yang et al., 2020). In *Fagopyrum tataricum*, the repressive activities of FtMYBs are directly enhanced by their interactions with FtSAD2 or FtJAZ1 (Zhang et al., 2018). The identification of protein–protein interactions between MYB TFs and other proteins provides clues about the regulation of gene expression and secondary metabolism during volatile biosynthesis.

In petunia, a network comprising three R2R3 MYB TFs (EOBI, EOBII, and ODO1) regulates flower-specific genes in the phenylpropanoid volatile biosynthesis pathway (Verdonk et al., 2005; Spitzer-Rimon et al., 2010, 2012; Van Moerkercke et al., 2011). Similarly, FaMYB10 regulates *FaEOBII* expression in strawberry (Medina-Puche et al., 2014, 2015). In *A. thaliana*, MYC2 interacts with DELLA protein to regulate the expression of sesquiterpene synthase genes (*TPS21* and *TPS11*), and the expression of *TPS21* and *TPS11* was modulated by phytohormones (Hong et al., 2012). Likewise, LcMYB1 interacts with LcbHLH to regulate the expression of key structural anthocyanin biosynthesis genes in *Litchi chinensis* (Lai et al., 2016). The roles of MYB proteins in scented ornamental plants uncovered in the current study sheds light on the evolution of this important transcription factor family, providing new insights into how they regulate the biosynthesis of secondary metabolic compounds, including terpenoids and phenylpropanoids, in plants. The key role of MYB TFs in controlling the biosynthesis of volatile compounds highlights the potential of engineering these TFs to enhance the economic value of ornamental plant species.

In conclusion, we demonstrated that the IAA-responsive, flower-specific R2R3-MYB TFs HcMYB1 and HcMYB2 function as activators of terpenoid and phenylpropanoid biosynthesis in *H. coronarium*. Both HcMYB1 and HcMYB2 interact with the promoter of *HcBSMT2*, encoding the key enzyme for methyl benzoate biosynthesis. Furthermore, HcMYB2 regulates the expression of *HcTPS5*, which plays a key role in linalool biosynthesis. Finally, we showed that auxin takes part in volatile biosynthesis by regulating the expression of R2R3-HcMYB transcription factor genes via protein–protein interactions in *H. coronarium*. Our findings provide important insights into

the roles of auxin and MYB TFs in the biosynthesis of floral scent compounds, laying the foundation for plant metabolic engineering efforts.

MATERIALS AND METHODS

Plant Material and Hormone Treatment

Hedychium coronarium plants were grown at South China Agricultural University under natural light conditions. For RNA extraction, the plant materials were harvested, immediately frozen in liquid nitrogen, and stored at −80°C. To analyze tissue-specific gene expression patterns, three tissues were used: fully open flowers, mature green leaves, and healthy rhizomes of 2-year-old *H. coronarium* plants. The process of flower development was divided into six stages: tight green bud stage (F1), white bud stage (F2), initial flowering stage (F3), half-open stage (F4), full-bloom stage (F5), and flower senescence (F6).

For IAA treatment, flowers at the F2 stage were cut into 35-cm pieces, placed in sterile water containing 100 μM IAA, and incubated for 12 h in a growth chamber under a 14/10 h light/dark cycle at 25°C. IAA and PCIB stock solution was prepared as per the manufacturer's protocol. Briefly, 18.79 mg IAA powder was dissolved in 1.5 mL methanol and diluted in 100 mL sterilized water. Likewise, 321 mg PCIB powder was dissolved and diluted as aforementioned conditions. Thereafter, the detached flowers were placed in a flask containing 100 mL solution and covered with a silver sheet to protect them from degradation. The control flowers were kept in with a similar amount of volume of sterile water without IAA under the same aforementioned conditions. Volatile content was analyzed in flowers at the full-bloom stage. After analysis, all samples were frozen in liquid nitrogen and stored at −80°C for further experiments. Three to five independent experiments was performed with each experimental variant. *A. thaliana* and *N. benthamiana* plants used for subcellular localization and BiFC assays were grown in a growth chamber at 24°C under a 12/12 light/dark cycle.

Sequence Alignment and Phylogenic Analysis

The sequences of *HcMYB1* and *HcMYB2* were obtained from a flower RNA-seq database for *H. coronarium* (SRP049915). The related protein sequences were retrieved from the NCBI database[2]. The protein sequences were aligned, and a phylogenetic tree based on the R2R3-MYB domain was constructed using Clustal Ω (Sievers et al., 2011) and MEGA X (Kumar et al., 2018) software.

RNA Isolation, cDNA Synthesis, and qRT-PCR

Total RNA was isolated from flowers at different stages of development and different organs using a HiPure Plant RNA Mini Kit (Magen) following the manufacturer's protocol. Each 1 μg RNA sample was reverse transcribed using a PrimeScript RT Reagent Kit with Genomic DNA Eraser (Takara) following the manufacturer's suggestions. To generate primers for reverse-transcription PCR (RT-PCR), the specific sequences of genes in *H. coronarium* were selected, and primers for the genes were designed with Primer 5.0 software (Abbas et al., 2020). The reaction mixtures (20 μl) included 10 μL pink SYBR mix, 7.2 μl distilled water, 0.4 μl each primer (10 μM), and 2.0 μl template cDNA. The expression values were calculated using the $2^{-\Delta\Delta Ct}$ method (Livak and Schmittgen, 2001). A similar procedure was performed for all treatments using specific primers listed in **Supplementary Table 2**.

Subcellular Localization Assay

The full-length fragments excluding the stop codon were fused with the *GFP* gene in the p35S-EGFP-1 vector using *Sma*I at the 5′ end and *Spe*I at the 3′ end. The isolation and transformation of *Arabidopsis* protoplasts were performed as described by Yoo et al. (2007). The protoplasts were observed and photographed at 18 h after transformation under a confocal laser-scanning microscope.

Virus-Induced Gene Silencing

Barley stripe mosaic virus (BSMV) was used for VIGS, as this system has successfully been used in monocots (Renner et al., 2009; Yuan et al., 2011). The pCaBSγ vector was linearized with *Apa*I before inserting the fragments. To specifically silence *HcMYB1* and *HcMYB2*, a 280-bp fragment of each gene (from the 3′ end) was amplified by PCR from *H. coronarium* cDNA and inserted into pCaBSγ to produce pCaBSγ:*HcMYB1* and pCaBSγ:*HcMYB2*, respectively. The cultures were harvested by centrifugation at 5000 rpm for 10 min and resuspended in infiltration buffer (10 mM $MgCl_2$, 0.1 mM acetosyringone, 10 mM MES, pH 5.6) to a final OD_{600} of ∼1. For VIGS, flowers at the F1 stage were collected, dipped in the bacterial suspension, and vacuum infiltrated at 0.8 MPa. After the vacuum was released, the flowers were washed in deionized water, placed into liquid MS medium, and cultured under a 12/12 h light/dark cycle at 16°C for 5 days. Total floral volatile compounds were analyzed at the full-bloom stage via gas chromatography-mass spectrometer (GC-MS); the experiment was replicated three to four times.

Yeast One-Hybrid Assay

The yeast-one-hybrid assay was performed using the Gold Yeast One Hybrid System (Clontech, Takara). To generate bait-specific reporter strains, a 1131-bp fragment (−1 to −1131 bp upstream of ATG) of the *HcBSMT2* promoter and a 1555-bp fragment (−1 to −1555 bp upstream of ATG) of the *HcTPS5* promoter were inserted into pAbAi to generate *HcBSMT2*-pAbAi and *HcTPS5*-pAbAi, respectively. The plasmids were integrated into the genome of yeast strain Y1H Gold (Clontech, Takara) via homologous recombination, and transformed colonies were selected on uracil-deficient synthetic dropout (SD/-Ura) medium. Different concentrations of the antibiotic AbA were used to select the bait strains. To generate the prey constructs, full-length *HcMYB1* and *HcMYB2* were cloned into pGAL4.

[2]http://www.ncbi.nlm.nih.gov/

The prey constructs were transformed into yeast cells harboring the bait constructs. The *in vivo* DNA-binding activity was determined based on the growth status of the transformed yeast cells on leucine-deficient synthetic dropout (SD/-Leu) medium supplemented with the selected concentration of AbA after 3–5 days of cultivation at 30°C.

Yeast Two-Hybrid Assay

The full-length *HcIAA2* and *HcIAA4* sequences were ligated into prey vector pGADT7 (AD). The coding sequences of *HcIAA4*, *HcMYB1*, and *HcMYB2* were cloned into bait vector pGBKT7 (BD). The prey construct and specific bait construct were co-transformed into yeast strain Y2HGold, which harbors four reporter genes (*Ade*, *MEL1*, *His*, and *AUR1*) under the control of a GAL4-responsive promoter. The empty pGADT7 vector was used as a blank control. The positive transformants that grew on SD/-Trp medium were inoculated onto SD plates (SD/-Leu-Trp-His-Ade) and incubated for 3–5 days at 30°C; transactivation activity was confirmed by growth on these plates. Yeast colonies expressing the α-galactosidase MEL1 turned blue upon the addition of X-α-Gal substrate (Clontech, TaKaRa). Primers used to amplify the genes are listed in **Supplementary Table 2**.

Dual-Luciferase Transient Expression Assay

To assay the transcriptional activities of *HcMYB1* and *HcMYB2*, the full-length coding regions were independently fused to the GAL4 DNA-binding domain ligated to the pBD vector and used as the effectors. The reporter vector was modified from the pGreenII 0800-LUC vector, which includes the firefly luciferase gene (*LUC*) driven by the CaMV *35S* minimal promoter with five repeats of upstream activating sequence, as well as the *35S*-promoter-driven *Renilla reniformis* luciferase gene (*REN*) as an internal control. To examine the binding of HcMYB1 and HcMYB2 to the promoters of *HcBSMT2* and *HcTPS5*, the promoters were cloned into the pGreenII 0800-LUC double-reporter vector (Hellens et al., 2005), whereas *HcMYB1* and *HcMYB2* were cloned into the pGreenII 62-SK vector as effectors. The effector and reporter plasmids were electroporated into *Agrobacterium tumefaciens* strain EHA105 and injected into *N. benthamiana* leaves with a needleless syringe. After 3–5 days, leaves were collected and LUC and REN activities examined via a dual-luciferase assay (Promega, United States) using a Luminoskan Ascent Microplate Luminometer (Thermo Fisher, United States) following the manufacturer's protocol. The transactivation ability and the binding activity of HcMYB1 and HcMYB2 are indicated by the ratio of LUC to REN. Four or five measurements were carried out for each combination and three independent experiments were performed.

Electrophoretic Mobility Shift Assay

For EMSA, pGEX-4T-1 (GE Healthcare) was used to generate the GST-HcMYB1 and GST-HcMYB2 expression vectors, which were transformed into *E. coli* strain *BM* Rosetta (DE3). The expression of the recombinant fusion proteins was induced by adding 0.5 mM isopropyl-β-D-thiogalactopyranoside (IPTG). Following incubation at 28°C for 8 h, the fusion proteins were purified using Glutathione Superflow Resin (Clontech) according to the manufacturer's instructions. The fragments (∼50 bp) containing putative MBE-binding sequences in the *HcTPS5* and *HcBSMT2* promoters were labeled with biotin. EMSA was carried out using a Light Shift Chemiluminescent EMSA Kit (Thermo Scientific) as previously described (Tan et al., 2019). The purified fusion protein was incubated with biotin-labeled DNA fragments and a 100-fold molar excess of unlabeled DNA fragments with the same sequences that were used as competitors; GST protein with labeled DNA was used as a negative control. The protein-DNA complexes were separated by 5% native polyacrylamide gel electrophoresis, detected based on chemiluminescence on a ChemiDoc MP Imaging System (Bio-Rad), and transferred onto a nylon membrane.

BiFC Assays

The full-length *HcMYB1* and *HcIAA4* sequences were separately inserted into PUC-SPYNE and PUC-SPYCE to form HcMYB1-YFP^N and HcIAA4-YFP^C, respectively. The empty vector (control) and recombinant plasmids were transformed into EHA105 competent cells, and different combinations of *Agrobacterium* cultures were co-infiltrated into *N. benthamiana* leaves. The plants were cultivated in an incubator under a 16 h/8 h light/dark cycle for 3 days, and the infiltrated leaves were visualized under a Leica DM RXA2 upright fluorescence microscope as described previously (Ke et al., 2019).

Measuring IAA Contents in *H. corium* Flowers

The flower samples were ground to a fine powder in liquid nitrogen (N_2) and transferred to 15 mL tubes containing 5 mL extraction solvent (2:1:0.002 [v/v/v] 2-propanol:H_2O:HCl). The samples were sonicated for 15 min and incubated at 4°C for 30 min with shaking (100 rpm). After adding 5 mL dichloromethane, the samples were incubated under the same conditions and centrifuged at 10,000 rpm for 10 min at 4°C. The samples were concentrated in the dark via aeration of the solvent mixture using nitrogen gas, followed by the addition of 1.0 mL methanol and purification through a Sep-Pak^{TM} C_{18} reverse-phase extraction cartridge. The samples were dried completely, dissolved in 200 μL methanol, and filtered through a 0.22-mm PTFE filter. The IAA standards were prepared by dissolving IAA standards in methanol (Sigma, United States). Chromatographic and mass spectrometric conditions were as described previously (Niu et al., 2014). The experiment was performed in three biological and three technical replicates.

Ultraperformance Liquid Chromatography-Tandem Mass Spectrometry (UPLC-MS/MS)

For the quantification of targeted hormones, flower samples were finely ground with liquid nitrogen following the protocol as described in Pan et al. (2010). Briefly, finely

grounded flower samples and an adequate amount of internal standard (IS) were placed in 15 mL centrifuge tubes and 2-propanol/H_2O/concentrated HCl (2:1:0.002, vol/vol/vol) were added to each tube followed by shaking at a speed of 5000 rpm for 10 min at 4°C. Thereafter, supernatants were transferred into a new tube and subjected under a gentle stream of highly purified nitrogen gas to a final volume of 3 mL and pH was adjusted to 8.0. Add twice the volume of petroleum ether to the solvent phase and shake at a speed of 5000 rpm for 10 min at 4°C and repeat this step. The sample solution was injected into the reverse-phase C_{18} Gemini HPLC column for UPLC-MS/MS analysis. The parameters of mass spectrometry for the measurement of hormones in the *H. coronarium* flowers are given in **Supplementary Table 4**. The experiment was performed in three biological and three technical replicates.

GC-MS Analysis

To analyze volatile compounds, a flower was placed in a 250-mL glass bottle and covered with an aluminum sheet; ethyl caprate was used as an internal standard. After 30 min of incubation, a PDMS fiber was inserted into the bottle, incubated for 30 min to adsorb volatiles, and injected into a gas chromatography-mass spectrometry system (Agilent) for volatile analysis as described previously (Yue et al., 2015). The experiment was performed in five to seven biological replicates.

Statistical Analysis

All data were analyzed using LSD with Origin software. P-values < 0.05 were considered to be significant.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**, further inquiries can be directed to the corresponding author/s.

AUTHOR CONTRIBUTIONS

YK, FA, and YF conceived and designed the experiment. YK, FA, and YZ contributed to the experiment and data analysis. YK and FA wrote the manuscript. FA, YF, and RY revised the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by the Key-Areas Research and Development Program of Guangdong Province (Grant No. 2020B020220007), National Natural Science Foundation of China (Grant No. 31770738), People's Livelihood Science and Technology Projects of Guangzhou (Grant No. 201903010054) to YF, and National Natural Science Foundation of China to RY (Grant No. 31870690).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2021.710826/full#supplementary-material

Supplementary Figure 1 | Phylogenetic analysis of six HcMYB proteins with *Arabidopsis* MYB protein family.

Supplementary Figure 2 | Correlation between *HcMYB1* and *HcMYB2* expression and the emission of volatile compounds.

Supplementary Figure 3 | A comparative analysis of transcript abundance of six *HcMYB* genes during flower development using RNA-seq data.

Supplementary Figure 4 | The protein expression level of HcMYB1 and HcIAA4 in different tissues.

Supplementary Figure 5 | The expression levels of key structural genes (*HcTPS5* and *HcBSMT2*) in different tissues.

Supplementary Figure 6 | Heatmap showing expression profiles (log_2TPM) of *HcBSMT* genes in different tissues.

Supplementary Figure 7 | Heatmap showing expression profiles (log_2TPM) of *HcTPS* genes in different tissues.

Supplementary Table 1 | MYB-core binding motifs in the promoter regions of *HcBSMTs* and *HcTPSs*.

Supplementary Table 2 | Primers used in this study.

Supplementary Table 3 | Genes used for phylogenetic analysis and their accession numbers.

Supplementary Table 4 | The parameters index of UPLC-MS/MS used for the measurement of hormones in the *H. coronarium* flowers.

REFERENCES

Abbas, F., Ke, Y., Yu, R., and Fan, Y. (2019). Functional characterization and expression analysis of two terpene synthases involved in floral scent formation in *Lilium* 'Siberia'. *Planta* 249, 71–93. doi: 10.1007/s00425-018-3006-7

Abbas, F., Ke, Y., Yu, R., Yue, Y., Amanullah, S., Jahangir, M. M., et al. (2017). Volatile terpenoids: multiple functions, biosynthesis, modulation and manipulation by genetic engineering. *Planta* 246, 803–816. doi: 10.1007/s00425-017-2749-x

Abbas, F., Ke, Y., Zhou, Y., Waseem, M., Yu, Y., Ashraf, U., et al. (2020). Cloning, functional characterization and expression analysis of *LoTPS5* from *Lilium* 'Siberia'. *Gene* 756:144921. doi: 10.1016/j.gene.2020.144921

Abbas, F., Ke, Y., Zhou, Y., Yu, Y., Waseem, M., Ashraf, U., et al. (2021a). Genome-wide analysis of ARF transcription factors reveals HcARF5 expression profile associated with the biosynthesis of β-ocimene synthase in *Hedychium coronarium*. *Plant Cell Rep*. 40, 1269–1284. doi: 10.1007/s00299-021-02709-1

Abbas, F., Ke, Y., Zhou, Y., Yu, Y., Waseem, M., Ashraf, U., et al. (2021b). Genome-wide analysis reveals the potential role of MYB transcription factors in floral scent formation in *Hedychium coronarium*. *Front. Plant Sci.* 12:623742. doi: 10.3389/fpls.2021.623742

Aharoni, A., De Vos, C. R., Wein, M., Sun, Z., Greco, R., Kroon, A., et al. (2001). The strawberry *FaMYB1* transcription factor suppresses anthocyanin and flavonol accumulation in transgenic tobacco. *Plant J.* 28, 319–332. doi: 10.1046/j.1365-313x.2001.01154.x

An, X. H., Tian, Y., Chen, K. Q., Liu, X. J., Liu, D. D., Xie, X. B., et al. (2014). MdMYB9 and MdMYB11 are Involved in the Regulation of the JA-Induced Biosynthesis of Anthocyanin

and Proanthocyanidin in Apples. *Plant Cell Physiol.* 56, 650–662. doi: 10.1093/pcp/pcu205

Báez, D., Pino, J. A., and Morales, D. (2011). Floral scent composition in *Hedychium coronarium* J. Koenig analyzed by SPME. *J. Essent. Oil Res.* 23, 64–67. doi: 10.1080/10412905.2011.9700460

Bedon, F., Bomal, C., Caron, S., Levasseur, C., Boyle, B., Mansfield, S. D., et al. (2010). Subgroup 4 R2R3-MYBs in conifer trees: gene family expansion and contribution to the isoprenoid-and flavonoid-oriented responses. *J. Exp. Bot.* 61, 3847–3864. doi: 10.1093/jxb/erq196

Cao, Y., Jia, H., Xing, M., Jin, R., Grierson, D., Gao, Z., et al. (2021). Genome-wide analysis of MYB gene family in Chinese bayberry (*Morella rubra*) and identification of members regulating flavonoid biosynthesis. *Front. Plant Sci.* 12:691384. doi: 10.3389/fpls.2021.691384

Chai, Y. M., Jia, H. F., Li, C. L., Dong, Q. H., and Shen, Y. Y. (2011). FaPYR1 is involved in strawberry fruit ripening. *J. Exp. Bot.* 62, 5079–5089. doi: 10.1093/jxb/err207

Colquhoun, T. A., Schwieterman, M. L., Wedde, A. E., Schimmel, B. C., Marciniak, D. M., Verdonk, J. C., et al. (2011). *EOBII* controls flower opening by functioning as a general transcriptomic switch. *Plant Physiol.* 156, 974–984. doi: 10.1104/pp.111.176248

Deluc, L., Barrieu, F., Marchive, C., Lauvergeat, V., Decendit, A., Richard, T., et al. (2006). Characterization of a grapevine R2R3-MYB transcription factor that regulates the phenylpropanoid pathway. *Plant Physiol.* 140, 499–511. doi: 10.1104/pp.105.067231

Deluc, L., Bogs, J., Walker, A. R., Ferrier, T., Decendit, A., Merillon, J. M., et al. (2008). The transcription factor VvMYB5b contributes to the regulation of anthocyanin and proanthocyanidin biosynthesis in developing grape berries. *Plant Physiol.* 147, 2041–2053. doi: 10.1104/pp.108.118919

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

Dubos, C., Le Gourrierec, J., Baudry, A., Huep, G., Lanet, E., Debeaujon, I., et al. (2008). MYBL2 is a new regulator of flavonoid biosynthesis in *Arabidopsis thaliana*. *Plant J.* 55, 940–953. doi: 10.1111/j.1365-313X.2008.03564.x

Dubos, C., Stracke, R., Grotewold, E., Weisshaar, B., Martin, C., and Lepiniec, L. (2010). MYB transcription factors in *Arabidopsis*. *Trends Plant Sci.* 15, 573–581. doi: 10.1016/j.tplants.2010.06.005

Dudareva, N., Klempien, A., Muhlemann, J. K., and Kaplan, I. (2013). Biosynthesis, function and metabolic engineering of plant volatile organic compounds. *New Phytol.* 198, 16–32. doi: 10.1111/nph.12145

Dudareva, N., Negre, F., Nagegowda, D. A., and Orlova, I. (2006). Plant volatiles: recent advances and future perspectives. *Crit. Rev. Plant Sci.* 25, 417–440. doi: 10.1080/07352680600899973

Fan, Y. P., Yu, R., Huang, Y., and Chen, Y. (2003). Studies on the essential constituent of *Hedychium flavum* and *H. coronarium*. *Acta Hortic. Sin.* 30:475. doi: 10.16420/j.issn.0513-353x.2003.04.030

Fan, Y.-P., Wang, X.-R., Yu, R.-C., and Yang, P. (2007). Analysis on the aroma components in several species of *Hedychium*. *Acta Hortic. Sin.* 34, 231–234. doi: 10.16420/j.issn.0513-353x.2007.01.049

Gershenzon, J., and Dudareva, N. (2007). The function of terpene natural products in the natural world. *Nat. Chem. Biol.* 3, 408–414. doi: 10.1038/nchembio.2007.5

Hellens, R. P., Allan, A. C., Friel, E. N., Bolitho, K., Grafton, K., Templeton, M. D., et al. (2005). Transient expression vectors for functional genomics, quantification of promoter activity and RNA silencing in plants. *Plant Methods* 1:13. doi: 10.1186/1746-4811-1-13

Hong, G. J., Xue, X. Y., Mao, Y. B., Wang, L. J., and Chen, X. Y. (2012). *Arabidopsis* MYC2 interacts with DELLA proteins in regulating sesquiterpene synthase gene expression. *Plant Cell* 24, 2635–2648. doi: 10.1105/tpc.112.098749

Jaradat, M. R., Feurtado, J. A., Huang, D., Lu, Y., and Cutler, A. J. (2013). Multiple roles of the transcription factor AtMYBR1/AtMYB44 in ABA signaling, stress responses, and leaf senescence. *BMC Plant Biol.* 13:192. doi: 10.1186/1471-2229-13-192

Jia, H. F., Chai, Y. M., Li, C. L., Lu, D., Luo, J. J., Qin, L., et al. (2011). Abscisic acid plays an important role in the regulation of strawberry fruit ripening. *Plant Physiol.* 157, 188–199. doi: 10.1104/pp.111.177311

Jia, H., Xie, Z., Wang, C., Shangguan, L., Qian, N., Cui, M., et al. (2017). Abscisic acid, sucrose, and auxin coordinately regulate berry ripening process of the Fujiminori grape. *Funct. Integr. Genomic* 17, 441–457. doi: 10.1007/s10142-017-0546-z

Jiang, Y., Liang, G., Yang, S., and Yu, D. (2014). *Arabidopsis* WRKY57 functions as a node of convergence for jasmonic acid–and auxin-mediated signaling in jasmonic acid–induced leaf senescence. *Plant Cell* 26, 230–245. doi: 10.1105/tpc.113.117838

Katiyar, A., Smita, S., Lenka, S. K., Rajwanshi, R., Chinnusamy, V., and Bansal, K. C. (2012). Genome-wide classification and expression analysis of MYB transcription factor families in rice and *Arabidopsis*. *BMC Genomics* 13:544. doi: 10.1186/1471-2164-13-544

Ke, M., Gao, Z., Chen, J., Qiu, Y., Zhang, L., and Chen, X. (2018). Auxin controls circadian flower opening and closure in the waterlily. *BMC Plant Biol.* 18:143. doi: 10.1186/s12870-018-1357-7

Ke, Y., Abbas, F., Zhou, Y., Yu, R., Yue, Y., Li, X., et al. (2019). Genome-wide analysis and characterization of the Aux/IAA family genes related to floral scent formation in *Hedychium coronarium*. *Int. J. Mol. Sci.* 20:3235. doi: 10.3390/ijms20133235

Kim, H. J., Park, K. J., and Lim, J. H. (2011). Metabolomic analysis of phenolic compounds in buckwheat (*Fagopyrum esculentum* M.) sprouts treated with methyl jasmonate. *J. Agric. Food Chem.* 59, 5707–5713. doi: 10.1021/jf200396k

Kranz, H. D., Denekamp, M., Greco, R., Jin, H., Leyva, A., Meissner, R. C., et al. (1998). Towards functional characterisation of the members of the R2R3−MYB gene family from *Arabidopsis thaliana*. *Plant J.* 16, 263–276. doi: 10.1046/j.1365-313x.1998.00278.x

Krizek, B. A. (2011). Auxin regulation of *Arabidopsis* flower development involves members of the aintegumenta-like/plethora (AIL/PLT) family. *J. Exp. Bot.* 62, 3311–3319. doi: 10.1093/jxb/err127

Kumar, S., Stecher, G., Li, M., Knyaz, C., and Tamura, K. (2018). MEGA X: molecular evolutionary genetics analysis across computing platforms. *Mol. Biol. Evol.* 35, 1547–1549. doi: 10.1093/molbev/msy096

Lai, B., Du, L. N., Liu, R., Hu, B., Su, W. B., Qin, Y. H., et al. (2016). Two LcbHLH transcription factors interacting with LcMYB1 in regulating late structural genes of anthocyanin biosynthesis in *Nicotiana* and *Litchi chinensis* during anthocyanin accumulation. *Front. Plant Sci.* 7:166. doi: 10.3389/fpls.2016.00166

Lan, J. B., Yu, R. C., Yu, Y. Y., and Fan, Y. P. (2013). Molecular cloning and expression of *Hedychium coronarium* farnesyl pyrophosphate synthase gene and its possible involvement in the biosynthesis of floral and wounding/herbivory induced leaf volatile sesquiterpenoids. *Gene* 518, 360–367. doi: 10.1016/j.gene.2013.01.007

Lavid, N., Wang, J., Shalit, M., Guterman, I., Bar, E., Beuerle, T., et al. (2002). O-methyltransferases involved in the biosynthesis of volatile phenolic derivatives in rose petals. *Plant Physiol.* 129, 1899–1907. doi: 10.1104/pp.005330

Li, J., Li, X., Guo, L., Lu, F., Feng, X., He, K., et al. (2006a). A subgroup of MYB transcription factor genes undergoes highly conserved alternative splicing in *Arabidopsis* and rice. *J. Exp. Bot.* 57, 1263–1273. doi: 10.1093/jxb/erj094

Li, J., Yang, X., Wang, Y., Li, X., Gao, Z., Pei, M., et al. (2006b). Two groups of MYB transcription factors share a motif which enhances trans-activation activity. *Biochem. Biophys. Res. Commun.* 341, 1155–1163. doi: 10.1016/j.bbrc.2006.01.077

Li, R., and Fan, Y. (2011). Molecular cloning and expression analysis of a terpene synthase gene, *HcTPS2*, in *Hedychium coronarium*. *Plant Mol. Biol. Rep.* 29, 35–42. doi: 10.1007/s11105-010-0205-1

Li, Y., Shan, X., Zhou, L., Gao, R., Yang, S., Wang, S., et al. (2019). The R2R3-MYB factor FhMYB5 from *Freesia hybrida* contributes to the regulation of anthocyanin and proanthocyanidin biosynthesis. *Front. Plant Sci.* 9:1935. doi: 10.3389/fpls.2018.01935

Liao, W., Yang, Y., Li, Y., Wang, G., and Peng, M. (2016). Genome-wide identification of cassava R2R3 MYB family genes related to abscission zone separation after environmental-stress-induced abscission. *Sci. Rep.* 6:32006. doi: 10.1038/srep32006

Liu, G., Ren, G., Guirgis, A., and Thornburg, R. W. (2009). The MYB305 transcription factor regulates expression of nectarin genes in the ornamental

tobacco floral nectary. *Plant Cell* 21, 2672–2687. doi: 10.1105/tpc.108.060079

Liu, J., Osbourn, A., and Ma, P. (2015). MYB transcription factors as regulators of phenylpropanoid metabolism in plants. *Mol. Plant* 8, 689–708. doi: 10.1016/j.molp.2015.03.012

Liu, L., Ramsay, T., Zinkgraf, M., Sundell, D., Street, N. R., Filkov, V., et al. (2015). A resource for characterizing genome−wide binding and putative target genes of transcription factors expressed during secondary growth and wood formation in *Populus*. *Plant J.* 82, 887–898. doi: 10.1111/tpj.12850

Livak, K. J., and Schmittgen, T. D. (2001). Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. *Methods* 25, 402–408. doi: 10.1006/meth.2001.1262

Luo, J., Butelli, E., Hill, L., Parr, A., Niggeweg, R., Bailey, P., et al. (2008). AtMYB12 regulates caffeoyl quinic acid and flavonol synthesis in tomato: expression in fruit results in very high levels of both types of polyphenol. *Plant J.* 56, 316–326. doi: 10.1111/j.1365-313X.2008.03597.x

Medina-Puche, L., Cumplido-Laso, G., Amil-Ruiz, F., Hoffmann, T., Ring, L., Rodríguez-Franco, A., et al. (2014). *MYB10* plays a major role in the regulation of flavonoid/phenylpropanoid metabolism during ripening of *Fragaria*× *ananassa* fruits. *J. Exp. Bot.* 65, 401–417. doi: 10.1093/jxb/ert377

Medina-Puche, L., Molina-Hidalgo, F. J., Boersma, M., Schuurink, R. C., López-Vidriero, I., Solano, R., et al. (2015). An R2R3-MYB transcription factor regulates eugenol production in ripe strawberry fruit receptacles. *Plant Physiol.* 168, 598–614. doi: 10.1104/pp.114.252908

Muhlemann, J. K., Klempien, A., and Dudareva, N. (2014). Floral volatiles: from biosynthesis to function. *Plant Cell Environ.* 37, 1936–1949. doi: 10.1111/pce.12314

Muhlemann, J. K., Maeda, H., Chang, C. Y., San Miguel, P., Baxter, I., Cooper, B., et al. (2012). Developmental changes in the metabolic network of snapdragon flowers. *PLoS One* 7:e40381. doi: 10.1371/journal.pone.0040381

Niu, Q., Zong, Y., Qian, M., Yang, F., and Teng, Y. (2014). Simultaneous quantitative determination of major plant hormones in pear flowers and fruit by UPLC/ESI-MS/MS. *Anal. Methods* 6, 1766–1773. doi: 10.1039/c3ay41885E

Oono, Y., Ooura, C., Rahman, A., Aspuria, E. T., Hayashi, K. I., Tanaka, A., et al. (2003). p-Chlorophenoxyisobutyric acid impairs auxin response in *Arabidopsis* root. *Plant Physiol.* 133, 1135–1147. doi: 10.1104/pp.103.027847

Pan, X., Welti, R. and Wang, X. (2010). Quantitative analysis of major plant hormones in crude plant extracts by high-performance liquid chromatographyŰmass spectrometry. *Nat. Protoc.* 5, 986–992. doi: 10.1038/nprot.2010.37

Perkins-Veazie, P. (1995). Growth and ripening of strawberry fruit. *Hortic. Rev.* 17, 267–297. doi: 10.1002/9780470650585.ch8

Pichersky, E., and Dudareva, N. (2007). Scent engineering: toward the goal of controlling how flowers smell. *Trends Biotechnol.* 25, 105–110. doi: 10.1016/j.tibtech.2007.01.002

Qi, T., Huang, H., Wu, D., Yan, J., Qi, Y., Song, S., et al. (2014). *Arabidopsis* DELLA and JAZ proteins bind the WD-repeat/bHLH/MYB complex to modulate gibberellin and jasmonate signaling synergy. *Plant Cell* 26, 1118–1133. doi: 10.1105/tpc.113.121731

Raguso, R. A. (2009). Floral scent in a whole−plant context: moving beyond pollinator attraction. *Funct. Ecol.* 23, 837–840. doi: 10.1111/j.1365-2435.2009.01643.x

Ramya, M., Kwon, O. K., An, H. R., Park, P. M., Baek, Y. S., and Park, P. H. (2017). Floral scent: regulation and role of MYB transcription factors. *Phytochem. Lett.* 19, 114–120. doi: 10.1016/j.phytol.2016.12.015

Ramya, M., Lee, S. Y., An, H. R., Park, P. M., Kim, N. S., and Park, P. H. (2019). MYB1 transcription factor regulation through floral scent in *Cymbidium* cultivar 'Sael Bit'. *Phytochem. Lett.* 32, 181–187. doi: 10.1016/j.phytol.2019.06.007

Reddy, V. A., Qian, W., Dhar, N., Kumar, N., Venkatesh, P. N., Rajan, C., et al. (2017). Spearmint R2R3−MYB transcription factor MsMYB negatively regulates monoterpene production and suppresses the expression of geranyl diphosphate synthase large subunit (*MsGPPS. LSU*). *Plant Biotechnol. J.* 15, 1105–1119. doi: 10.1111/pbi.12701

Renner, T., Bragg, J., Driscoll, H. E., Cho, J., Jackson, A. O., and Specht, C. D. (2009). Virus-induced gene silencing in the culinary ginger (*Zingiber officinale*): an effective mechanism for down-regulating gene expression in tropical monocots. *Mol. Plant* 2, 1084–1094. doi: 10.1093/mp/ssp033

Rommens, C. M., Richael, C. M., Yan, H., Navarre, D. A., Ye, J., Krucker, M., et al. (2008). Engineered native pathways for high kaempferol and caffeoylquinate production in potato. *Plant Biotechnol. J.* 6, 870–886. doi: 10.1111/j.1467-7652.2008.00362.x

Rushton, P. J., Somssich, I. E., Ringler, P., and Shen, Q. J. (2010). WRKY transcription factors. *Trends Plant Sci.* 15, 247–258. doi: 10.1016/j.tplants.2010.02.006

Shin, B., Choi, G., Yi, H., Yang, S., Cho, I., Kim, J., et al. (2002). *AtMYB21*, a gene encoding a flower−specific transcription factor, is regulated by COP1. *Plant J.* 30, 23–32. doi: 10.1046/j.1365-313x.2002.01264.x

Shin, R., Burch, A. Y., Huppert, K. A., Tiwari, S. B., Murphy, A. S., Guilfoyle, T. J., et al. (2007). The *Arabidopsis* transcription factor MYB77 modulates auxin signal transduction. *Plant Cell* 19, 2440–2453. doi: 10.1105/tpc.107.050963

Sievers, F., Wilm, A., Dineen, D., Gibson, T. J., Karplus, K., Li, W., et al. (2011). Fast, scalable generation of high−quality protein multiple sequence alignments using Clustal Omega. *Mol. Syst. Biol.* 7:539. doi: 10.1038/msb.2011.75

Spitzer-Rimon, B., Farhi, M., Albo, B., Cna'ani, A., Ben Zvi, M. M., Masci, T., et al. (2012). The R2R3-MYB–like regulatory factor EOBI, acting downstream of EOBII, regulates scent production by activating *ODO1* and structural scent-related genes in petunia. *Plant Cell* 24, 5089–5105. doi: 10.1105/tpc.112.105247

Spitzer-Rimon, B., Marhevka, E., Barkai, O., Marton, I., Edelbaum, O., Masci, T., et al. (2010). *EOBII*, a gene encoding a flower-specific regulator of phenylpropanoid volatiles' biosynthesis in petunia. *Plant Cell* 22, 1961–1976. doi: 10.1105/tpc.109.067280

Tan, H., Man, C., Xie, Y., Yan, J., Chu, J., and Huang, J. (2019). A crucial role of GA-regulated flavonol biosynthesis in root growth of *Araboidopsis*. *Mol. Plant* 12, 521–537. doi: 10.1016/j.molp.2018.12.021

Uimari, A., and Strommer, J. (1997). Myb26: a MYB−like protein of pea flowers with affinity for promoters of phenylpropanoid genes. *Plant J.* 12, 1273–1284. doi: 10.1046/j.1365-313x.1997.12061273.x

Van Moerkercke, A., Haring, M. A., and Schuurink, R. C. (2011). The transcription factor emission of benzenoids II activates the MYB *ODORANT1* promoter at a MYB binding site specific for fragrant petunias. *Plant J.* 67, 917–928. doi: 10.1111/j.1365-313X.2011.04644.x

Verdonk, J. C., Haring, M. A., van Tunen, A. J., and Schuurink, R. C. (2005). ODORANT1 regulates fragrance biosynthesis in petunia flowers. *Plant Cell* 17, 1612–1624. doi: 10.1105/tpc.104.028837

Wu, Q., Tao, X., Ai, X., Luo, Z., Mao, L., Ying, T., et al. (2018). Contribution of abscisic acid to aromatic volatiles in cherry tomato (*Solanum lycopersicum* L.) fruit during postharvest ripening. *Plant Physiol. Biochem.* 130, 205–214. doi: 10.1016/j.plaphy.2018.06.039

Yamaguchi, N., Wu, M. F., Winter, C. M., Berns, M. C., Nole-Wilson, S., Yamaguchi, A., et al. (2013). A molecular framework for auxin-mediated initiation of flower primordia. *Dev. Cell* 24, 271–282. doi: 10.1016/j.devcel.2012.12.017

Yan, H., Zhang, H., Wang, Q., Jian, H., Qiu, X., Wang, J., et al. (2011). Isolation and identification of a putative scent-related gene *RhMYB1* from rose. *Mol. Biol. Rep.* 38, 4475–4482. doi: 10.1007/s11033-010-0577-1

Yang, Z., Li, Y., Gao, F., Jin, W., Li, S., Kimani, S., et al. (2020). MYB21 interacts with MYC2 to control the expression of terpene synthase genes in flowers of *Freesia hybrida* and *Arabidopsis thaliana*. *J. Exp. Bot.* 71, 4140–4158. doi: 10.1093/jxb/eraa184

Yoo, S. D., Cho, Y. H., and Sheen, J. (2007). *Arabidopsis* mesophyll protoplasts: a versatile cell system for transient gene expression analysis. *Nat. Protoc.* 2:1565. doi: 10.1038/nprot.2007.199

Yoshida, K., Oyama-Okubo, N., and Yamagishi, M. (2018). An R2R3-MYB transcription factor ODORANT1 regulates fragrance biosynthesis in lilies (*Lilium* spp.). *Mol. Breed.* 38, 1–14. doi: 10.1007/s11032-018-0902-2

Yuan, C., Li, C., Yan, L., Jackson, A. O., Liu, Z., Han, C., et al. (2011). A high throughput barley stripe mosaic virus vector for virus induced gene silencing in monocots and dicots. *PLoS One* 6:e26468. doi: 10.1371/journal.pone.0026468

Yue, Y., Wang, L., Yu, R., Chen, F., He, J., Li, X., et al. (2021). Coordinated and high-level expression of biosynthetic pathway genes is responsible for

the production of a major floral scent compound methyl benzoate in *Hedychium coronarium*. *Front. Plant Sci.* 12:650582. doi: 10.3389/fpls.2021.650582

Yue, Y., Yu, R., and Fan, Y. (2015). Transcriptome profiling provides new insights into the formation of floral scent in *Hedychium coronarium*. *BMC Genomics* 16:470. doi: 10.1186/s12864-015-1653-7

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, X., He, Y., Li, L., Liu, H., and Hong, G. (2021). Involvement of the R2R3-MYB transcription factors MYB21 and its homologs in regulating the stamen flavonols accumulation in *Arabidopsis*. *J. Exp. Bot.* 72, 4319–4332. doi: 10.1093/jxb/erab156

Zhao, Y., Xing, L., Wang, X., Hou, Y. J., Gao, J., Wang, P., et al. (2014). The ABA receptor PYL8 promotes lateral root growth by enhancing MYB77-dependent transcription of auxin-responsive genes. *Sci. Signal.* 7:ra53. doi: 10.1126/scisignal.2005051

Zhou, M., and Memelink, J. (2016). Jasmonate-responsive transcription factors regulating plant secondary metabolism. *Biotechnol. Adv.* 34, 441–449. doi: 10.1016/j.biotechadv.2016.02.004

Zhou, M., Zhang, K., Sun, Z., Yan, M., Chen, C., Zhang, X., et al. (2017). LNK1 and LNK2 corepressors interact with the MYB3 transcription factor in phenylpropanoid biosynthesis. *Plant Physiol.* 174, 1348–1358. doi: 10.1104/pp.17.00160

Zhu, N., Cheng, S., Liu, X., Du, H., Dai, M., Zhou, D. X., et al. (2015). The R2R3-type MYB gene *OsMYB91* has a function in coordinating plant growth and salt stress tolerance in rice. *Plant Sci.* 236, 146–156. doi: 10.1016/j.plantsci.2015.03.023

Zvi, M. M. B., Shklarman, E., Masci, T., Kalev, H., Debener, T., Shafir, S., et al. (2012). PAP1 transcription factor enhances production of phenylpropanoid and terpenoid scent compounds in rose flowers. *New Phytol.* 195, 335–345. doi: 10.1111/j.1469-8137.2012.04161.x

Zou, X., Neuman, D., and Shen, Q. J. (2008). Interactions of two transcriptional repressors and two transcriptional activators in modulating gibberellin signaling in aleurone cells. *Plant Physiol.* 148, 176–186. doi: 10.1104/pp.108.123653

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.

frontiers in Plant Science

ORIGINAL RESEARCH
published: 23 November 2021
doi: 10.3389/fpls.2021.762330

Integration of Metabolome and Transcriptome Reveals the Relationship of Benzenoid–Phenylpropanoid Pigment and Aroma in Purple Tea Flowers

*Xin Mei[1,2], Shihua Wan[1], Chuyuan Lin[1], Caibi Zhou[2], Liuhong Hu[2], Chan Deng[2] and Lingyun Zhang[1]**

[1] College of Horticulture, South China Agricultural University, Guangzhou, China, [2] College of Biological Science and Agriculture, Qiannan Normal University for Nationalities, Duyun, China

OPEN ACCESS

Edited by:
Xiumin Fu,
South China Botanical Garden, Chinese Academy of Sciences (CAS), China

Reviewed by:
Ying Zhou,
Hainan Institute of Zhejiang University, China
JunYan Zhu,
Anhui Agricultural University, China
Xinghui Li,
Nanjing Agricultural University, China

***Correspondence:**
Lingyun Zhang
zhanglingyun@scau.edu.cn

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

Received: *21 August 2021*
Accepted: *22 October 2021*
Published: *23 November 2021*

Citation:
Mei X, Wan S, Lin C, Zhou C, Hu L, Deng C and Zhang L (2021) Integration of Metabolome and Transcriptome Reveals the Relationship of Benzenoid–Phenylpropanoid Pigment and Aroma in Purple Tea Flowers. Front. Plant Sci. 12:762330. doi: 10.3389/fpls.2021.762330

Tea (*Camellia sinensis*) flowers are normally white, even though the leaves could be purple. We previously discovered a specific variety with purple leaves and flowers. In the face of such a phenomenon, researchers usually focus on the mechanism of color formation but ignore the change of aroma. The purple tea flowers contain more anthocyanins, which belong to flavonoids. Meanwhile, phenylalanine (Phe), derived from the shikimate pathway, is a precursor for both flavonoids and volatile benzenoid–phenylpropanoids (BPs). Thus, it is not clear whether the BP aroma was attenuated for the appearance of purple color. In this study, we integrated metabolome and transcriptome of petals of two tea varieties, namely, Zijuan (ZJ) with white flowers and Baitang (BT) with purple flowers, to reveal the relationship between color (anthocyanins) and aroma (volatile BPs). The results indicated that in purple petals, the upstream shikimate pathway promoted for 3-deoxy-D-arabino-heptulosonate 7-phosphate synthase (DAHPS) was elevated. Among the increased anthocyanins, delphinidin-3-*O*-glucoside (DpG) was extremely higher; volatile BPs, including benzyl aldehyde, benzyl alcohol, acetophenone (AP), 1-phenylethanol, and 2-phenylethanol, were also enhanced, and AP was largely elevated. The structural genes related to the biosynthesis of volatile BPs were induced, while the whole flavonoid biosynthesis pathway was downregulated, except for the genes *flavonoid 3′-hydroxylase* (*F3′H*) and *flavonoid 3′,5′-hydroxylase* (*F3′5′H*), which were highly expressed to shift the carbon flux to delphinidin, which was then conjugated to glucoside by increased bronze-1 (BZ1) (UDP-glucose: flavonoid 3-*O*-glucosyltransferase) to form DpG. Transcription factors (TFs) highly related to AP and DpG were selected to investigate their correlation with the differentially expressed structural genes. TFs, such as MYB, AP2/ERF, bZIP, TCP, and GATA, were dramatically expressed and focused on the regulation of genes in the upstream synthesis of Phe (*DAHPS*; *arogenate*

dehydratase/prephenatedehydratase) and the synthesis of AP (*phenylacetaldehyde reductase*; *short-chain dehydrogenase/reductase*), Dp (*F3′H*; *F3′5′H*), and DpG (*BZ1*), but inhibited the formation of flavones (*flavonol synthase*) and catechins (*leucoanthocyanidin reductase*). These results discovered an unexpected promotion of volatile BPs in purple tea flowers and extended our understanding of the relationship between the BP-type color and aroma in the tea plant.

Keywords: *Camellia sinensis*, petal, anthocyanin, volatile, transcription factor, acetophenone, delphinidin-3-*O*-glucoside, transportation

INTRODUCTION

Benzenoids, especially phenylpropanoids, are the main sources of plant color and aroma. They help the plant to attract pollinators or deter enemies. Humans also gain health benefits such as antioxidation or improving memory (Cheng et al., 2016).

Benzenoids and phenylpropanoids (BPs) are a group of secondary metabolites originating from two aromatic amino acids phenylalanine (Phe) and tyrosine in the shikimic acid pathway. Phenylpropanoids contain an aromatic ring and a three-carbon propene tail. The simplest phenylpropanoid, cinnamic acid, is derived from the elimination of ammonia from Phe and transferred by cinnamate-4 hydroxylase (C4H) and 4-coumarate-CoA ligase (4CL); it leads to the biosynthesis of other phenylpropanoids, including flavonoids (Stevenson and Aslam, 2006). In flavonoid biosynthesis, chalcone is first generated by chalcone synthase (CHS) and then is isomerized to naringenin by chalcone isomerase (CHI). Naringenin is oxidized to dihydroflavonol by flavanone 3-hydroxylase (F3H). Flavonoid 3′-hydroxylase (F3′H) and flavonoid 3′,5′-hydroxylase (F3′5′H) add more hydroxyls to the secondary benzene ring of dihydroflavonol, which later direct to different types of catechins or anthocyanidins. Dihydroflavonol is reduced to leucoanthocyanidin, which is transferred to (+)-catechins and anthocyanidins by leucoanthocyanidin reductase (LAR) and anthocyanidin synthase (ANS), respectively. Anthocyanidins could be transferred to (−)-catechins by anthocyanidin reductase (ANR) or linked with glucose by UDP-glucose: flavonoid 3-*O*-glucosyltransferase (UFGT) to form anthocyanins. It has been widely proved that these structure genes were influenced by many transcription factors (TFs), such as *MYB*, *ERF*, *bZIP*, *TCP*, *GATA*, and so on (Wei et al., 2016; Chen et al., 2017; Jian et al., 2019; Gao et al., 2020; Fu et al., 2021). The important plant pigment anthocyanidins can be classified mainly into two groups, namely, flavonoids and phenolics. Anthocyanidins are generally conjugated with sugar molecules to form anthocyanins and accumulated in the vacuolar sap of the epidermal tissues. Anthocyanins can be divided into three subclasses, namely, anthocyanidin derivatives, non-acylated anthocyanidin glucoside, and acylated anthocyanidin glucoside (e.g., caffeoylated anthocyanin and malonylated anthocyanin) (Chandra Singh et al., 2020). The common anthocyanidin derivatives in plant include pelargonidin (Pg), cyanidin (Cy), delphinidin (Dp), peonidin (Pn), petunidin (Pt), and malvidin (Mv) (Chandra Singh et al., 2020). Moreover starting from cinnamate, common phenylpropanoid volatiles are synthesized, including phenylacetaldehyde, acetophenone (AP), 1-phenylethanol (1-PE), 2-phenylethanol (2-PE), and 2-phenylacetate. Other benzenoid volatiles are also generated from Phe but not *via* cinnamate, including benzaldehyde (Bald), benzyl alcohol (Balc), benzyl acetate, and methyl salicylate (MeSA) (**Figure 1**; Oliva et al., 2017). The diversity and richness of aroma are two reasons why tea [*Camellia sinensis* (L.) Kuntze] leaf has been made to be a popular non-alcoholic beverage. The main volatile BPs in tea leaves are 2-PE, Bald, Balc, and MeSA, while AP and 1-PE are unique in tea flower (Zhou et al., 2015; He et al., 2016). Volatile BPs and non-volatile BPs have a common upstream pathway. They both come from Phe, which is generated by the shikimic acid pathway. Phe enters Phe metabolism and Phe synthesis, respectively. The former pathway continues to form volatile BPs, and the latter forms flavonoids, including anthocyanins. The ideal material for studying BP pigments and aroma should preferably contain as many anthocyanins and volatile BPs as possible. Zijuan (ZJ; *C. sinensis* var. *kitamura*) is a special tea cultivar with purple buds, stems, and leaves, but it still lacked AP and 1-PE (He et al., 2016). Moreover, like other tea varieties, the flowers of ZJ are also white, which means low content of anthocyanins. Fortunately, we found a perfect material in the tea plantation of Baitang town, Boluo county, Guangdong province, China. Just near the ZJ, a natural mutant Baitang (BT; *C. sinensis* var. Baitang) has purple leaves and flowers (Zhou et al., 2020). These beautiful pink tea flowers provide a better model for studying the metabolism of BPs in tea plant and the relationship between color and scent caused by BPs. In our previous article, the developmental process of tea flower was divided into five stages, including mature flower bud (Stage 1), preopening (Stage 2), initial opening (Stage 3), half bloom (Stage 4), and full bloom

Abbreviations: BPs, benzenoids and phenylpropanoids; 1-PE, 1-phenylethanol; 2-PE, 2-phenylethanol; 4CL, 4-coumarate-CoA ligase; ADT, arogenate dehydratase/prephenate dehydratase; ANR, anthocyanidin reductase; ANS, anthocyanidin synthase; AP, acetophenone; Balc, benzyl alcohol; Bald, benzaldehyde; BPBT, benzyl alcohol *O*-benzoyltransferase; BT, Baitang; BZ1, bronze-1; C4H, cinnamate-4 hydroxylase; CHI, chalcone isomerase; CHS, chalcone synthase; Cy, cyanidin; DAHPS, 3-deoxy-D-arabino-heptulosonate 7-phosphate synthase; DFR, dihydroflavonol 4-reductase; Dp, delphinidin; DpG, delphinidin-3-*O*-glucoside; F3′5′H, flavonoid 3′,5′-hydroxylase; F3H, flavanone 3-hydroxylase; F3′H, flavonoid 3′-hydroxylase; FLS, flavonol synthase; GA3P, glyceraldehyde 3-phosphate; GST, glutathione *S*-transferase; LAR, leucoanthocyanidin reductase; MATE, multidrug and toxin extrusion; MeSA, methyl salicylate; MRP, multidrug resistance-associated protein; Mv, malvidin; PAR, phenylacetaldehyde reductase; PEP, phosphoenolpyruvate; Pg, pelargonidin; Phe, phenylalanine; Pn, peonidin; Pt, petunidin; SAMT, salicylic acid carboxyl methyltransferase; SDR, short-chain dehydrogenase/reductase; TFs, transcription factors; TPM, Transcripts Per Kilobase of exon model per Million mapped reads; UFGT, UDP-glucose: flavonoid 3-*O*-glucosyltransferase; ZJ, Zijuan.

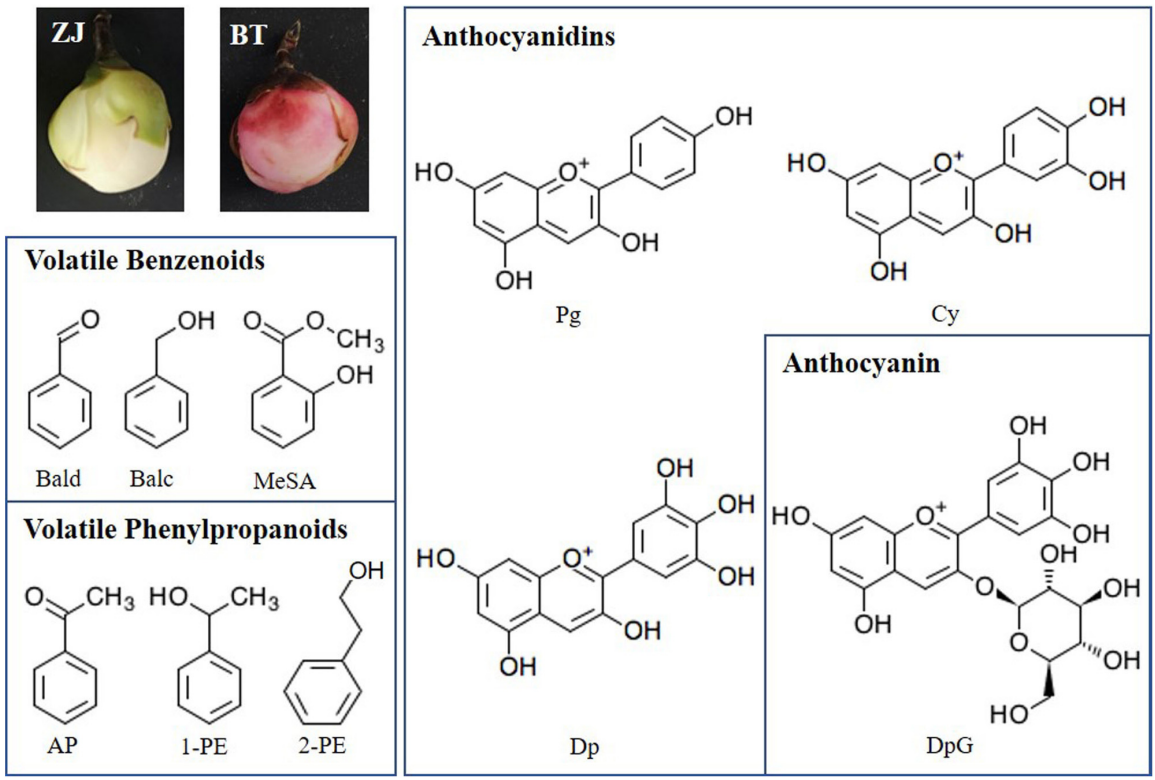


FIGURE 1 | White [Zijuan (ZJ) variety] and purple [Baitang (BT) variety] tea flowers and structure of main benzenoid–phenylpropanoids (BPs) detected in this study. Bald, benzaldehyde; Balc, benzyl alcohol; MeSA, methyl salicylate; AP, acetophenone; 1-PE, 1-phenylethanol; 2-PE, 2-phenylethanol; Pg, pelargonidin; Cy, cyanidin; Dp, delphinidin; DpG, Dp-3-*O*-glucoside.

(Stage 5). We found that total anthocyanin contents were similar between Stage 2 and Stage 3 and between Stage 4 and Stage 5, and anthocyanin had been generated in S1. To distinguish the obvious differences between stages, this time, we combined the similar stages defined last time and redivided tea flowers into three stages, namely, mature flower bud (Stage 1), before blooming (Stage 2), and blooming (Stage 3). In addition, the floral aroma will generally be released in large quantities after blooming. Considering the formation of anthocyanin in early stages and the amount of aroma, Stage 2 was thus selected as the experimental object in this study.

In this study, to discover the contribution of volatile and non-volatile BPs to the scent and color in purple tea flowers, BP-type aroma and anthocyanins were determined in white (ZJ variety) and purple (BT variety) tea petals (**Figure 1**), gene expression profile in the relevant pathways of biosynthesis and anthocyanin transportation were identified, and possible TFs for regulating these changes between ZJ and BT were analyzed.

MATERIALS AND METHODS

Sample Preparation

The pink-flowered BT (*C. sinensis* L. var. Baitang) and the white-flowered ZJ (*C. sinensis* L. var. *kitamura*) were cultured in the tea plantation of Baitang town, Guangdong, China. Their fresh flowers of Stage 2 (preopening) were plucked, and the petals were collected and immediately fixed in liquid nitrogen on December 12, 2019. Each variety had three biological replicates. Each replicate was then ground in liquid nitrogen and stored at −80°C for standby.

Transcriptome

Transcriptome of tea petals was sequenced and analyzed as described earlier (Zhou et al., 2020). In brief, 1 g of petals was extracted by using Sangon Total RNA Purification Kit (Shanghai Sangon Biotechnology Co., Ltd., Shanghai, China). The total RNA was sequenced and assembled on the Illumina HiSeq 2500 platform by the Biomarker Technologies Corporation (Beijing, China). The raw data have been uploaded to National Genomics Data Center, China, National Center for Bioinformation (NGDC, CNCB) with an accession number CRA005021.[1] The clean reads were mapped and annotated following the reference genome (Wei et al., 2018) on TPIA website[2] by using HISAT2 program (Xia et al., 2019). The levels of gene expression were estimated by Transcripts Per Kilobase of exon model per Million mapped reads (TPM) method. The significant difference between the two varieties was examined by DESeq2 R package (1.10.1).

[1]https://bigd.big.ac.cn/gsa/browse/CRA005021
[2]http://tpdb.shengxin.ren/

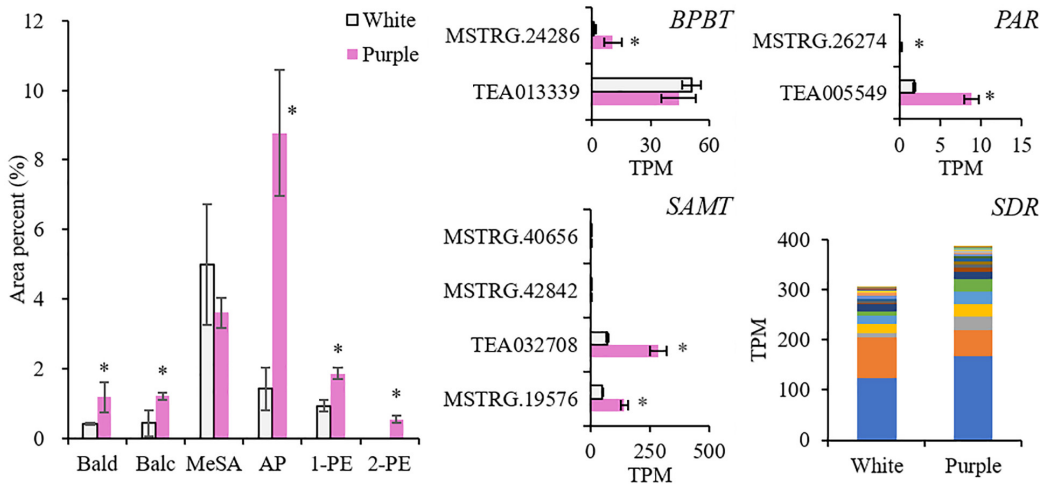


FIGURE 2 | Contents of volatile benzenoid–phenylpropanoids (BPs) and related genes in white [Zijuan (ZJ) variety] and purple [Baitang (BT) variety] tea flowers. The stacked columns of *SDR* contain the gene members with significant difference between the two tea varieties. *Indicates significant difference ($p < 0.05$) between the white and the purple. Bald, benzaldehyde; Balc, benzyl alcohol; MeSA, methyl salicylate; AP, acetophenone; 1-PE, 1-phenylethanol; 2-PE, 2-phenylethanol; BPBT, benzoyl-CoA:benzyl alcohol/2-phenylethanol benzoyltransferase; SAMT, salicylic acid carboxyl methyltransferase; PAR, phenylacetaldehyde reductase; SDR, short-chain dehydrogenase/reductase.

Differentially expressed genes were filtered by selecting a fold change greater than 2 and a false discovery rate less than 0.01.

Quantitative Real-Time PCR

The quantitative real-time PCR (qRT-PCR) was performed as described before (Zhou et al., 2020). Briefly, total RNA was extracted from petals by using RNAqueous^{TM} Total RNA Isolation Kit (Thermo, MA, United States). Primers of selected gene members were designed with the Primer Premier 5.0 software and listed in **Supplementary Table 4**. β-Actin was used as the reference gene. The fluorescence PCR reagent was the Hieff^{TM} qPCR SYBR Green Master Mix (No Rox) (Yaesen Biotech Co., Ltd., Shanghai, China). The experiment and analysis were carried out on the LightCycler® 480 II Real-Time System (Roche, CA, United States). Metabolome was carried out by using liquid chromatography–mass spectrometry.

The determination of the non-volatile metabolome of tea petals was described as earlier (Zhou et al., 2020). Briefly, 100 mg powder samples were extracted in 1.0 ml methanol (70%) at 4°C for 24 h, and 5 μl supernatant was injected into ultra-performance liquid chromatography (UPLC, Shimadzu Co., Kyoto, Japan) with a mass system (MS, Applied Biosystems 6500 Q TRAP, MA, United States). Metabolites were identified using MWDB (Metware Database, Metware Biotechnology Co., Ltd., Wuhan, China) and subject to the partial least squares (PLS) discriminant analysis. The significant dissimilarities of metabolites were set as the variable importance (VIP) ≥1 and the fold change ≥2 or ≤0.5.

Volatile Profile

Two grams of frozen samples were accurately weighed and heated in an air bath at 45°C for 5 min. Then, a manual injector with 65 μm polydimethylsiloxane/divinylbenzene extraction head was inserted to conduct headspace extraction under the condition of 45°C air bath. After 40 min of extraction, the manual injector was inserted into the injection port of the gas chromatograph–mass spectroscopy (GC–MS) immediately, and the instrument started to collect data at the same time. GC column: DB-5MS (30 m × 250 μm × 0.25 μm); injection port temperature: 250°C; carrier gas: helium (99.999%); flow rate: 1.6 ml/min; temperature program: 50°C for 3 min, then increasing to 265°C at the speed of 4.0°C/min, and keeping for another 5 min. MS ionization mode: EI; electron energy: −70 eV; quality scanning range: 33–600 U; ion source temperature: 220°C. The mass spectrum data obtained from GC–MS analysis were searched in NIST98.L standard spectrum library. The relevant mass spectrum data were checked, and the base peak, mass nucleus ratio, and relative abundance were analyzed. The structures and names of aromatic compounds represented by each peak were confirmed. The relative content of components is obtained by the ratio of the peak area of each aroma component to the total peak area.

Statistical Analysis

Pearson's correlation between the content of metabolites and the levels of gene expression was analyzed by using R (version 4.0.3). The correlation network was drawn by Cytoscape (v3.8.2). Other figures were presented by Excel 2010.

RESULTS

Volatile Benzenoid–Phenylpropanoids and Related Genes

The most abundant volatile BPs in both the white and purple tea flowers were MeSA and two phenylpropanoid volatiles, namely, AP and 1-PE (**Figure 2**, **Supplementary Figure 1**,

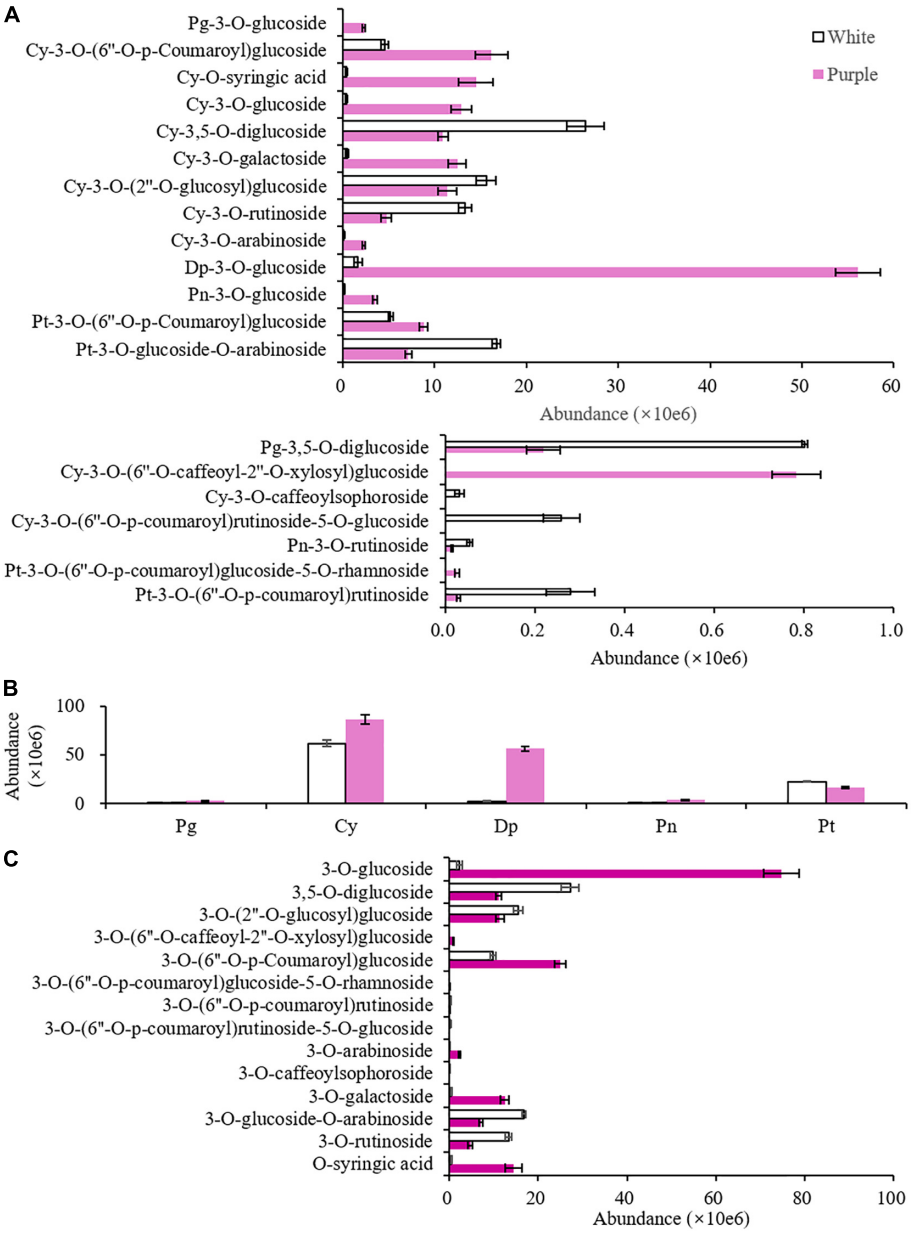


FIGURE 3 | Non-volatile benzenoid–phenylpropanoids (BPs) with significant difference ($p < 0.05$) between the white [Zijuan (ZJ) variety] and purple [Baitang (BT) variety] petals of tea plants. **(A)** Peak area of all anthocyanins. **(B)** Aggregation of the same aglycone. **(C)** Aggregation of the same glycosyl. Pg, pelargonidin; Cy, cyanidin; Dp, delphinidin; Pn, peonidin; Pt, petunidin.

and **Supplementary Table 1**). MeSA was the dominant volatile BPs in the white flowers of ZJ, while in the purple flowers of BT, it was AP, which was 6.2 times higher than that in ZJ. Except for MeSA, other common volatile BPs were significantly elevated in purple flowers: 1-PE, Bald, and Balc were 2, 2.9, and 2.8 times higher, respectively; and 2-PE was even not detected in white flowers. In terms of the total amounts, the dominant aroma from BPs was changed in the purple tea flowers, where the phenylpropanoids (AP, 1-PE, and 2-PE) were higher than other benzenoids (Bald, Balc, and MeSA), while in the normal whiter flowers, it was opposite. It indicated that the synthetic pathways of phenylpropanoids were promoted in the purple tea flowers.

Then, we investigated four enzyme genes responsible for the direct synthesis of benzyl benzoate, MeSA, 2-PE, and AP, which were benzyl alcohol *O*-benzoyltransferase (BPBT; KO id is K19861), salicylic acid carboxyl methyltransferase (SAMT; KO id is K21483), phenylacetaldehyde reductase (PAR), and short-chain dehydrogenase/reductase (SDR), respectively (Zhou et al., 2015; Cheng et al., 2016). They were all enhanced in the

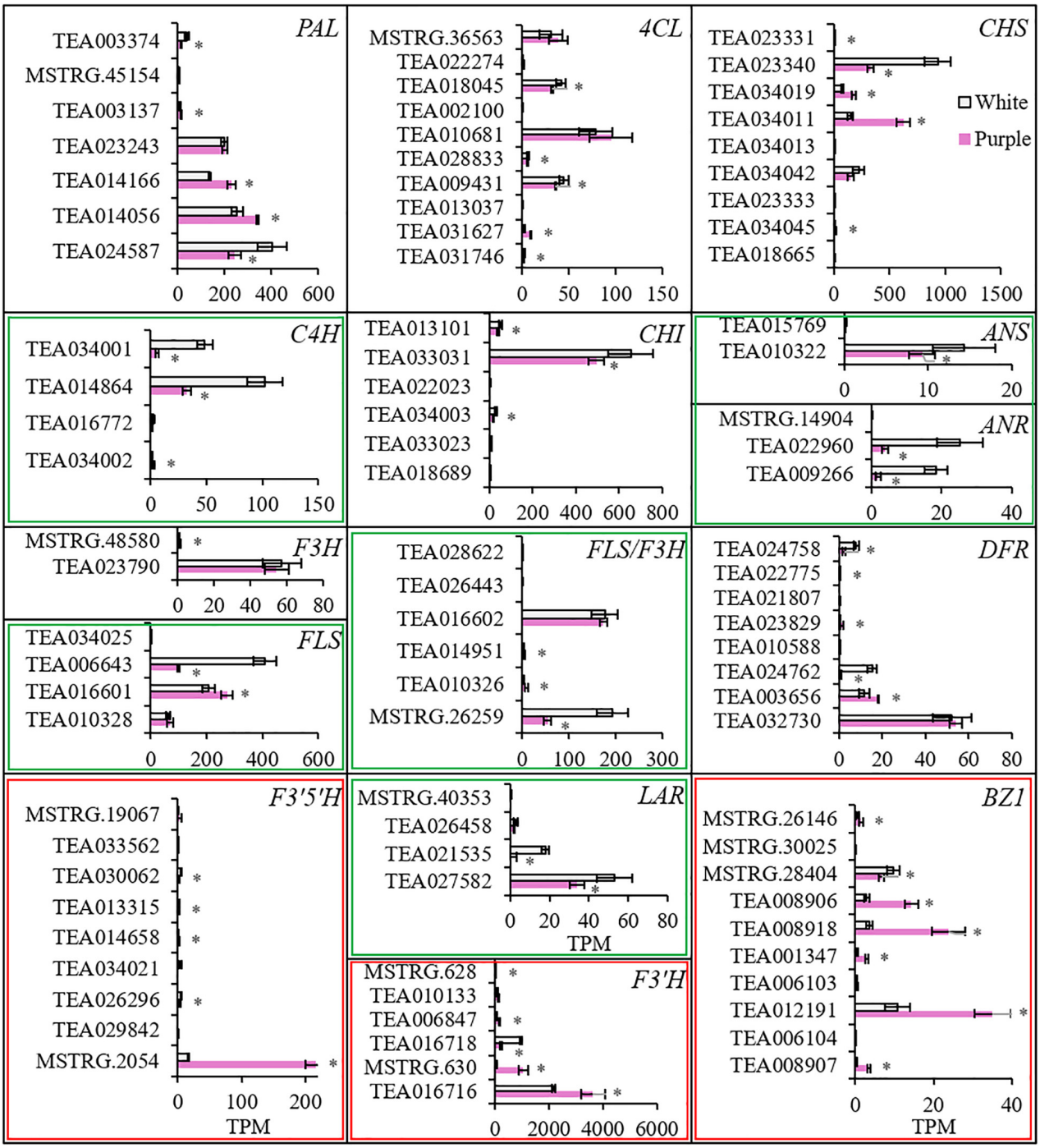


FIGURE 4 | The expression levels of genes involved in biosynthesis of anthocyanins. *Indicates significant difference ($p < 0.05$) between the white and the purple. PAL, phenylalanine ammonia lyase; 4CL, 4-coumarate-CoA ligase; C4H, cinnamate-4 hydroxylase; CH, chalcone synthase; CHI, chalcone isomerase; F3H, flavanone 3-hydroxylase; F3′H, flavonoid 3′-hydroxylase; F3′5′H, flavonoid 3′,5′-hydroxylase; FLS, flavonol synthase; DFR, dihydroflavonol 4-reductase; LAR, leucoanthocyanidin reductase; ANS, anthocyanidin synthase; ANR, anthocyanidin reductase; BZ1, bronze-1 (i.e., anthocyanidin 3-*O*-glucosyltransferase or UFGT).

purple petals. A stacked column was used to illustrate the 22 gene members of *SDR* with significant differences between the two varieties. The sum of the 22 *SDR* members was 1.3-fold higher in the purple petals and so was the member (TEA025281) with the highest expression level. *BPBT* and *PAR* had low expression levels but were 15 and 5 times higher, respectively. Two members of *SAMT* were highly induced in tea flowers, and they were 3–4 times higher in purple petals. These results

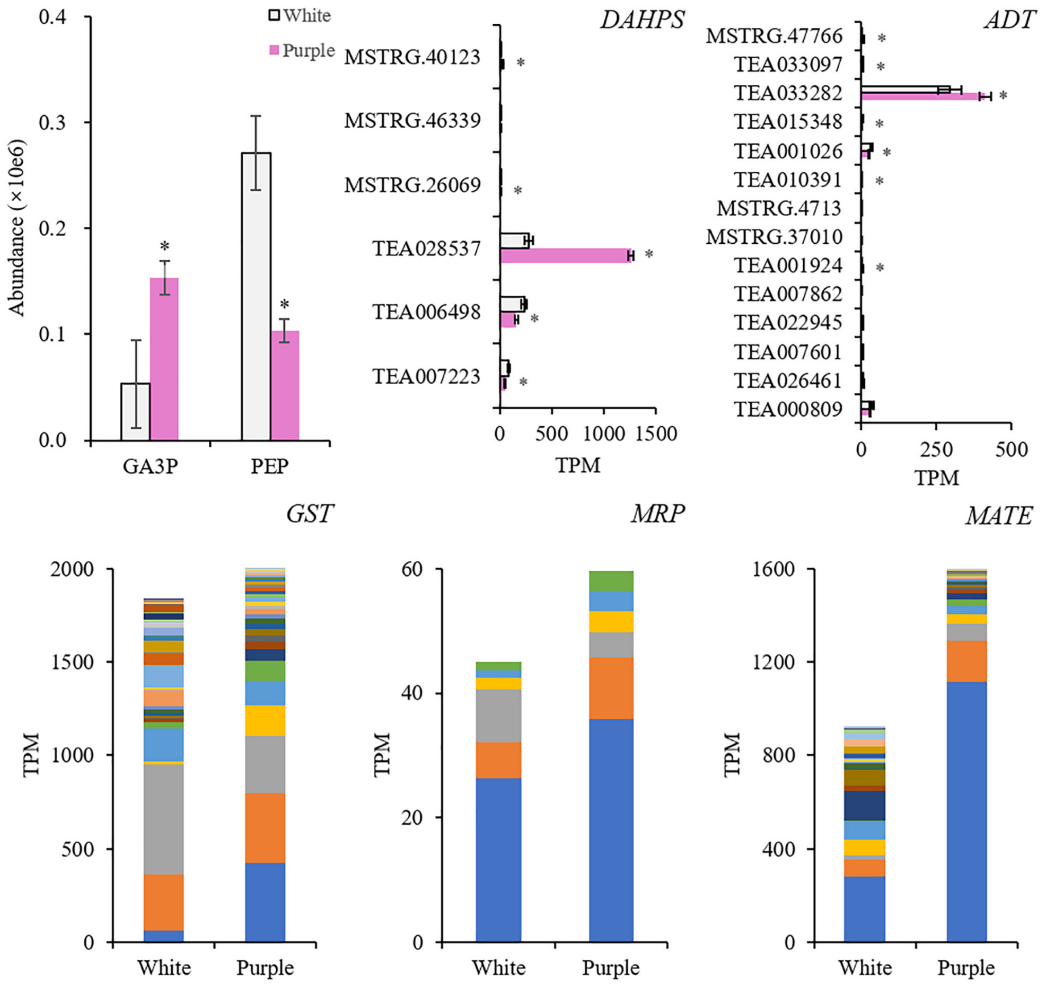


FIGURE 5 | Key compounds and genes in the upstream biosynthetic pathways of benzenoid–phenylpropanoids (BPs), and the genes related to the transportation of anthocyanins. *Indicates significant difference ($p < 0.05$) between the white and the purple. GA3P, glyceraldehyde 3-phosphate; PEP, phosphoenolpyruvate; DAHPS, 3-deoxy-D-arabino-heptulosonate 7-phosphate synthase; ADT, arogenate dehydratase/prephenate dehydratase; GST, glutathione *S*-transferase; MRP, multidrug resistance-associated protein; MATE, multidrug and toxin extrusion.

proved that the biosynthesis of volatile BPs was stronger in the purple tea flowers.

Non-volatile Benzenoid–Phenylpropanoids and Related Genes

A total of 20 anthocyanins were detected in the petals of the 2 varieties, and 5 anthocyanidins were found as aglycones, namely, Pg, Cy, Dp, Pn, and Pt. Mv did not exist in tea. Each variety had two or three undetected anthocyanins, namely, petunidin-3-*O*-(6″-*O*-*p*-coumaroyl)glucoside-5-*O*-rhamnoside, cyanidin-3-*O*-(6″-*O*-caffeoyl-2″-*O*-xylosyl)glucoside, and pelargonidin-3-*O*-glucoside in ZJ and cyanidin-3-*O*-(6″-*O*-*p*-coumaroyl)rutinoside-5-*O*-glucoside and cyanidin-3-*O*-caffeoylsophoroside in BT (**Figure 3A** and **Supplementary Table 2**). The total content of these 20 anthocyanins was 1.9 times higher in the purple petals than that in the white petals. According to the types of aglycones, 20 anthocyanins can be divided into 2 Pg, 11 Cy, 1 Dp, 2 Pn, and 4 Pt, of which Cy and Dp were the dominant aglycones (**Figure 3B**). Although Dp has only one glycosidic form, it was increased largely in the purple petals. The dominant anthocyanins in the purple petals were delphinidin-3-*O*-glucoside (DpG, Mirtillin), Cy-3-*O*-(6″-*O*-*p*-coumaroyl)glucoside, Cy-*O*-syringic acid, Cy-3-*O*-glucoside (Kuromanin), and Cy-3-*O*-galactoside, while in white flowers, they were Cy-3,5-*O*-diglucoside (Cyanin), Pt-3-*O*-glucoside-*O*-arabinoside, Cy-3-*O*-(2″-*O*-glucosyl)glucoside, and Cy-3-*O*-rutinoside (Keracyanin). Apart from aglycones, the other half of these 20 anthocyanins contained 13 glycosides and 1 organic acid, which was syringic acid, a member of benzoic acids (**Figure 3C**). In white petals of ZJ, the most anthocyanins were 3,5-*O*-diglucosides, and in purple petals of BT, 3-*O*-glucosides were extremely elevated. As a result, the biggest difference of anthocyanins between the two tea varieties was DpG, which might be the

TABLE 1 | The candidate transcription factors (TFs) involved in the biosynthesis of acetophenone (AP) and delphinidin-3-*O*-glucoside (DpG).

TF family	Gene ID	Expression level (TPM)		Correlation coefficient		Annotation
		ZJ	BT	AP	DpG	
AP2/ERF	TEA030964	9.9 ± 1.0	24.1 ± 3.2	0.962	0.965	APETALA2 (*Camellia sinensis*)
	TEA004341	4.4 ± 0.4	8.9 ± 1.3	0.979	0.960	ERF118-like (*Nicotiana sylvestris*)
bZIP	TEA032800	12.5 ± 1.2	19.8 ± 1.4	0.956	0.963	bZIP 18 (*Camellia sinensis*)
	TEA014041	0.7 ± 0.2	1.5 ± 0.2	0.968	0.952	TGA2.3 (*Vitis vinifera*)
EMB	MSTRG.32209	N.D.	1.6 ± 0.2	0.956	0.983	EMB1444-like (*Camellia sinensis*)
GATA	TEA027958	N.D.	0.4 ± 0.0	0.957	0.996	GATA (*Actinidia chinensis* var. *chinensis*)
	TEA022802	2.1 ± 0.8	13.5 ± 2.5	0.954	0.962	GATA 26-like (*Nicotiana tabacum*)
HS	TEA000588	1.1 ± 0.0	1.8 ± 0.2	0.971	0.953	Heat stress TF B-3-like (*Quercus suber*)
MYB	TEA024999	0.9 ± 0.2	2.5 ± 0.3	0.970	0.972	PHL5 (*Vitis vinifera*)
	MSTRG.43318	4.4 ± 0.6	15.3 ± 1.7	0.969	0.983	ETC3 (*Camellia sinensis*)
	TEA002399	7.7 ± 0.6	26.7 ± 1.6	0.951	0.997	MYB61 (*Arabidopsis thaliana*)
NAC	MSTRG.28161	N.D.	5.0 ± 1.2	0.987	0.973	NAC 29-like (*Camellia sinensis*)
TCP	TEA009154	3.2 ± 1.2	15.5 ± 2.6	0.962	0.976	TCP4 (*Glycine max*)
WRKY	TEA012360	0.2 ± 0.1	1.3 ± 0.3	0.984	0.959	WRKY 40 (*Actinidia chinensis* var. *chinensis*)

main reason for the purple color of the BT flowers, and the biosynthesis of Dp and glucose had probably been enhanced in the purple petals.

To investigate the enhanced biosynthesis progresses for producing anthocyanins, the transcriptomes of petals were sequenced for ZJ and BT. The expression profiles of genes involved in flavonoid biosynthesis pathway showed that each gene had several members, and some were higher in white petals and others were converse (**Figure 4** and **Supplementary Table 3**). However, the sum of the expression of members with significant differences in each gene indicated that most genes in purple petals were downregulated, especially *C4H*, *FLS*, *LAR*, and *ANR*, which were largely decreased. But *F3′H* and *F3′5′H* were significantly upregulated. In particularly, *F3′5′H* was elevated by eight times, which proved its critical role in generating Dp. Moreover, due to the little amount of *ANR*, Dp could not be reduced to epigallocatechin and thereby fluxed into the synthesis of DpG, which was catalyzed by bronze-1 (BZ1). BZ1 is an anthocyanidin 3-*O*-glucosyltransferase, also known as UFGT. The total expression level of *BZ1* gene members was three times higher in the purple petals than that in the white.

To further reveal the different synthetic mechanisms of BPs between the two tea varieties, the upstream compounds and genes were detected (**Figure 5**). In purple petals, glyceraldehyde 3-phosphate (GA3P) increased 2.9 times, while phosphoenolpyruvate (PEP) was only 38% of that in the white petals. It indicated that the carbon metabolism in the purple petals was promoted, but it might not flux into the downstream metabolism of Phe. The decrease of PEP might be due to the highly induced enzyme 3-deoxy-D-arabino-heptulosonate 7-phosphate synthase (DAHPS). The main member of *DAHPS* gene TEA028537 was 4.6-fold higher in the purple petals. Arogenate dehydratase/prephenate dehydratase (ADT) catalyzes arogenate to produce Phe. Its main member TEA033282 was elevated a little in purple petals, which was 1.3-fold higher. In addition, synthesized anthocyanins need to be transported to vacuole before displaying color. This progress is related to glutathione *S*-transferase (GST), multidrug resistance-associated protein (MRP), and multidrug and toxin extrusion (MATE). The expression levels of these gene members with significant differences were added up, and the total amount was all higher in the purple petals, which were 1.1, 1.3, and 1.7 times higher, respectively. The main members (TEA015341, TEA013531, and TEA006958) of these three genes in the purple petals were induced 6.9, 1.4, and 3.9 times higher, respectively.

Transcription Factors Involved in the Biosynthesis of Benzenoid–Phenylpropanoids

The results above suggested that in purple flowers, the most characteristic volatile and non-volatile BPs were AP and DpG, respectively. The Pearson's correlation coefficients were calculated between these two BPs and the TFs, which were significantly different between the two tea varieties. There were 14 TFs whose correlation coefficients were greater than 0.95 both with AP and DpG (**Table 1**). They belong to the families of *AP2/ERF*, *bZIP*, *EMP*, *GATA*, *HS*, *MYB*, *NAC*, *TCP*, and *WRKY* and were 1.6–6.7 times higher in the purple petals. *MYB*, *AP2/ERF*, *bZIP*, *TCP*, and *GATA* had the most expression levels.

Correlations between the 14 TFs and the structural genes narrated above were analyzed to discover the potential regulatory mechanisms (**Figure 6**). Most of these TFs are highly related to *PAR*, *SAMT*, and *SDR*, which contributed to the biosynthesis of volatile BPs; and *4CL*, *BZ1*, *CHS*, *F3′H*, *F3′5′H*, *FLS*, and *LAR*, which were responsible for non-volatile BPs; and *ADT* and *DAHPS* in the upstream biosynthesis of Phe. By contrast, these TFs did not focus on a specific gene for anthocyanin transportation, neither on BPBT in the biosynthesis of benzyl

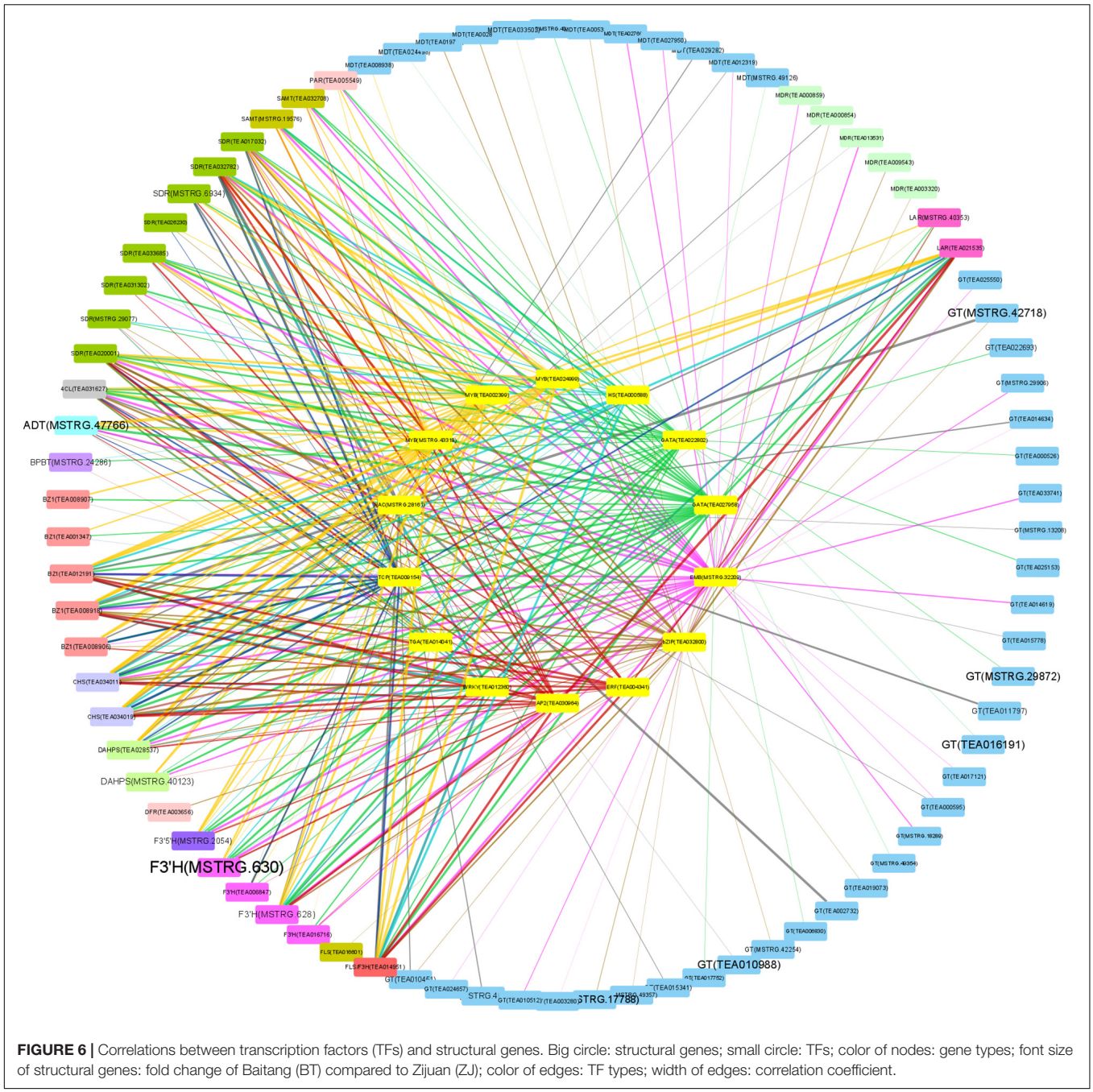


FIGURE 6 | Correlations between transcription factors (TFs) and structural genes. Big circle: structural genes; small circle: TFs; color of nodes: gene types; font size of structural genes: fold change of Baitang (BT) compared to Zijuan (ZJ); color of edges: TF types; width of edges: correlation coefficient.

benzoate and *PAL*, *C4H*, *CHI*, *ANS*, *ANR*, *F3H*, and *DFR* in the pathway of flavonoid biosynthesis.

Validation of Gene Expression Levels by Real-Time PCR

Some gene members with high expression levels determined by RNA-seq were selected (**Supplementary Table 4**) to be further validated by real-time PCR. The expression profiles of the two methods were consistent, and the correlation coefficient between $2^{-\Delta\ ct}$ and TPM was 0.845, showing a good correlation (**Figure 7**). These results indicated that the transcriptome data were credible.

DISCUSSION

The floral color of common *Camellia*, such as *C. reticulata* and *C. sasanqua*, ranges from white to red. However, the flowers of tea have been observed to be white in the past (Chen et al., 2020), though it has purple leaves. Exceptionally, a new mutation with purple flowers was found recently, which provides an

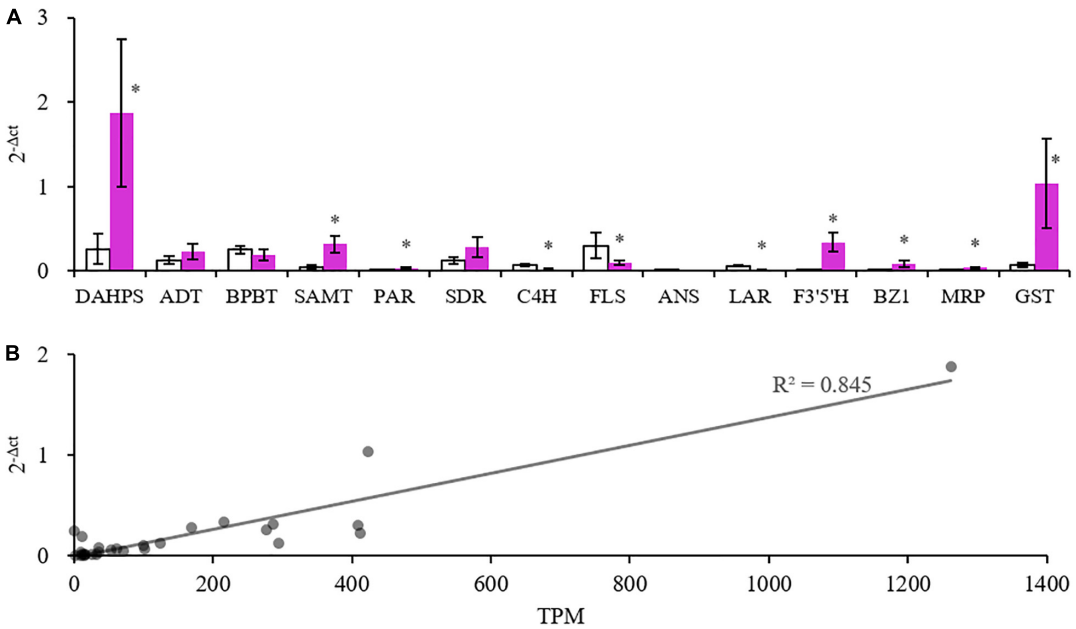


FIGURE 7 | Quantitative real-time polymerase chain reaction (PCR) (qRT-PCR) validation of some selected genes. **(A)** Expression values of genes. *Indicates significant difference ($p < 0.05$) between the white and purple petals. **(B)** Correlation of qRT-PCR results and Transcripts Per Kilobase of exon model per Million mapped reads (TPM) values of the selected genes.

opportunity for us to study the color formation in tea plant. The purple color is from the pigment of anthocyanins (Liu et al., 2016). Furthermore, floral scents are different from the aroma emitting from tea leaves, such as AP, a major fragrance in tea flowers (Zhou et al., 2015). Because anthocyanins and AP are all generated from Phe metabolism, it attracted us to find out whether the occurrence of purple color leads to the attenuation of aroma in the related pathways.

Carbohydrate metabolism was extremely important for the biosynthesis of flavonoids. The production of PAL and flavonoids in strawberry leaves depended on a supply of carbon dioxide (Creasy, 1968). Herein, the increase of GA3P in purple petals indicated the stimulation of carbon flux exchange. The induction of *DAHPS* and *ADT* further suggested that the shikimate pathway had been directed to produce Phe. The shikimate pathway links the primary metabolism and the downstream biosynthesis of BPs (Oliva et al., 2017). However, the decrease of PEP in the upstream pathway and the downregulation of genes in the main route of phenylpropanoid and flavonoid biosynthesis showed no obvious evidence that a large quantity of carbon flux had entered the pathway to produce BPs. The stimulated carbon metabolism might mainly contribute to gluconeogenesis and pentose phosphate pathway to synthesize the glycosyl moiety of anthocyanins. In flavonoid biosynthesis pathway, only *F3′H* and *F3′5′H* were elevated to increase the downstream Cy and Dp. In tea flowers, whether in white or purple color, Cy was conjugated to several carbohydrates, while Dp was only conjugated to glucoside. But the contents of total Cy and its glycosides in white petals were close to that in the purple petals. Thus, as *BZ1* was induced, Dp-3-*O*-glucoside became the mainly increased and dominant anthocyanin in purple petals. Although the carbon flux might not increase, it seems that anthocyanins and flavonoids are always in different channels. The carbon flux in dark-exposed *Begonia semperflorens* channeled into flavonoids from anthocyanins (Zhang et al., 2016). We also found the inhibition of *FLS*, *LAR*, and *ANR* in the branching way to produce flavones and flavanols (catechins). Taken together, anthocyanins in purple petals were probably due to the shift of carbon flux from flavones and flavanols but little from upstream pathways (**Figure 8**). The promoted upstream pathway, i.e., the shikimate pathway, might flux into other aromatic derivatives, such as volatile BPs. Although *F′3′5′H* and *BZ1* had critical effects on the formation of anthocyanins, the highly expressed gene members were different from the identified members in tea leaf (Zhao et al., 2017; Su et al., 2018; Guo et al., 2019).

As for the volatile BPs, the contents of aroma and related genes were increased in the purple petals, except for a few indexes (e.g., MeSA) with no significant difference. It means that the biosynthesis pathways of volatile BPs were generally promoted but not reduced with the increasing purple color as we thought before. It was similar in purple cherry tomato, where the types of volatile compounds were less than that in red and orange tomatoes, but phenolic volatiles increased (Silva Souza et al., 2020). Whether white tea flowers smell better needs a more comprehensive investigation. In tea leaves, single blue or red light-induced *PAL* gene and volatile benzenoids but not volatile phenylpropanoids (Fu et al., 2015). It seemed that *PAL* might only affect benzenoid aroma. However, the findings of this study showed that BPs increased in purple flowers despite the downregulation of *PAL*, which means *PAL* may not be a

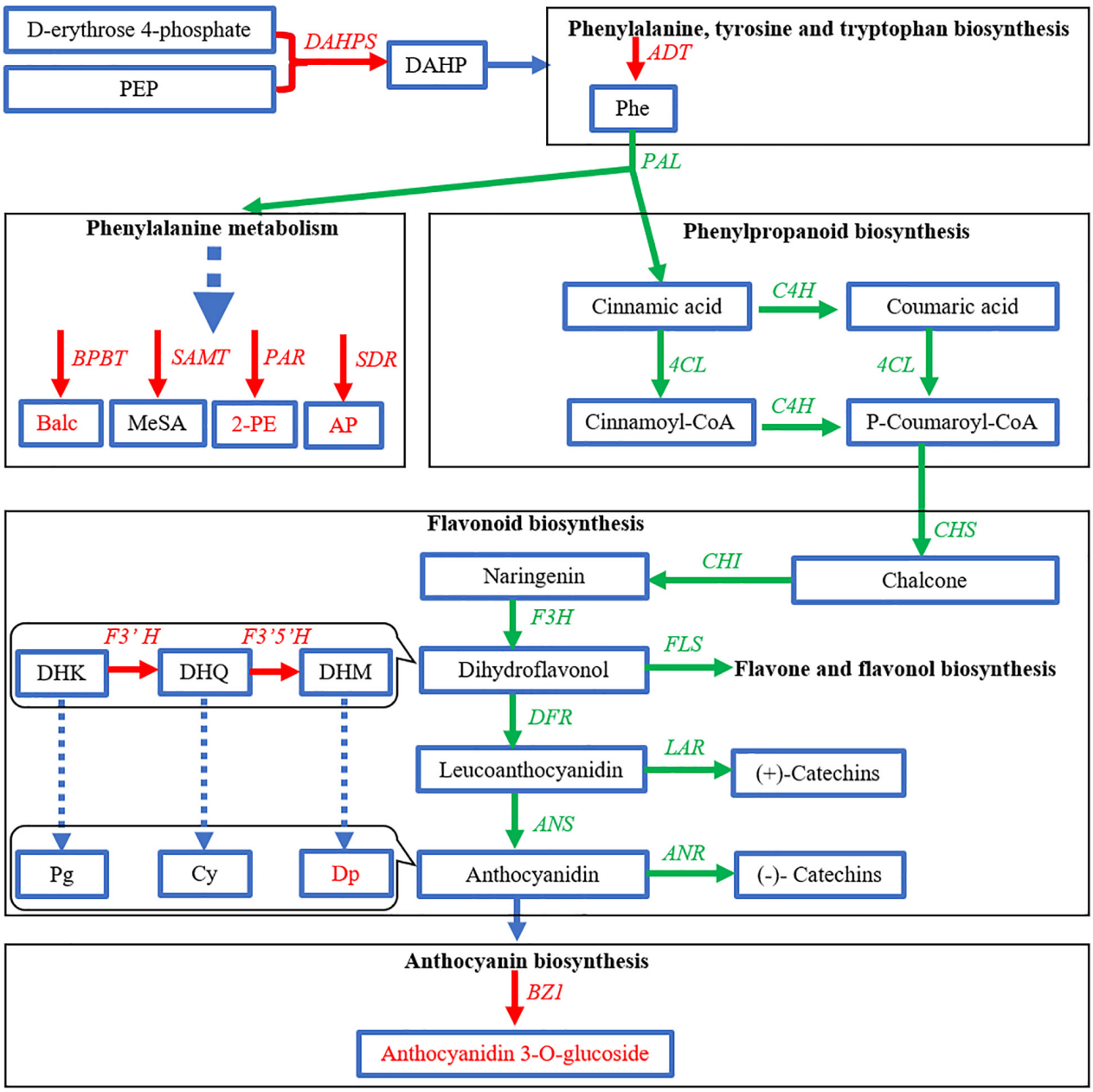


FIGURE 8 | Expression profile of genes related to biosynthesis of benzenoid–phenylpropanoids (BPs) in purple petals [Baitang (BT) variety] compared to that in white petals [Zijuan (ZJ) variety]. Green and red arrows indicate lower and higher in the purple petals, respectively.

decisive factor. Volatile phenylpropanoids are special aroma in tea flower. In particular, 1-PE and AP are highly accumulated in tea flower but not in tea leaf and increase during floral development (Chen et al., 2018). In the white petals of ZJ variety, 1-PE and AP were indeed the main phenylpropanoid aroma. However, they were surprisingly even much higher in the purple petals of BT variety, and AP became the highest volatile BPs, which obviously improved the aroma characteristics of tea flower. So far, the formation mechanism of AP has been not very clear. The increased 1-PE was probably the precursor/metabolite of AP, and SDR might be the potential enzyme (Zhou et al., 2015), but its members did not show as much change as AP. The reason for the increase of AP in purple tea flowers is worth further investigation.

Furthermore, from the results, we could see that anthocyanins also existed in the white petals. Except for the concentration, the color was affected by pH and anthocyanin transportation into vacuole. Three types of mediation have been proposed for anthocyanin transport: GST, membrane transporter, and membrane vesicle (Zhao et al., 2020). GST sequestrates anthocyanin from cytosol to the vacuole. *PpGST1* correlated well

with anthocyanin accumulation in peach fruit (Zhao et al., 2020). The second mechanism relies on MATE transporters and ABC transporters located in the tonoplast. *ZmMRP3* was first reported as an anthocyanin transporter and controlled by the regulators of anthocyanin biosynthesis (Goodman et al., 2004). The flower color of Asiatic Hybrid Lilies (*Lilium* spp.) is half purple (lower part) and half white (upper part), which is similar to BT flowers. An MATE gene *LhDTX35* might function as a carrier protein to transport anthocyanins (Xu et al., 2020). Furthermore, a flavonoid carrier similar to mammalian bilitranslocase (BTL) was identified in grape, and it may participate in long distance transport of flavonoids for its presence in vascular bundles (Petrussa et al., 2013). Thus, the transportation of anthocyanins is related to GST, MRP, MATE, and BTL-homolog. Herein, in tea flowers, GST, MATE, and MRP have been detected. Their main members and the total amount were higher in the purple petals, which indicated a promotion of anthocyanin transportation. Flavonoids in tea flowers may not be transported from tea leaves, since BTL-homolog was not detected in the flowers of both tea varieties. Many genes involved in the synthesis of flavonoids have been stimulated during the tea flower development (Chen et al., 2020). However, more experiments are needed to validate the sources of flavonoids in tea flower.

Since the biosynthetic pathways of volatile and non-volatile BPs are not fully understood, it is difficult to find a key enzyme or its gene as a hub gene controlling the synthetic direction of BPs for pigment or aroma. Nevertheless, we may get some clues from the analysis of TFs. In this study, we found 14 TFs were highly related to AP and DpG, of which *MYB*, *AP2/ERF*, *bZIP*, *TCP*, and *GATA* had relatively higher expression levels. They all play important roles in elevation of aroma and pigment. *SlMYB75* promoted accumulation of anthocyanin and volatile aroma production in tomato fruits (Jian et al., 2019). *AP2/ERF* regulates the expression of ethylene responsive genes in plant development and flower abscission (Gao et al., 2020). *bZIP* affected anthocyanin accumulation in apple and aroma in banana. *FvTCP9* dramatically promotes the genes involved in fruit color and aroma metabolism in strawberry fruits (Wei et al., 2016). *GATA* promoted crucial genes of Ehrlich pathway to enhance 2-PE production in *Saccharomyces cerevisiae* (Chen et al., 2017) and regulate anthocyanin accumulation in *Camellia japonica* petals (Fu et al., 2021). The 14 TFs herein were also highly related to the genes in the biosynthesis of Phe and BPs. This is in accordance with the fact that these TFs were highly related to the content of AP and DpG. These TFs focused on the genes for synthesis of AP (*PAR* and *SDR*), Dp (*F3′H* and *F3′5′H*), and DpG (*BZ1*) but inhibited the formation of flavones (*FLS*) and catechins (*LAR*). This indicated that in spite of the promotion of the upstream Phe synthesis (*ADT* and *DAHPS*), these TFs only changed the synthetic direction to Dp and blocked other flavonoids in the pathway of flavonoid biosynthesis.

CONCLUSION

It is an interesting thing to know what happens to the aroma when the flower color mutates naturally. In the purple petals of BT tea flowers, the BP-type aroma increased as the BP-type color occurred. AP and DpG were largely accumulated and become the dominant volatile and non-volatile BPs. The structural genes in the shikimate pathway and the biosynthesis of volatile BPs were promoted. Most genes in flavonoid biosynthesis were downregulated, and the efflux was directed to accumulate Dp. Many TFs were involved in regulating the purple color and its related aroma. The outcome of this study revealed the relationship between the BP-type color and aroma in tea plant. To further clarify the regulatory mechanism, more identification of the functions of relevant genes will be needed in the future.

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: https://ngdc.cncb.ac.cn/gsa/browse/CRA005021, PRJCA006284.

AUTHOR CONTRIBUTIONS

LZ and XM designed the research. XM wrote the manuscript. SW, LZ, and XM collected the experimental materials. SW, CL, and LH conducted the experiment involved in this manuscript. CL, CZ, and CD helped with data analysis. LZ revised the manuscript critically for important intellectual content. All authors listed here contributed and approved the manuscript.

FUNDING

National Key Research & Development Program (2018YFD1000601); National Natural Science Foundation of China (32072628); Qiannan Normal University for Nationalities (2020qnsyrc08); and Fund for Less Developed Regions of the National Natural Science Foundation of China (32160727).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2021.762330/full#supplementary-material

Supplementary Figure 1 | Total ion chromatogram (TIC) of tea flower from gas chromatography-mass spectroscopy (GC–MS).

REFERENCES

Chandra Singh, M., Kelso, C., Price, W. E., and Probst, Y. (2020). Validated liquid chromatography separation methods for identification and quantification of anthocyanins in fruit and vegetables: a systematic review. *Food Res. Int.* 138:109754. doi: 10.1016/j.foodres.2020.109754

Chen, D., Chen, G., Sun, Y., Zeng, X., and Ye, H. (2020). Physiological genetics, chemical composition, health benefits and toxicology of tea (*Camellia sinensis*

L.) flower: a review. *Food Res. Int.* 137:109584. doi: 10.1016/j.foodres.2020.109584

Chen, X., Wang, Z., Guo, X., Liu, S., and He, X. (2017). Regulation of general amino acid permeases Gap1p, GATA transcription factors Gln3p and Gat1p on 2-phenylethanol biosynthesis via Ehrlich pathway. *J. Biotechnol.* 242, 83–91. doi: 10.1016/j.jbiotec.2016.11.028

Chen, Y., Zhou, Y., Zeng, L., Dong, F., Tu, Y., and Yang, Z. (2018). Occurrence of functional molecules in the flowers of tea (*Camellia sinensis*) plants: evidence for a second resource. *Molecules* 23:790. doi: 10.3390/molecules23040790

Cheng, S., Fu, X., Mei, X., Zhou, Y., Du, B., Watanabe, N., et al. (2016). Regulation of biosynthesis and emission of volatile phenylpropanoids/benzenoids in petunia× hybrida flowers by multi-factors of circadian clock, light, and temperature. *Plant Physiol. Biochem.* 107, 1–8. doi: 10.1016/j.plaphy.2016.05.026

Creasy, L. L. (1968). The significance of carbohydrate metabolism in flavonoid synthesis in strawberry leaf disks. *Phytochemistry* 7, 1743–1749. doi: 10.1016/S0031-9422(00)86645-5

Fu, M., Yang, X., Zheng, J., Wang, L., Yang, X., Tu, Y., et al. (2021). Unraveling the regulatory mechanism of color diversity in *Camellia japonica* petals by integrative transcriptome and metabolome analysis. *Front. Plant Sci.* 12:685136. doi: 10.3389/fpls.2021.685136

Fu, X., Chen, Y., Mei, X., Katsuno, T., Kobayashi, E., Dong, F., et al. (2015). Regulation of formation of volatile compounds of tea (*Camellia sinensis*) leaves by single light wavelength. *Sci. Rep.* 5:16858. doi: 10.1038/srep16858

Gao, J., Zhang, Y., Li, Z., and Liu, M. (2020). Role of ethylene response factors (ERFs) in fruit ripening. *Food Qual. Saf.* 4, 15–20. doi: 10.1093/fqsafe/fyz042

Goodman, C. D., Casati, P., and Walbot, V. (2004). A multidrug resistance–associated protein involved in anthocyanin transport in Zea mays. *Plant Cell* 16, 1812–1826. doi: 10.1105/tpc.022574

Guo, L., Gao, L., Ma, X., Guo, F., Ruan, H., Bao, Y., et al. (2019). Functional analysis of flavonoid 3'-hydroxylase and flavonoid 3',5'-hydroxylases from tea plant (*Camellia sinensis*), involved in the B-ring hydroxylation of flavonoids. *Gene* 717:144046. doi: 10.1016/j.gene.2019.144046

He, C., Guo, X., Yang, Y., Xie, Y., Ju, F., and Guo, W. (2016). Characterization of the aromatic profile in "zijuan" and "pu-erh" green teas by headspace solid-phase microextraction coupled with GC-O and GC-MS. *Anal. Methods* 8, 4727–4735. doi: 10.1039/C6AY00700G

Jian, W., Cao, H., Yuan, S., Liu, Y., Lu, J., Lu, W., et al. (2019). SlMYB75, an MYB-type transcription factor, promotes anthocyanin accumulation and enhances volatile aroma production in tomato fruits. *Hortic. Res.* 6:22. doi: 10.1038/s41438-018-0098-y

Liu, L., Zhang, L.-Y., Wang, S.-L., and Niu, X.-Y. (2016). Analysis of anthocyanins and flavonols in petals of 10 *Rhododendron* species from the Sygera mountains in Southeast Tibet. *Plant Physiol. Biochem.* 104, 250–256. doi: 10.1016/j.plaphy.2016.03.036

Oliva, M., Bar, E., Ovadia, R., Perl, A., Galili, G., Lewinsohn, E., et al. (2017). Phenylpyruvate contributes to the synthesis of fragrant benzenoid–phenylpropanoids in Petunia × hybrida flowers. *Front. Plant Sci.* 8:769. doi: 10.3389/fpls.2017.00769

Petrussa, E., Braidot, E., Zancani, M., Peresson, C., Bertolini, A., Patui, S., et al. (2013). Plant flavonoids—biosynthesis, transport and involvement in stress responses. *Int. J. Mol. Sci.* 14, 14950–14973. doi: 10.3390/ijms140714950

Silva Souza, M. A., Peres, L. E., Freschi, J. R., Purgatto, E., Lajolo, F. M., and Hassimotto, N. M. (2020). Changes in flavonoid and carotenoid profiles alter volatile organic compounds in purple and orange cherry tomatoes obtained by allele introgression. *J. Sci. Food Agric.* 100, 1662–1670. doi: 10.1002/jsfa.10180

Stevenson, P. C., and Aslam, S. N. (2006). The chemistry of the genus *Cicer L. Stud. Nat. Prod. Chem.* 33, 905–956. doi: 10.1016/S1572-5995(06)80043-8

Su, X., Wang, W., Xia, T., Gao, L., Shen, G., and Pang, Y. (2018). Characterization of a heat responsive UDP: flavonoid glucosyltransferase gene in tea plant (*Camellia sinensis*). *PLoS One* 13:e0207212. doi: 10.1371/journal.pone.0207212

Wei, C., Yang, H., Wang, S., Zhao, J., Liu, C., Gao, L., et al. (2018). Draft genome sequence of *Camellia sinensis* var. sinensis provides insights into the evolution of the tea genome and tea quality. *Proc. Natl. Acad. Sci. U.S.A.* 115, E4151–E4158. doi: 10.1073/pnas.1719622115

Wei, W., Hu, Y., Cui, M.-Y., Han, Y.-T., Gao, K., and Feng, J.-Y. (2016). Identification and transcript analysis of the TCP transcription factors in the diploid woodland strawberry *Fragaria vesca*. *Front. Plant Sci.* 7:1937. doi: 10.3389/fpls.2016.01937

Xia, E., Li, F., Tong, W., Li, P., Wu, Q., Zhao, H., et al. (2019). Tea plant information archive: a comprehensive genomics and bioinformatics platform for tea plant. *Plant Biotechnol. J.* 17, 1938–1953. doi: 10.1111/pbi.13111

Xu, H., Yang, P., Cao, Y., Tang, Y., He, G., Xu, L., et al. (2020). Cloning and functional characterization of a flavonoid transport-related MATE gene in asiatic hybrid lilies (*Lilium* spp.). *Genes* 11:418. doi: 10.3390/genes11040418

Zhang, K. M., Guo, M. L., He, D., Wu, R. H., and Li, Y. H. (2016). The Inhibition effect and excessive carbon flux resulting from blocking anthocyanin biosynthesis under darkness in *Begonia semperflorens*. *J. Plant Growth Regul.* 35, 22–30. doi: 10.1007/s00344-015-9503-z

Zhao, X., Wang, P., Li, M., Wang, Y., Jiang, X., Cui, L., et al. (2017). Functional characterization of a new tea (*Camellia sinensis*) flavonoid glycosyltransferase. *J. Agric. Food Chem.* 65, 2074–2083. doi: 10.1021/acs.jafc.6b05619

Zhao, Y., Dong, W., Zhu, Y., Allan, A. C., Lin-Wang, K., and Xu, C. (2020). PpGST1, an anthocyanin-related glutathione S-transferase gene, is essential for fruit coloration in peach. *Plant Biotechnol. J.* 18, 1284–1295. doi: 10.1111/pbi.13291

Zhou, C., Mei, X., Rothenberg, D. O., Yang, Z., Zhang, W., Wan, S., et al. (2020). Metabolome and transcriptome analysis reveals putative genes involved in anthocyanin accumulation and coloration in white and pink tea (*Camellia sinensis*) flower. *Molecules* 25:190. doi: 10.3390/molecules25010190

Zhou, Y., Zhang, L., Gui, J., Dong, F., Cheng, S., Mei, X., et al. (2015). Molecular cloning and characterization of a short-chain dehydrogenase showing activity with volatile compounds isolated from *Camellia sinensis*. *Plant Mol. Biol. Rep.* 33, 253–263. doi: 10.1007/s11105-014-0751-z

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.

frontiers in Plant Science

ORIGINAL RESEARCH
published: 06 December 2021
doi: 10.3389/fpls.2021.697212

Identification of a Strong Anthocyanin Activator, *VbMYBA*, From Berries of *Vaccinium bracteatum* Thunb.

*Ya-Ling Zhang[1], Kui Lin-Wang[2], Nick W. Albert[2], Caitlin Elborough[2], Richard V. Espley[2], Christelle M. Andre[2] and Zhi-Zhen Fang[1]**

[1] Fruit Research Institute, Fujian Academy of Agricultural Sciences, Fuzhou, China, [2] The New Zealand Institute for Plant and Food Research Limited, Mt Albert Research Centre, Auckland, New Zealand

OPEN ACCESS

Edited by:
Lourdes Gómez-Gómez, University of Castilla-La Mancha, Spain

Reviewed by:
Ivana Tomaz, University of Zagreb, Croatia
Kaili Chen, Northwest A&F University, China

***Correspondence:**
Zhi-Zhen Fang
Fzhzh2008@163.com

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

Received: *04 May 2021*
Accepted: *09 November 2021*
Published: *06 December 2021*

Citation:
Zhang Y-L, Lin-Wang K, Albert NW, Elborough C, Espley RV, Andre CM and Fang Z-Z (2021) Identification of a Strong Anthocyanin Activator, VbMYBA, From Berries of Vaccinium bracteatum Thunb. Front. Plant Sci. 12:697212. doi: 10.3389/fpls.2021.697212

Wufanshu (*Vaccinium bracteatum* Thunb.), which is a wild member of the genus *Vaccinium*, accumulates high concentration of anthocyanin in its berries. In this study, the accumulated anthocyanins and their derivatives in Wufanshu berries were identified through UHPLC–MS/MS analysis. Candidate anthocyanin biosynthetic genes were identified from the transcriptome of Wufanshu berries. qRT-PCR analyses showed that the expression of anthocyanin structural genes correlated with anthocyanin accumulation in berries. The R2R3-MYB, *VbMYBA*, which is a homolog of anthocyanin promoting R2R3-MYBs from other *Vaccinium* species, was also identified. Transient expression of *VbMYBA* in *Nicotiana tabacum* leaves confirmed its role as an anthocyanin regulator, and produced a higher anthocyanin concentration when compared with blueberry *VcMYBA* expression. Dual-luciferase assays further showed that VbMYBA can activate the *DFR* and *UFGT* promoters from other *Vaccinium* species. VbMYBA has an additional 23 aa at the N terminus compared with blueberry VcMYBA, but this was shown not to affect the ability to regulate anthocyanins. Taken together, our results provide important information on the molecular mechanisms responsible for the high anthocyanin content in Wufanshu berries.

Keywords: ***Vaccinium Bracteatum*** **Thunb., anthocyanin biosynthesis, transcriptome, VbMYBA, R2R3 MYB**

Abbreviations: 4CL, 4-coumaroyl:CoA-ligase; ANS, anthocyanin synthase; PAL, phenylalanine ammonia-lyase; C4H, cinnamate-4-hydroxylase; CHS, chalcone synthase; COG, Clusters of Orthologous Groups database; DAD, diode array detector; eggNOG, evolutionary genealogy of genes: Non-supervised Orthologous Groups database; ESI, electrospray ion; F3H, flavanone-3-hydroxylase; F3′H, flavonoid3′-hydroxylase; F3′5′H, flavonoid 3′,5′-hydroxylase; FPKM, fragments per kilobase per million mapped reads; DFR, dihydroflavonol-4-reductase; GST, glutathione S-transferase; HPLC, high performance liquid chromatography; KEGG, Kyoto Encyclopedia of Genes and Genomes; KOG, euKaryotic Ortholog Groups of proteins database; MBW, MYB-bHLH-WD40; MS, mass spectrometry; Nr, NCBI non-redundant protein database; Pfam, Pfam-protein family database; qRT-PCR, real-time quantitative RT-PCR; RNA-Seq, RNA-sequencing; Swiss-Prot, Swiss-Prot protein database; UFGT, UDPglucose:flavonoid-3-*O*-glucosyltransferase; GO, Gene Ontology database; UHPLC, ultra-high performance chromatography system.

INTRODUCTION

Wufanshu (*Vaccinium bracteatum* Thunb.) is a wild blueberry species belonging to the genus *Vaccinium*. It is an evergreen shrub or small tree widely distributed in hilly areas of China (Wang et al., 2013; Fan et al., 2019). In China, it is a traditional herbal medicine used to treat many diseases, such as inflammation, diarrhea, and skin eruptions (Wang et al., 2004; Fan et al., 2019; Zheng et al., 2019). Leaf juice of Wufanshu was used to stain and cook well-known local traditional food named Wu Fan in China (Zhang et al., 2014). Pigments derived from leaves of Wufanshu were also widely used to dye hair and clothing (Xu et al., 2017; Fan et al., 2020). Wufanshu leaves are rich in health-promoting compounds such as polysaccharides and flavonoids (Wang et al., 2013; Zhang et al., 2014; Fan et al., 2018). It has been reported that its extracts have health beneficial activities, such as tyrosinase inhibiting, anti-inflammatory, anti-fatigue, anti-diabetic and anti-cancer (Wang et al., 2010, 2013; Landa et al., 2014; Oh et al., 2018a,b; Fan et al., 2019). More particularly, the presence of secondary plant metabolites, such as anthocyanins (Lee et al., 2014), in Wufanshu berry extracts has been associated with its health-promoting activities such as anti-proliferative, anti-inflammatory and antidepressant-like activity (Landa et al., 2014; Oh et al., 2018a). The amount of anthocyanins in fruits of *Vaccinium* species varies between species and cultivars (Lohachoompol et al., 2008; Stevenson and Scalzo, 2012; Li et al., 2017). Cho et al. (2012) showed that the anthocyanin concentration in berries of *V. bracteatum* Thunb. was higher than that of *V. corymbosum*, but the basis for this is unknown.

Anthocyanins belong to the flavonoid class of polyphenolics and are produced by a complex biosynthetic pathway. This starts with the condensation of three molecules of malonyl CoA and one molecule of 4-coumaroyl CoA by phenylalanine ammonia-lyase (PAL), cinnamate-4-hydroxylase (C4H), 4-coumaroyl:CoA-ligase (4CL), chalcone synthase (CHS). The naringenin chalcone formed from this is then converted to naringenin flavanone by chalcone isomerase (CHI), and then hydroxylated by flavanone-3-hydroxylase (F3H), to form dihydrokaempferol. At this stage the pathway can branch to form different anthocyanins, such as cyanidin or delphinidin. Further hydroxyl groups are added at this stage by flavonoid3′-hydroxylase (F3′H) or flavonoid 3′, 5′-hydroxylase (F3′5′H) before conversion by dihydroflavonol-4-reductase (DFR). The formation of anthocyanidins (aglycones) is controlled by anthocyanidin synthase (ANS), and these are glycosylated by UDPglucose:flavonoid-3-*O*-glucosyltransferase (UFGT) to form anthocyanins (Holton and Cornish, 1995). Additional chemical modifications of the basic anthocyanin molecules with additional sugars, methyl and acyl groups can generate a large variety of anthocyanins that differ in color and chemical properties. Many of these biosynthetic steps are shared with proanthocyanidin (condensed tannin) biosynthesis, but the activity of UFGT is necessary for anthocyanin production. Thus, it is a key point of regulation in fruit species that have complex flavonoid profiles, such as grape and *Vaccinium* spp. (Kobayashi et al., 2002; Günther et al., 2020).

In *Vaccinium* species, such as blueberry, the most common anthocyanins are aglycones of cyanidin (cyanidin 3-galactoside, cyanidin 3-arabinoside and cyanidin 3-glucoside), peonidin (peonidin 3-glucoside, peonidin 3-galactoside and peonidin 3-arabinoside), delphinidin (delphinidin 3-galactoside, delphinidin 3-glucoside and delphinidin 3-arabinoside), malvidin (malvidin 3-arabinoside), and petunidin (petunidin 3-galactoside, petunidin 3-arabinoside and petunidin 3-glucoside) (Lohachoompol et al., 2008). The network that regulates anthocyanin biosynthesis in plants has been studied extensively (Jaakola, 2013). While all three components are necessary for anthocyanin regulation, the central component of the network that regulates anthocyanin biosynthesis is the MYB-bHLH-WD40 (MBW) complex (Nesi et al., 2001; Gonzalez et al., 2008; Albert et al., 2014) and R2R3-MYBs have been shown to be the key regulators across a range of plant species (Espley et al., 2007; Hichri et al., 2011; Albert et al., 2014; Feng et al., 2015; Tuan et al., 2015; Jin et al., 2016; Butelli et al., 2017; Plunkett et al., 2018; Zhang et al., 2018). This is because the MYB protein is usually the limiting component of the MBW complex, and because the MYB can activate the expression of bHLH and WDR genes through a widely conserved hierarchy (Albert et al., 2014, 2021). In *Vaccinium* species, R2R3-MYBs have been reported to be involved in anthocyanin accumulation (Nguyen et al., 2017; Plunkett et al., 2018; Die et al., 2020) and down regulation of their expression was associated with a reduction or loss of anthocyanin accumulation in berries (Primetta et al., 2015; Yang et al., 2018; Die et al., 2020). The regulators of anthocyanin biosynthesis of Wufanshu have not yet been identified. Understanding the regulatory mechanisms in these the highly anthocyanic berries will increase our knowledge of anthocyanin accumulation in the *Vaccinium* genus.

In this study, we report on the identification of *VbMYBA*, which is closely related to the R2R3-MYB anthocyanin activator *VcMYBA* from blueberries (*V. corymbosum* and *V. virgatum*) (Plunkett et al., 2018) and show that its anthocyanin promoting activity is stronger than *VcMYBA*. Our results will help the development of new *Vaccinium* cultivars that can accumulate more anthocyanins in their berries.

MATERIALS AND METHODS

Plant Materials

All berries at different stages were harvested from wild Wufanshu plants in Lianjiang, in the city of Fuzhou (Fujian Province, China). Green berries were harvested on June 19, 2016. Red and black berries were harvested on November 16, 2016 and divided into two groups according to color. We estimated that these berry stages approximately correspond with stages 5, 6, and 8, described by for blueberry (Zifkin et al., 2011). Berries sampled from three different plants were used as biological replicates. For each plant, berries with similar size and color were pooled together. After harvesting, all samples were transported to lab and immediately frozen in liquid nitrogen and kept at −80°C for further analysis.

Total Anthocyanin Content Analysis in Wufanshu Berries

Approximately 1 g (FW) of whole berries was ground to fine powder in liquid nitrogen and extracted with 10 mL extraction solution (0.05% HCl in methanol) at 4°C for 24 h. Anthocyanin content was quantified using pH differential method as described previously (Fang et al., 2016).

Anthocyanins Characterization of Wufanshu Berries

Extraction, identification and quantification of metabolites were carried out at Suzhou BioNovoGene (Suzhou, China). Approximately 0.2 g (FW) of the berry powder was transferred to a 2 mL centrifuge tube containing 1 mL of water: methanol: formic acid (66.5: 28.5: 5, v/v/v), incubate on ice for 30 min and centrifuged at 18,756 × *g* for 10 min. The tube containing the mixture was centrifuged at 12,000 rpm for 10 min. The supernatants were freeze-dried and the extracts were resuspended with 0.2 mL 5% formic acid. An Ultra-high performance chromatography system (UHPLC, Vanquish, Thermo Fisher Scientific, Waltham, MA, United States) with a Waters HSS T3 (2.1 × 100 mm, 1.8 μm) column was used. The parameters used were as follows: flow rate of 300 μL min^{-1}, column oven temperature of 40°C, and sample size of 2 μL. The mobile phases were 0.1% formic acid in water (A) and 0.1% formic acid in acetonitrile (B). The gradient elution procedure was as follows: 0-4.0 min, 10% A; 4.0-12.0 min, 10–60% A; 12.0–18.0 min, 60% A constant; 18.0–18.1 min, 60-10% A; 18.1–26.0 min, 10% A constant. Mass spectrometry (MS) analysis was carried out on a Thermo Q Exactive instrument (Thermo Fisher Scientific, Waltham, MA, United States) equipped with an electrospray ion (ESI) source in positive mode. The capillary voltage was set to 3000 V and the capillary temperature was set to 320°C. The pressure of sheath gas and auxiliary gas were set at 40 and 10 arbitrary units, respectively. Metabolites were detected using full-scan/MS2 mode with a resolution of 70000. ESI spectra were acquired through the information-dependent acquisition mode in Xcalibur 4.1 software (Thermo Fisher Scientific, Waltham, MA, United States). The dynamic exclusion time was set to 6s. For MS1, full MS spectra between 200 and 1500 mass-to-charge ratio (m/z) were recorded. MS/MS scans were recorded between 200 and 2000 m/z. Data dependent acquisition survey scans were acquired in 100 ms and the 10 most abundant product ion scans were collected. Each metabolite was confirmed based on their exact molecular weights, then the fragment information obtained according to the MS/MS mode was matched in database built by BioNovogene to identify metabolites. Searches were performed using the following settings: precursor ion m/z tolerance: ±10ppm; MS/MS m/z tolerance: ±20ppm.

RNA Extraction, cDNA Library Construction and Sequencing

For RNA-Seq, berries at green or black stage from three replicates were mixed and subjected to RNA extraction. Total RNA extraction, cDNA library construction, and sequencing were performed by staff at Beijing BioMarker Technologies (Beijing, China). RNA extraction was carried out using EZNA Plant RNA Kit (Omega Bio-tek) RNA purity was checked using the NanoPhotometer® spectrophotometer (IMPLEN, CA, United States). RNA concentration was measured using Qubit® RNA Assay Kit in Qubit 2.0 Fluorometer (Life Technologies, CA, United States). RNA integrity was assessed using the RNA Nano 6000 Assay Kit of the Agilent Bioanalyzer 2100 system (Agilent Technologies, CA, United States). cDNA library was generated using NEBNext®Ultra™ RNA Library Prep Kit for Illumina®(NEB, United States) following manufacturer's recommendation. Library quality was assessed using the Agilent Bioanalyzer 2100 system. Sequencing of the cDNA libraries were carried out on the Illumina HiSeq™ X Ten sequencing platform. The RNA-seq reads have been deposited in the NCBI Short Read Archive and are accessible under PRJNA694726.

De novo Assembly, UniGene Annotation and Expression Quantification of Unigenes

The raw reads were processed through in-house perl scripts and *de novo* assembled into unigenes with Trinity (Grabherr et al., 2011) as described previously (Zhang et al., 2018). Annotation of the assembled unigenes was performed using BLASTx (E-value<10–5) searches against eight public databases including the NCBI non-redundant protein database (Nr), Swiss-Prot protein database (Swiss-Prot), the Gene Ontology database (GO), the Clusters of Orthologous Groups database (COG), the euKaryotic Ortholog Groups of proteins database (KOG), the Kyoto Encyclopedia of Genes and Genomes (KEGG), the Pfam-protein family database (Pfam) and the evolutionary genealogy of genes: Non-supervised Orthologous Groups database (eggNOG). For calculation of gene expression level, clean RNA-Seq reads were mapped to the assembled unigenes using Bowtie (Langmead et al., 2009) and the expression levels of unigenes were calculated with fragments per kilobase per million mapped reads (FPKM) using RSEM (Li and Dewey, 2011).

qRT-PCR (Real-Time Quantitative RT-PCR) Analysis

Total RNA was extracted from green, red and black Wufanshu berries using EZNA Plant RNA Kit (Omega Bio-tek). First strand cDNA was prepared from 500 ng total RNA HiScript III RT SuperMix for qPCR with gDNA wiper (Vazyme, Nanjing, China). qRT-PCR was performed using the Eppendorf Realplex^4 real-time PCR system (Hamburg, Germany) in a total volume of 20 μL in each well containing 10 μL of 2 × ChamQ Universal SYBR qPCR Master Mix (Vazyme, Nanjing, China), 6 μL of cDNA (in 1:30 dilution), and 0.4 μL 10 μM primers. qRT-PCR conditions were 30 s at 95°C, followed by 40 cycles of 5 s at 95°C, 15 s at 60°C, followed by 60°C to 95°C melting curve detection. *Actin* gene (c87909.graph_c0) was used as the reference. The expression levels were calculated using the $2^{-\Delta\Delta CT}$ method. Three biological and four technical replications were performed. Primers for qRT-PCR were listed in **Supplementary Table 1**. Linear regression analysis of FPKM and qPCR was performed using Minitab® 18.

Cloning of *VbMYBA* and Prediction of Amino Acid Sequence

cDNA of black Wufanshu berries was generated using a First-Strand cDNA synthesis kit (Fermentas, United States) and used as template for gene cloning. The cDNA sequence encoding VbMYBA was isolated using I-5^{TM} 2 × High-Fidelity Master Mix (MCLAB, San Francisco, CA) with primers (forward 5′-GGCAGCTTACATGAAAATTCTCC-3′ and reverse 5′-CAAACAAAGAAATGCTTGCCG-3′) designed according to unigene sequence. The generated PCR products were purified and subsequently cloned into the pEASY-Blunt Zero vector using pEASY-Blunt Zero Cloning Kit (TransGen, Beijing, China) and sequenced. The open-reading frame of the sequence was predicted using GenBank ORF finder[1].

Multiple Sequence Alignment and Phylogenetic Analysis

Multiple sequence alignment of R2R3 MYB amino acid sequences from Wufanshu, blueberry, kiwifruit and peach was performed by CLUSTALW[2]. Shading of the alignment results was performed using ESPript 3.0 (Robert and Gouet, 2014). Phylogenetic tree was constructed using the maximum likelihood method with 1000 bootstrap replicates by MEGA-X. R2R3 MYB amino acid sequences used for phylogenetic analysis include *Arabidopsis thaliana* AtMYB4 (At4G38620), AtMYB11 (AT3G62610), AtMYB12 (AT2G47460), AtMYB75 (AT1G56650), AtMYB90 (AT1G66390), AtMYB111 (AT5G49330), AtMYB113 (AT1G66370), AtMYB114 (AT1G66380), AtMYB123 (AT5G35550); *Actinidia chinensis* AcMYB10 (PSS35990), and AcMYB110 (AHY00342); *Fragaria* × *ananassa* FaMYB1 (AAK84064.1), FaMYB9 (JQ989281), FaMYB10 (ABX79947.1); *Gossypium hirsutum* GhMYB1 (AAA33067.1), GhMYB6 (AAN28286.1) GhMYB10 (ABR01222.1), and GhMYB36 (AF336284); *Lotus japonicus* LjMYB12 (AB334529); *Myrciaria cauliflora* McMYB (MH383068); *Malus* × *domestica* MdMYB1 (XP_028963316.1), MdMYB9 (DQ267900), MdMYB10 (EU518249), MdMYB11 (DQ074463); *Medicago truncatula* MtLAP1 (ACN79541.1), and MtMYB14 (XP_003594801.1); *Prunus cerasifera* PcMYB10.1 (KP772281) and PcMYB10.2 (KP772282); *Petunia hybrida* PhMYB27 (AHX24372), PhAN2 (AB982128), PhDEEP PURPLE (ADQ00393.1), PhPURPLE HAZE (ADQ00388.1) and PhMYB4 (ADX33331.1); *Prunus persica* PpMYB10.1 (XM_007216468), PpMYB10.2 (XM_007216161) and PpMYB18 (KT159234); *Prunus salicina* PsMYB10.1 (MK105923), PsMYB10.2 (MK340932), and PsMYB18 (MK284223); PtMYB14 (ACR83705.1); *Pyrus pyrifolia* PyMYB10 (GU253310); *Solanum lycopersicum* SlMYB12 (ACB46530.1); *Trifolium arvense* TaMYB14 (AFJ53053.1); *Trifolium repens* TrMYB4 AMB27079), TrMYB133 (AMB27081), and TrMYB134 (AMB27082); *Vaccinium ashei* VaMYB (QOQ50851.1), *Vaccinium corymbosum* VcMYBA (MH105054); *Vaccinium uliginosum* VuMYBC2; *Vitis vinifera* VvMYBA2 (BAD18978), VvMYBF1 (ACV81697), VvMYBC2-L1 (AFX64995.1), VvMYB-L2 (ACX50288.2), and VvMYBPA2 (ACK56131.1).

Vector Construction

Full-length coding sequence of *VbMYBA* was isolated using 2 × Phanta Max Master Mix (Vazyme, Nanjing, China) with primer VbMYBAOEF and VbMYBAOER and inserted into pSAK277. A fragment (70–855 bp) of *VbMYBA* was amplified by primer VbMYB-NdelOEF and VbMYB-NdelOER. Fusion PCR was carried out to generate VcMYBA-Nadd (*VcMYBA* with 1–69 bp of *VbMYBA*) were constructed as following: the N-terminal of VbMYBA (1–69 bp) was amplified by primer VbMYBNF and VbMYBNR, and the *VcMYBA* contains overlap with N-terminal of *VbMYBA* (1–69 bp) was amplified by VcMYBF and VcMYBR. Then the two fragments were fused through PCR. Finally, the VcMYBA-Nadd was amplified by primer VbMYB-NaddOEF and VbMYB-NaddOER, and inserted into pSAK277. The insertion of PCR fragments into pSAK277 was conducted using ClonExpress Ultra One Step Cloning Kit (Vazyme, Nanjing, China). Primer sequences used for the vector construction are listed in **Supplementary Table 2**. The promoter of *VcDFR* was previously isolated (Plunkett et al., 2018) and the promoter of *VavUFGT* was isolated from *Vaccinium virgatum* using the primers VavpUFGT-F (CTCCACATTTTTAACCTGGTGCAC) and VavpUFGT-R (CATGGTTATATTTTTGGTGGT), and cloned into pGreenII 0800-LUC (Hellens et al., 2005).

Transient Transformation in Tobacco Leaf and Quantification of Anthocyanins

Transient color assays were carried out in young leaves of *Nicotiana tabacum* seedlings grown in the greenhouse as described previously (Fang et al., 2021). All the constructs were transformed into *Agrobacterium tumefaciens strain GV3101*, and incubated at 28°C for 2 days. *Agrobacterium* cultures carrying constructs were resuspended in infiltration buffer containing 10 mM $MgCl_2$ and 100 μM acetosyringone (pH = 5.7) and incubated at room temperature without shaking for 2 h before infiltration. Mixed bacterial cultures were injected into the abaxial side of young leaves. Many studies have suggested that basic helix-loop-helix (bHLH) transcription factors are indispensable partners of R2R3-MYBs (Espley et al., 2007; Gonzalez et al., 2008), in particular those belonging to the bHLH2 subgroup (PhAN1/AtTT8). Tobacco leaves express WDR and bHLH1 (PhJAF13/AtEGL3) orthologs, but require a bHLH2 (PhAn1/AtTT8) protein to fully regulate anthocyanin biosynthesis (Montefiori et al., 2015). Because the ability for the MYB to regulate the endogenous tobacco *bHLH2* gene is variable in transient assays, a *bHLH2* (*PsbHLH3*) gene was co-transformed. We failed to clone the anthocyanin-associated bHLH from Wufanshu, and so we chose the plum anthocyanin-activating PsbHLH3 for transient color assays in tobacco leaves (Fang et al., 2021). This is a close homolog of peach PpbHLH3, which showed stronger anthocyanin-promoting activity than the homologous apple MdbHLH3 when coinfiltrated with blueberry VcMYBA (Plunkett et al., 2018). Separate strains containing *VbMYBA* and *PsbHLH3* fused to the 35S promoter in the

[1]https://www.ncbi.nlm.nih.gov/orffinder/

[2]https://www.genome.jp/tools-bin/clustalw

pSAK277 vector and empty pSAK277 vector were infiltrated or co-infiltrated into the abaxial leaf surface. Blueberry *VcMYBA* (Plunkett et al., 2018) co-infiltrated with *PsbHLH3* was used as positive control. Each infiltration was performed using three leaves from the same plants. Photographs were taken seven days after infiltration.

For quantification of anthocyanins, 10 mg of freeze-dried tobacco leaves from the infiltrated area was mixed in 1 mL of methanol: water: formic acid (80: 19.5: 0.5, v/v/v) and shaken for 4 h at room temperature. The tube containing the mixture was centrifuged at 10,000 × *g* for 15 min. The supernatants were filtered through a 0.45 μm PTFE syringe filter and submitted to high performance liquid chromatography (HPLC) analysis according to a method reported by Andre et al. (2012) with a few modifications. Briefly, quantification of the anthocyanins was performed using a Dionex Ultimate 3000 system (Sunnyvale, CA, United States) equipped with a diode array detector (DAD). A 5 μL aliquot was injected onto a Thermo C18 Acclaim PolarAdvantage II column (150 × 2.1 mm i.d.; 3 μm particle size) (Waltham, MA, United States). The mobile phases were (A) H_2O with 5% formic acid and (B) MeCN with 0.1% formic acid. The flow rate was 0.35 mL min^{-1}, and the column temperature was 35°C. The 28 min gradient was as follows: 0-5 min, 7% B constant; 5-10 min, 7-12% B; 10-20 min, 12-25% B; 20-21 min, 25-100% B; 21-24 min, 100% B constant; 24 min, 7% B; 24-28 min, 7% B re-equilibration time. Monitoring was set at 520 nm for quantification. Anthocyanins were identified by their spectral data and were quantified as cyanidin-3-glucoside using five-point calibration curves. A validation standard was injected after every 10th injection.

Dual-Luciferase Assay

Dual-luciferase assays were conducted in *Nicotiana benthamiana* leaves as reported previously (Lin-Wang et al., 2010). All the promoter constructs were individually transformed into *Agrobacterium* strain GV3101 that contains the pSoup plasmid using the electroporation method. Agrobacteria cultivation and infiltration preparation were performed according to the same protocol as described above for the transient color assay.

RESULTS

Accumulation of Anthocyanin in Berries of Wufanshu

The skin of unripe Wufanshu berries was initially green, then turned red and finally black when fully ripen (**Figure 1A**). The total content in anthocyanins increased during the ripening of berries and reached 516.97 mg/100 g FW at the full ripe stage (**Figure 1B**). Individual anthocyanin compounds were characterized in ripe Wufanshu berries using UHPLC–MS/MS analysis. Delphinidin, cyanidin, malvidin, peonidin, petunidin, as well as pelargonidin derivatives could be identified based on their mass spectral data. Delphinidin 3-galactoside and 3-glucoside predominated the anthocyanin profile (**Supplementary Table 3**).

RNA Sequencing and *de novo* Assembly of Wufanshu Berry Transcriptomes

To identify genes responsible for anthocyanin accumulation in the berries of Wufanshu and investigate the underlying molecular mechanisms, transcriptomes of berries at green and black stages were generated by RNA-Seq. A total of 13.49 Gb clean reads were generated (**Supplementary Table 4**) and *de novo*-assembled into 87,811 unigenes. The average length of unigenes was 672.45 nt. Over one-half of the unigenes (61.59%) were shorter than 500 bp, and only 18.59% (16,319) of the unigenes were longer than 1 kb (**Supplementary Table 5**).

All unigenes were annotated against eight public protein databases; NR, Swiss-Prot, Pfam, GO, COG, KOG, eggNOG4.5 and KEGG. In total, 49.37% (43,355) of the unigenes showed significant BLAST matches to known sequences in these databases (**Supplementary Table 6**). According to the

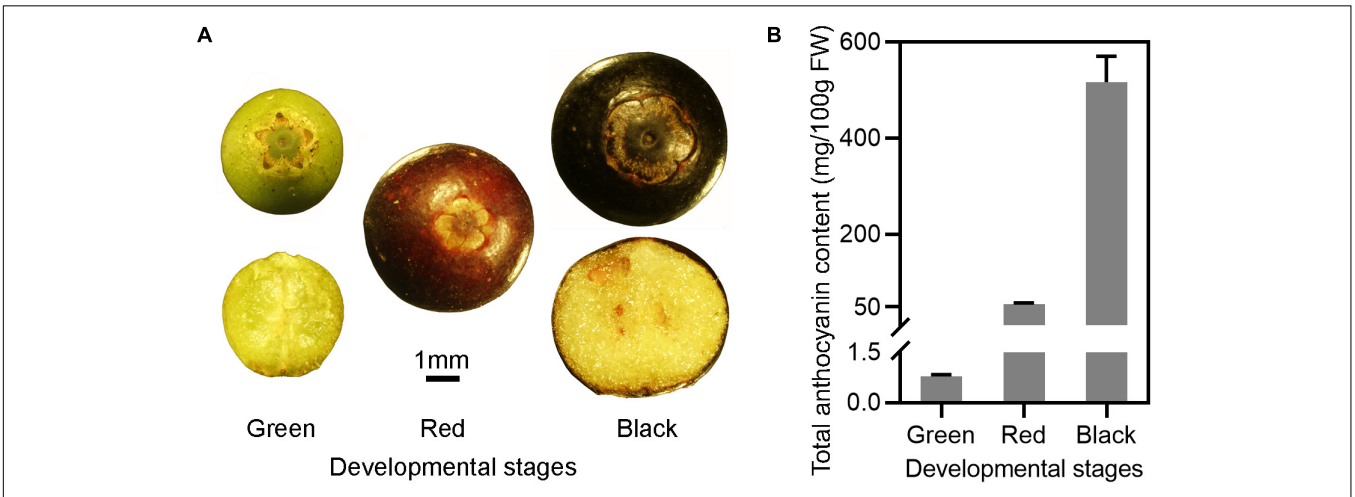


FIGURE 1 | Anthocyanin accumulation in the fruits of Wufanshu. **(A)** Fruits of Wufanshu at different stages. **(B)** Anthocyanin content in fruits of Wufanshu at different ripening stages. Error bars represent standard error of three replicates. Error bars represent standard error of five replicates.

KEGG enrichment analysis, 168 unigenes were assigned to phenylpropanoid biosynthesis pathway (ko00940), 48 unigenes were assigned to flavonoid biosynthesis pathway (ko00941) and fourteen genes were involved in the anthocyanin biosynthesis pathway (ko00942) (**Supplementary Table 7**). Twenty-seven of these genes, which were more abundant in transcript levels in black Wufanshu berries (**Supplementary Table 8**), were identified as candidate genes involved in anthocyanin biosynthesis. c81310.graph_c0 was annotated as leucoanthocyanidin dioxygenase according to Swissprot database and c84705.graph_c0 was annotated as UDP-glucose: flavonoid 3-*O*-glucosyltransferase according to Nr database (**Supplementary Table 8**). In addition, c72505.graph_c0 was predicted to encode a homolog of cyclamen CkmGST3, reported to be involved in anthocyanin accumulation (Kitamura et al., 2012).

Anthocyanin Biosynthetic Genes Are Upregulated in the Berries of Wufanshu

Based on the RNA-Seq data, the expression levels of 11 anthocyanin biosynthetic genes with high FPKM values (maximum FPKM > 40) were analyzed by qRT-PCR. Our results indicated that the expression of *VbPAL* (*phenylalanine ammonia-lyase*, c84812.graph_c0), *Vb4CL* (*4-coumaroyl:CoA-ligase*, c83421.graph_c0), *VbC4H* (*cinnamate-4-hydroxylase*, c88472.graph_c0), *VbCHS* (*chalcone synthase*, c82998.graph_c0), *VbF3H* (*flavanone-3-hydroxylase*, c85319.graph_c0 and c78308.graph_c0), *VbF3′H* (*flavonoid3′-hydroxylase*, c88140.graph_c0), *VbF3′5′H* (c94417.graph_c2), *VbDFR* (*dihydroflavonol-4-reductase*, c76847.graph_c0), *VbUFGT* (*UDPglucose:flavonoid-3-O-glucosyltransferase*, c84705.graph_c0), and *VbGST* (*glutathione S-transferase*, c81305.graph_c0) were upregulated during ripening of Wufanshu berries (**Figures 2A–K**). Linear regression analysis showed that qRT-PCR value of the analyzed anthocyanin biosynthetic genes was significantly correlated with that FPKM value of them (**Figure 2L**).

Identification of *VbMYBA* From Wufanshu Berries

To identify candidate anthocyanin MYB activators, amino acid sequence of VcMYBA (Plunkett et al., 2018) from blueberry was

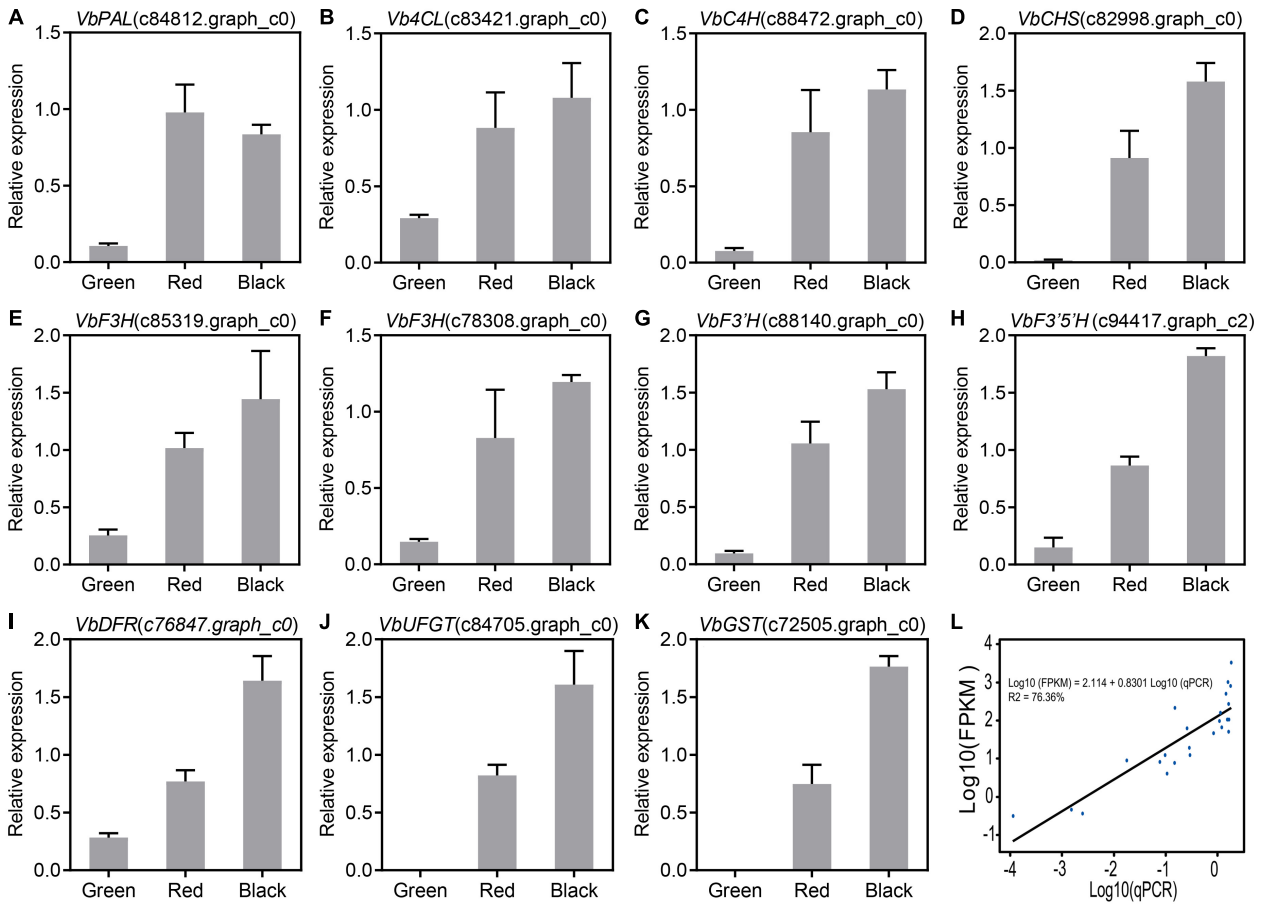


FIGURE 2 | Expression of candidate genes involved in anthocyanin biosynthesis. **(A–K)** qRT-PCR analysis of candidate genes involved in anthocyanin biosynthesis. *Actin* was used as the reference gene. The error bars represent the standard error of three biological replicates. **(L)** Linear regression analysis of gene expression levels obtained from RNA-Seq and quantitative real-time PCR data.

used as query to BLAST against the assembled unigene sequences using TBtools (Chen et al., 2020). The BLAST analysis enabled the identification of a unigene (c67315.graph_c0) encoding MYB protein. The full-length open-reading frame of c67315.graph_c0 was amplified from cDNA of Wufanshu berries and confirmed by sequencing. The gene was designated as *VbMYBA* (GenBank Accession No. MW543447) (**Supplementary Figure 1**). BLASTp search against NCBI non-redundant protein sequences showed that VbMYBA protein have highest matching score to VcMYB1 (AYC35399.1) from *V. corymbosum*. Sequence alignment analysis indicated that VbMYBA protein showed high sequence identity to blueberry VcMYB1 and VcMYBA and contains conserved R2R3 domain and SG6 motif for anthocyanin-promoting MYBs (**Figure 3**). Nineteen amino acids, three in each R domain and 12 in C-terminal, were different between VbMYBA and VcMYBA (**Figure 3**). In addition, the N-terminal of VbMYBA protein was 6 aa and 23 aa longer than blueberry VcMYB1 and VcMYBA, respectively (**Figure 3**). Phylogenetic analysis showed that VbMYBA belongs to SG6 clade, which contains anthocyanin-activating MYB proteins from other plant species and is closely related to the anthocyanin-activators from *Vaccinium* species (**Figure 4**).

VbMYBA Is a Strong Anthocyanin Activator

Multiple sequences alignment and phylogenetic analysis suggested that VbMYBA was an anthocyanin-activator, similar to VcMYBA. RNA-Seq data along with qRT-PCR analysis indicated that the expression of *VbMYBA* was closely associated with that of anthocyanin biosynthetic genes and anthocyanin accumulation (**Figure 5A**). The function of VbMYBA was verified using tobacco leaf transient expression assays. Our results indicated that infiltration of *VbMYBA* with or without *PsbHLH3* resulted in strong red pigmentation at infiltration sites 5 days after transformation, but only faint

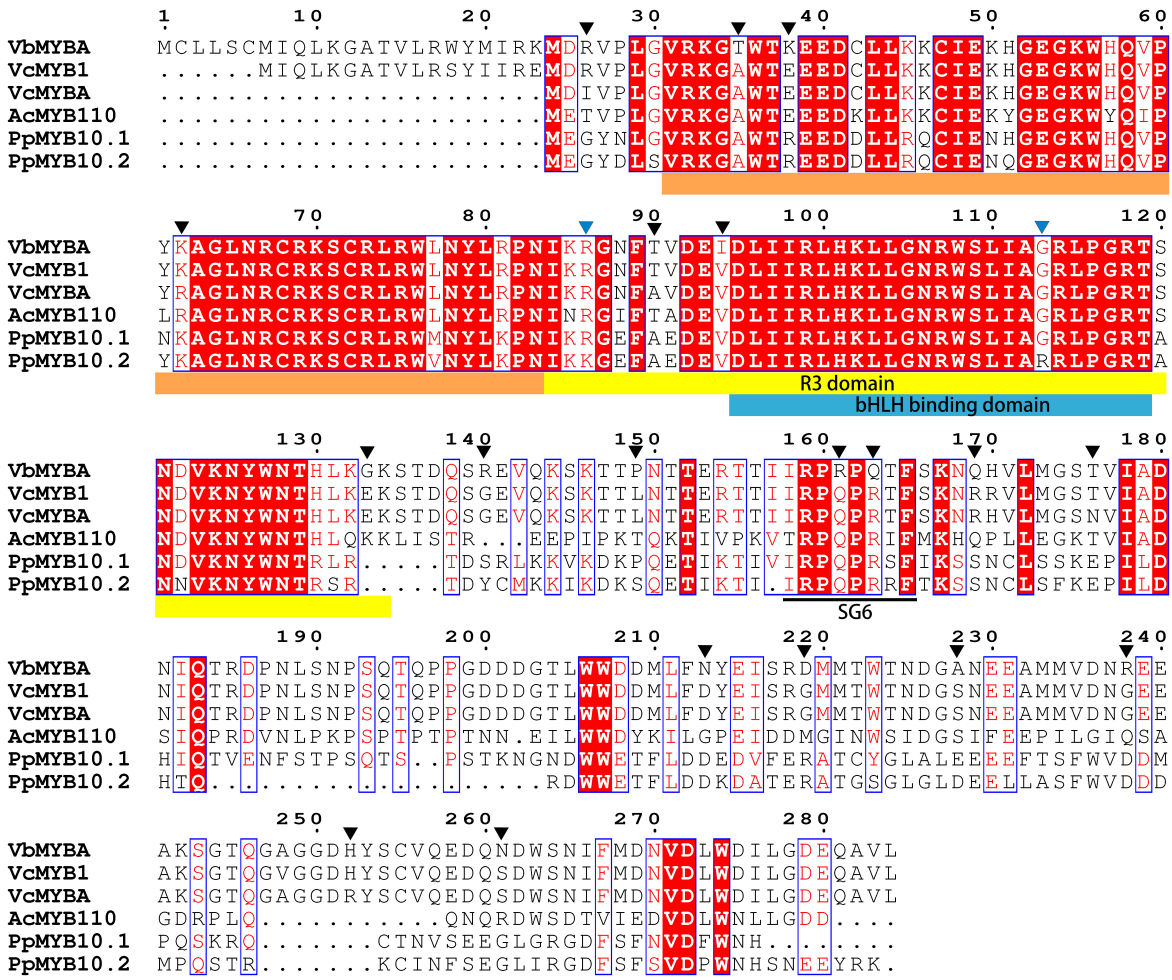


FIGURE 3 | Sequence alignment of VbMYBA and anthocyanin MYB activators from blueberry and peach. R2, R3, and bHLH binding domain are highlighted in marigold yellow, yellow, and blue colors, conserved SG6 motif for anthocyanin-promoting MYBs are indicated in black line. Amino acids that are different between VbMYB and blueberry VcMYB were indicate with black triangles. Amino acids that were reported to affect the anthocyanin-promoting activity of MYB proteins (Zhou et al., 2018) are indicate with blue triangles.

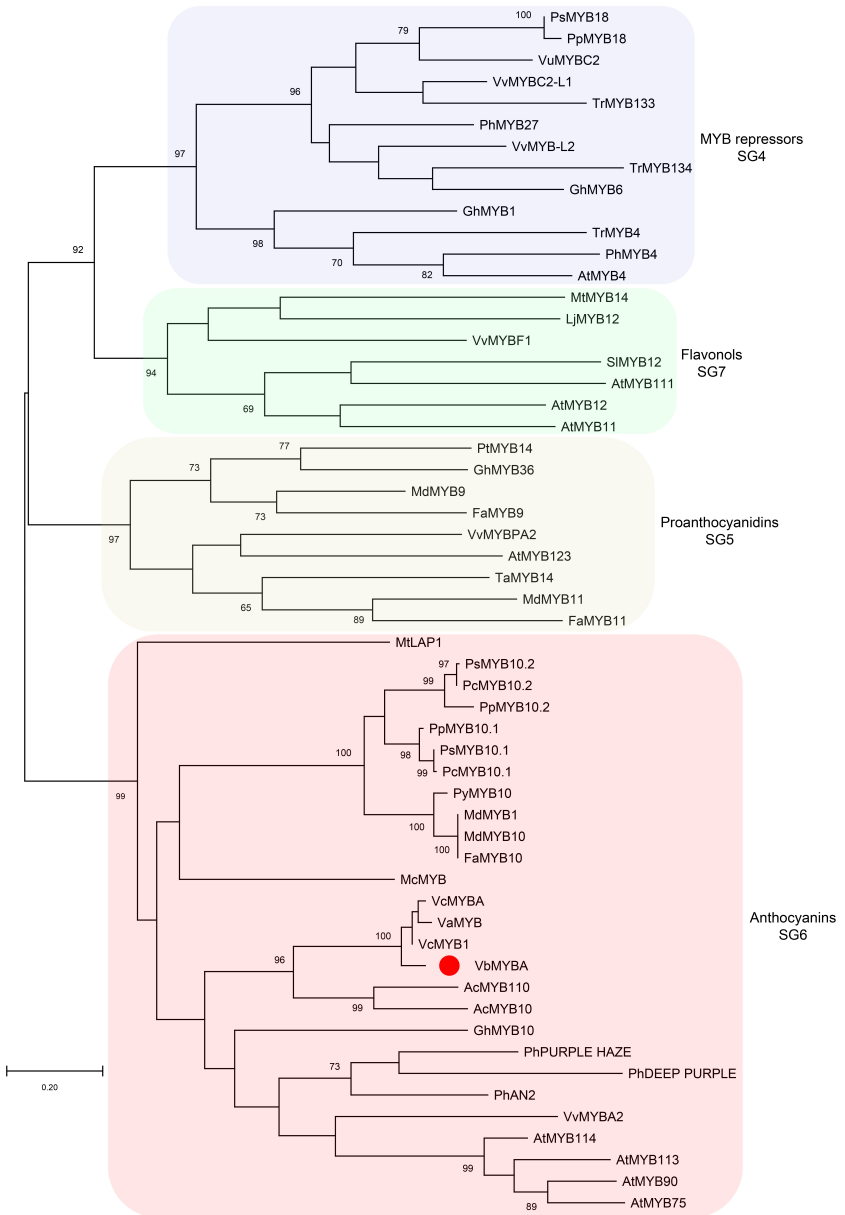


FIGURE 4 | Phylogenetic analysis of VbMYBA and R2R3 MYBs from other plant species. Phylogenetic tree was constructed using maximum-likelihood method by MEGA-X software. VbMYB from Wufanshu fruits is indicated with red dot. Phylogenetic groups are highlighted with different color shades. Numbers indicate node support >65% from 1000 bootstrap replicates. The scale bar represents 0.2 substitutions per site.

red coloration was found when *VcMYBA* and *PsbHLH3* were co-infiltrated (**Figure 5B**). Anthocyanin content analysis in leaves of *N. tabacum* indicated that no anthocyanin was detected in the infiltrated areas transformed with empty vector or PsbHLH3 (**Figure 5C**). A high concentration of anthocyanin was detected in tobacco leaves infiltrated with *VbMYBA*, while co-infiltration of *VbMYBA* and *PsbHLH3* did not result in stronger anthocyanin accumulation in tobacco leaves (**Figure 5C**). However, the anthocyanin content in tobacco leaves infiltrated with *VcMYBA* and *PsbHLH3* was significantly lower than that in tobacco leaves infiltrated with *VbMYBA* with or without *PsbHLH3* (**Figure 5C**). Among the anthocyanins detected in tobacco leaves, anthocyanin 1 and 2, which were likely correspond to a cyanidin-3-rutinoside and delphinidin-3-rutinoside, respectively, according to their UV spectra and literature data, were the predominant forms (**Figure 5C**). In addition, we employed dual-luciferase assays to investigate the function of *VbMYBA*. The promoters of *VcDFR* and *VavUFGT* from blueberry (*V. corymbosum* and *V. virgatum*, respectively) were fused to a luciferase reporter. Infiltration of *VbMYBA* led to strong activation of promoters *VcDFR* and *VavUFGT* (**Figures 5D,E**). These results demonstrate that VbMYBA is a

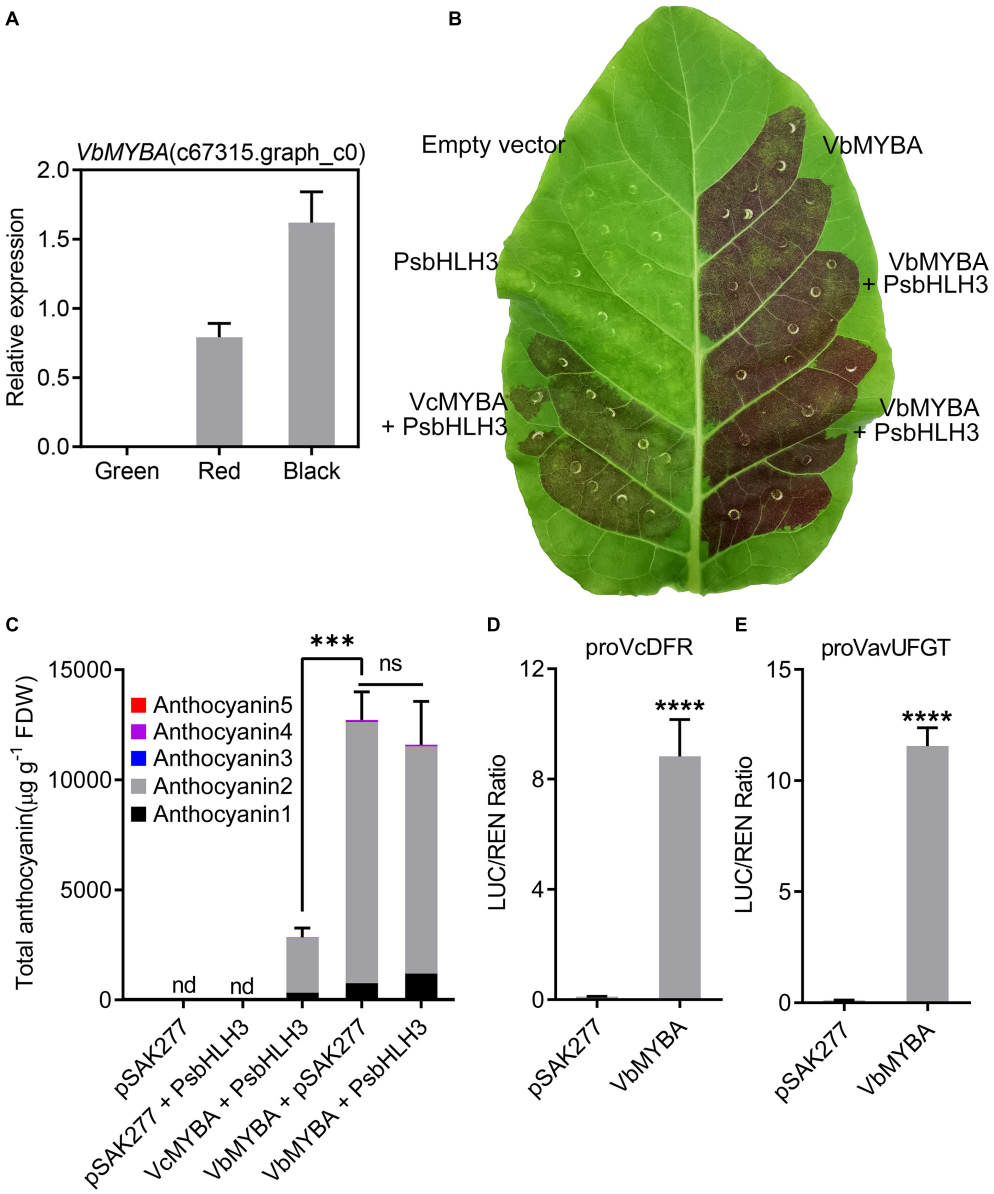


FIGURE 5 | Functional analysis of the *VbMYBA* gene using transient assay in tobacco. **(A)** Relative gene expression of *VbMYBA* in the fruits of Wufanshu. **(B)** Transient expression of *VbMYBA* in tobacco leaf. The photo was taken 5 days after infiltration. Blueberry *VcMYBA*/*PsbHLH3* was used as positive control. **(C)** Anthocyanin content in transformed leaves of tobacco. Anthocyanins were measured as cyanidin-3-galactoside equivalents. Each value represents the mean of three biological replicates. FDW, freeze dry weight. nd, non-detected. Asterisks (***) indicate significant differences among treatments according to *t*-test ($P < 0.0002$). ns, non-significant. Analysis of the interaction of the *VbMYBA* gene with the promoters of VcDFR gene **(D)** from blueberry and *VavUFGT* gene **(E)** from *V. virgatum* using dual-luciferase reporter assay *Nicotiana benthamiana* leaves. Error bars represent the SEs for three replicates. Asterisks (****) indicate significant differences among treatments according to *t*-test ($P < 0.0001$).

strong anthocyanin activator and is capable of activating the promoters of *DFR* and *UFGT* from blueberry.

N-Terminal Amino Acid Residues 1-23 of VbMYBA Is Not Responsible for Its Higher Anthocyanin-Promoting Activity

Since the sequence for VbMYBA contained a 23 amino acid-residue insertion in the N-terminal region compared with VcMYBA, we hypothesized that the N-terminal amino acid residues 1-23 of VbMYBA (N^{1-23}) are responsible for the capacity for higher anthocyanin-promoting activity. To investigate whether the N^{1-23} is responsible for the divergence of anthocyanin-promoting activity between VbMYBA and VcMYBA, two mutant constructs were generated by deleting the N^{1-23} in VbMYBA (VbMYBA-Ndel) and fusing the N^{1-23} to VcMYBA (VcMYBA-Nadd) (**Figure 6A**). Anthocyanin-promoting activities of the mutant constructs

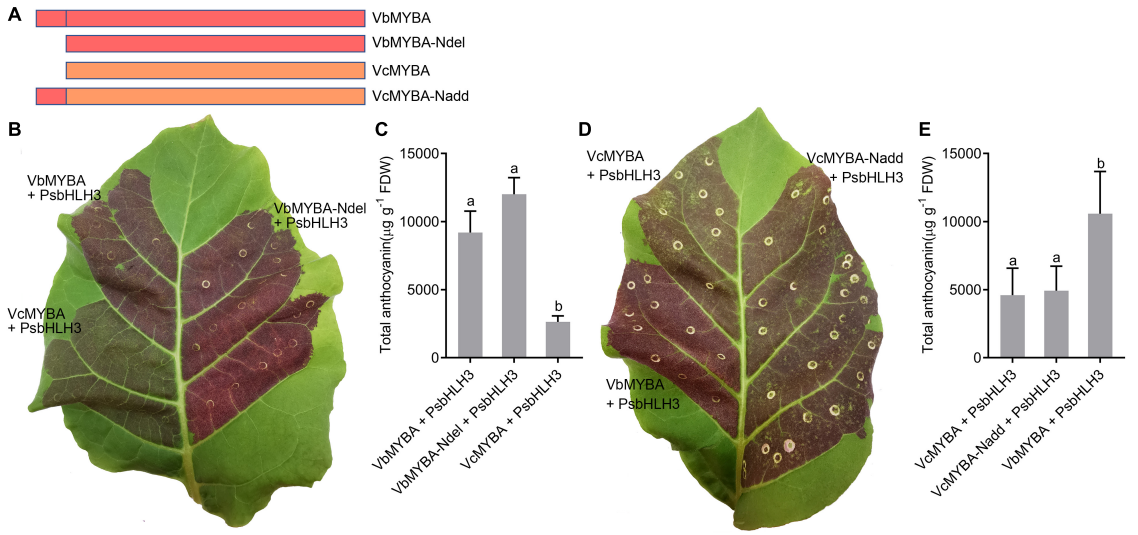


FIGURE 6 | Effect of N-terminal amino acid residues 1-23 of VbMYBA on anthocyanin-promoting activity. **(A)** Schematic diagram of mutants of VbMYBA and VcMYBA. **(B,D)** Transient expression of *VbMYBA*, *VcMYBA* and their mutants in tobacco leaf. The photo was taken 7 days after infiltration. **(C,E)** Anthocyanin content in transformed tobacco leaves. Error bars represent the standard errors for three replicates. Different lowercase letters indicate significant differences among treatments according to one-way ANOVA ($P < 0.01$).

were tested by tobacco leaf transient expression assays. Our results indicated that deletion of N^{1-23} in VbMYBA or insertion of N^{1-23} to VcMYBA have no significant effects on anthocyanin content in infiltrated tobacco leaves (**Figures 6B–E**). These results demonstrated that the N-terminal amino acid residues 1-23 of VbMYBA was not responsible for the divergence of anthocyanin-promoting activity between VbMYBA and VcMYBA.

DISCUSSION

Fruits of *Vaccinium* species are becoming increasingly popular due to their high levels of health-promoting bioactive compounds such as anthocyanins (Krikorian et al., 2010; Norberto et al., 2013; Chu et al., 2017; Shi et al., 2017; Zhou et al., 2020). In this study, the mechanism involved in anthocyanin accumulation in Wufanshu berries was investigated using UHPLC-MS/MS, RNA-Seq, dual-luciferase assays and transient color assays in tobacco leaves.

Anthocyanin composition and content has been reported to be highly variable among *Vaccinium* species and cultivars (Lohachoompol et al., 2008; You et al., 2011; Stevenson and Scalzo, 2012). Our results indicated that ripe Wufanshu berries are rich in a complex set of anthocyanins, including delphinidin, cyanidin, malvidin, peonidin, pelargonidin, and petunidin derivatives. The total anthocyanin content in Wufanshu berries is comparable to the highest content in blueberries as reported by other studies (Stevenson and Scalzo, 2012; Kim et al., 2013; Günther et al., 2020). Pelargonidin and its derivatives were shown to be absent in fruits of *Vaccinium* species such as highbush blueberry (Chung et al., 2016; Chai et al., 2021) and bilberry (Jaakola et al., 2002). However, four derivatives of pelargonidin (pelargonidin-3-*O*-galactoside, pelargonidin-3-*O*-glucoside, pelargonidin-3-*O*-rutinoside and pelargonidin-3,5-*O*-diglucoside) were detected in ripe Wufanshu berries (**Supplementary Table 3**). In future, it will be interesting to investigate anthocyanin composition of the berries of *V. bracteatum* species and elucidate the mechanisms and pathway enzymes responsible for the accumulation of pelargonidin and its derivatives.

Real-time quantitative RT-PCR analyses showed that whole pathway of anthocyanin biosynthetic genes was upregulated in Wufanshu berries during ripening. *R2R3 MYB* genes have been reported to act as anthocyanin activators in blueberry (Nguyen et al., 2017; Plunkett et al., 2018). Blast analysis enabled us to identify *VbMYBA*, a homolog of the blueberry *R2R3 MYB* gene *VcMYBA* (Plunkett et al., 2018), from the transcriptome of Wufanshu berries. Sequence alignment and phylogenetic analysis suggested that the sub-group 6 VbMYBA is likely to be an anthocyanin activator. qRT-PCR analyses indicated that the expression of *VbMYBA* was well correlated with anthocyanin biosynthetic gene expression and anthocyanin accumulation in Wufanshu berries. Overexpression of *VbMYBA* with or without *PsbHLH3* induced anthocyanin accumulation and red pigmentation in tobacco leaves. Dual-luciferase assays also show that VbMYBA can strongly activate the promoter of anthocyanin biosynthetic genes *VcDFR* and *VavUFGT* from other *Vaccinium* species. These results suggested that VbMYBA is an anthocyanin activator.

The anthocyanin-promoting ability of different *MYB* genes varied significantly (Zhou et al., 2018). Plunkett et al. (2018) show that transient overexpression of blueberry *VcMYBA* with *bHLH3* genes can induce a greater concentration of anthocyanin pigments in tobacco than apple *MdMYB10* and peach *PpMYB10.1*. Zhou et al. (2018) demonstrated that two

amino acid changes, Arg/Lys^{66} and Gly/Arg^{93}, in the R3 repeat is responsible for anthocyanin-promoting activity divergence between peach *PpMYB10.1* and *PpMYB10.2*. They further demonstrated that reciprocal substitution of Arg/Gly^{93} between PpMYB10.1 and PpMYB10.2 affect their binding affinity to PpbHLH3 and suggested that bHLH-binding affinity is a key factor that determine the anthocyanin-promoting activity of MYB genes (Zhou et al., 2018). However, sequence alignment showed that the two amino acid and bHLH-binding domain were identical between VcMYBA and VbMYBA (**Figure 3**). The N-terminal region of VbMYBA is longer than VcMYBA, but deletion of 23 amino acid-residue in the N-terminal region of VbMYBA did not reduce its activity. In addition, ligation of the amino acid-residues to the N-terminal of VcMYBA did not enhance its anthocyanin promoting activity. These results suggest that the divergence of anthocyanin-promoting activity between VcMYBA and VbMYBA is likely to be due to the 19 different amino acids that distributed in R domains and C-terminal region. R2 and R3 domains of R2R3-MYBs play pivotal roles in interaction with DNA and determining their DNA binding affinity (Jia et al., 2004; Jiang et al., 2004). It has been reported that mutation of the R2 and R3 domains caused significant alteration in promoter target specificity and DNA-binding affinity of MYB proteins (Williams and Grotewold, 1997; Hernandez et al., 2004; Heppel et al., 2013). The C-terminus contains activation or repression and play an important role in determining the function of MYB proteins (Dubos et al., 2010; Feller et al., 2011). Variation of C-terminus also affect function of MYB proteins (Liu et al., 2016). Further study would be needed to identify the key amino acids responsible for the high anthocyanin-promoting activity of VbMYBA. It is also noteworthy that 23 amino acid-residue in the N-terminal region of VbMYBA should be further verified at a protein level.

It has been demonstrated that bHLHs act as a crucial component in the regulation of anthocyanin biosynthesis. In Arabidopsis, the bHLH factors EGL3 and TT8 are necessary for anthocyanin accumulation in seedlings (Gonzalez et al., 2008). Several studies have indicated that bHLHs are essential for the anthocyanin activating activity of MYB proteins, such as potato R2R3 MYBs (Liu et al., 2016), peach PpMYB10.2 (Zhou et al., 2018), and apple MdMYB10 (Espley et al., 2007) in *Nicotiana tabacum* leaves. Infiltration of some *MYB*s, such as peach *PpMYB10.1* (Zhou et al., 2015) and *PpMYB10.4* (Zhou et al., 2014), Chinese bayberry *MrMYB1* (Niu et al., 2010), litchi *LcMYB1* (Lai et al., 2016), alone activated weak anthocyanin pigmentation in *N. tabacum* leaves, while co-infiltration with *bHLH*s significantly enhanced anthocyanin accumulation. Kiwifruit *AcMYB110* alone was able to induce strong anthocyanin accumulation in *N. tabacum* leaves (Peng et al., 2019). Similarly, infiltration of *VbMYBA* also triggered strong anthocyanin pigmentation in tobacco leaves and co-infiltration with *PsbHLH3* has no significant effects on anthocyanin accumulation (**Figures 5A,B** and **Supplementary Figure 1**). One possibility is that *VbMYBA* can recruit bHLHs of *N. tabacum* to activate anthocyanin biosynthetic genes. Intriguingly, its homolog *VcMYBA*, which has a same bHLH-binding domain, cannot activate anthocyanin accumulation in *N. tabacum* leaves when infiltrated alone (**Supplementary Figure 2**). It will be interesting to investigate the underlying mechanisms that are responsible for MYBs, such as VbMYBA and AcMYB110, to activate anthocyanin accumulation without bHLHs in tobacco leaves.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are publicly available. This data can be found here: The raw transcriptome data have been deposited in the NCBI Sequence Read Archive under accession number PRJNA694726.

AUTHOR CONTRIBUTIONS

Z-ZF supervised the project. Y-LZ and Z-ZF wrote the manuscript. Y-LZ, Z-ZF, NA, CE, and CA participate in the experiments. KL-W, NA, CE, and RE provided scientific suggestion and revised the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

This research was funded by the Project of Fujian Academy of Agricultural Sciences (AA2018-16) and the China Scholarship Council (201909350001). Work at Plant & Food Research was carried out under The New Zealand Ministry of Business, Innovation & Employment (MBIE) contract C11 × 1704 "Filling the void: boosting the nutritional content of NZ fruit".

ACKNOWLEDGMENTS

We thank Monica Dragulescu and her team for plant care.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2021.697212/full#supplementary-material

REFERENCES

Albert, N. W., Butelli, E., Moss, S. M. A., Piazza, P., Waite, C. N., Schwinn, K. E., et al. (2021). Discrete bHLH transcription factors play functionally overlapping roles in pigmentation patterning in flowers of *Antirrhinum majus*. *New Phytol.* 231, 849–863. doi: 10.1111/nph.17142

Albert, N. W., Davies, K. M., Lewis, D. H., Zhang, H., Montefiori, M., Brendolise, C., et al. (2014). A conserved network of transcriptional activators and repressors regulates anthocyanin pigmentation in eudicots. *Plant Cell* 26, 962–980. doi: 10.1105/tpc.113.122069

Andre, C. M., Greenwood, J. M., Walker, E. G., Rassam, M., Sullivan, M., Evers, D., et al. (2012). Anti-inflammatory procyanidins and triterpenes in 109 apple varieties. *J. Agric. Food Chem.* 60, 10546–10554. doi: 10.1021/jf302809k

Butelli, E., Garcia-Lor, A., Licciardello, C., Las Casas, G., Hill, L., Reforgiato Recupero, G., et al. (2017). Changes in anthocyanin production during domestication of *Citrus*. *Plant Physiol.* 173, 2225–2242. doi: 10.1104/pp.16.01701

Chai, Z., Herrera-Balandrano, D. D., Yu, H., Beta, T., Zeng, Q., Zhang, X., et al. (2021). A comparative analysis on the anthocyanin composition of 74 blueberry cultivars from China. *J. Food Comp. Anal.* 102:104051. doi: 10.1016/j.jfca.2021.104051

Chen, C., Chen, H., Zhang, Y., Thomas, H. R., Frank, M. H., He, Y., et al. (2020). TBtools - an integrative toolkit developed for interactive analyses of big biological data. *Mol. Plant* 13, 1194–1202. doi: 10.1016/j.molp.2020.06.009

Cho, H.-S., Cho, Y.-S., and Cho, J.-A. J. A. H. (2012). Eco-physiological and horticultural characteristics of two Korean wild Vacciniums. *Acta hortic.* 926, 149–156.

Chu, W., Gao, H., Cao, S., Fang, X., Chen, H., and Xiao, S. J. F. C (2017). Composition and morphology of cuticular wax in blueberry (*Vaccinium* spp.) fruits. *Food Chem.* 219, 436–442.

Chung, S. W., Yu, D. J., and Lee, H. J. (2016). Changes in anthocyanidin and anthocyanin pigments in highbush blueberry (*Vaccinium corymbosum* cv. Bluecrop) fruits during ripening. *Hortic. Environ. Biotechnol.* 57, 424–430. doi: 10.1007/s13580-016-0107-8

Die, J. V., Jones, R. W., Ogden, E. L., Ehlenfeldt, M. K., and Rowland, L. J. (2020). Characterization and analysis of anthocyanin-related genes in wild-type blueberry and the pink-fruited mutant cultivar 'Pink Lemonade': new insights into anthocyanin biosynthesis. *Agronomy* 10:1296.

Dubos, C., Stracke, R., Grotewold, E., Weisshaar, B., Martin, C., and Lepiniec, L. (2010). MYB transcription factors in *Arabidopsis*. *Trends Plant Sci.* 15, 573–581. doi: 10.1016/j.tplants.2010.06.005

Espley, R. V., Hellens, R. P., Putterill, J., Stevenson, D. E., Kutty-Amma, S., and Allan, A. C. (2007). Red colouration in apple fruit is due to the activity of the MYB transcription factor, MdMYB10. *Plant J.* 49, 414–427. doi: 10.1111/j.1365-313X.2006.02964.x

Fan, M., Fan, Y., Huang, W., Wang, L., Li, Y., Qian, H., et al. (2018). Tentative characterization of precursor compounds and co-factors of pigment formation in production of 'wu mi' from *Vaccinium bracteatum* Thunb. Leaves. *Food Chem.* 262, 199–205. doi: 10.1016/j.foodchem.2018.04.101

Fan, M., Fan, Y., Rao, Z., Li, Y., Qian, H., Zhang, H., et al. (2019). Comparative investigation on metabolite changes in 'wu mi' production by *Vaccinium bracteatum* Thunb. leaves based on multivariate data analysis using UPLC–QToF–MS. *Food Chem.* 286, 146–153. doi: 10.1016/j.foodchem.2019.01.144

Fan, M., Lian, W., Li, T., Fan, Y., Rao, Z., Li, Y., et al. (2020). Metabolomics approach reveals discriminatory metabolites associating with the blue pigments from *Vaccinium bracteatum* thunb. Leaves at different growth stages. *Ind. Crops Prod.* 147:112252. doi: 10.1016/j.indcrop.2020.112252

Fang, Z. Z., Zhou, D. R., Ye, X. F., Jiang, C. C., and Pan, S. L. (2016). Identification of candidate anthocyanin-related genes by transcriptomic analysis of 'Furongli' plum (*Prunus salicina* Lindl.) during fruit ripening using RNA-Seq. *Front. Plant Sci.* 7:1338. doi: 10.3389/fpls.2016.01338

Fang, Z.-Z., Lin-Wang, K., Zhou, D.-R., Lin, Y.-J., Jiang, C.-C., Pan, S.-L., et al. (2021). Activation of PsMYB10.2 transcription causes anthocyanin accumulation in flesh of the red-fleshed mutant of 'Sanyueli' (*Prunus salicina* Lindl.). *Front. Plant Sci.* 12:1167. doi: 10.3389/fpls.2021.680469

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

Feng, S., Sun, S., Chen, X., Wu, S., Wang, D., and Chen, X. (2015). PyMYB10 and PyMYB10.1 interact with bHLH to enhance anthocyanin accumulation in pears. *PLoS One* 10:e0142112. doi: 10.1371/journal.pone.0142112

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., and Amit, I. (2011). Full-length transcriptome assembly from RNA-Seq data without a references genome. *Nat. Biotechnol.* 29, 644–652. doi: 10.1038/nbt.1883

Günther, C. S., Dare, A. P., McGhie, T. K., Deng, C., Lafferty, D. J., Plunkett, B. J., et al. (2020). Spatiotemporal modulation of flavonoid metabolism in blueberries. *Front. Plant Sci.* 11:545. doi: 10.3389/fpls.2020.00545

Hellens, R. P., Allan, A. C., Friel, E. N., Bolitho, K., Grafton, K., Templeton, M. D., et al. (2005). Transient expression vectors for functional genomics, quantification of promoter activity and RNA silencing in plants. *Plant Methods* 1:13. doi: 10.1186/1746-4811-1-13

Heppel, S. C., Jaffe, F. W., Takos, A. M., Schellmann, S., Rausch, T., Walker, A. R., et al. (2013). Identification of key amino acids for the evolution of promoter target specificity of anthocyanin and proanthocyanidin regulating MYB factors. *Plant Mol. Biol.* 82, 457–471. doi: 10.1007/s11103-013-0074-8

Hernandez, J. M., Heine, G. F., Irani, N. G., Feller, A., Kim, M.-G., Matulnik, T., et al. (2004). Different mechanisms participate in the R-dependent activity of the R2R3 MYB transcription factor C1. *J. Biol. Chem.* 279, 48205–48213. doi: 10.1074/jbc.M407845200

Hichri, I., Barrieu, F., Bogs, J., Kappel, C., Delrot, S., and Lauvergeat, V. (2011). Recent advances in the transcriptional regulation of the flavonoid biosynthetic pathway. *J. Exp. Bot.* 62, 2465–2483. doi: 10.1093/jxb/erq442

Holton, T. A., and Cornish, E. C. (1995). Genetics and biochemistry of anthocyanin aiosynthesis. *Plant Cell* 7, 1071–1083. doi: 10.1105/tpc.7.7.1071

Jaakola, L. (2013). New insights into the regulation of anthocyanin biosynthesis in fruits. *Trends Plant Sci.* 18, 477–483. doi: 10.1016/j.tplants.2013.06.003

Jaakola, L., Määttä, K., Pirttilä, A. M., Törrönen, R., Kärenlampi, S., and Hohtola, A. (2002). Expression of genes involved in anthocyanin biosynthesis in relation to anthocyanin, proanthocyanidin, and flavonol levels during bilberry fruit development. *Plant Physiol.* 130, 729–739.

Jia, L., Clegg, M. T., and Jiang, T. (2004). Evolutionary dynamics of the DNA-binding domains in putative R2R3-MYB genes identified from rice subspecies indica and japonica genomes. *Plant Physiol.* 134, 575–585. doi: 10.1104/pp.103.027201

Jiang, C., Gu, X., and Peterson, T. (2004). Identification of conserved gene structures and carboxy-terminal motifs in the Myb gene family of *Arabidopsis* and *Oryza sativa* L. ssp. indica. *Genome Biol.* 5:R46. doi: 10.1186/gb-2004-5-7-r46

Jin, W., Wang, H., Li, M., Wang, J., Yang, Y., Zhang, X., et al. (2016). The R2R3 MYB transcription factor *PavMYB10.1* involves in anthocyanin biosynthesis and determines fruit skin colour in sweet cherry (*Prunus avium* L.). *Plant Biotechnol. J.* 14, 2120–2133. doi: 10.1111/pbi.12568

Kim, J. G., Kim, H. L., Kim, S. J., and Park, K.-S. (2013). Fruit quality, anthocyanin and total phenolic contents, and antioxidant activities of 45 blueberry cultivars grown in Suwon, Korea. *J. Zhejiang Univ. Sci. B* 14, 793–799. doi: 10.1631/jzus.B1300012

Kitamura, S., Akita, Y., Ishizaka, H., Narumi, I., and Tanaka, A. (2012). Molecular characterization of an anthocyanin-related glutathione S-transferase gene in cyclamen. *J. Plant Physiol.* 169, 636–642. doi: 10.1016/j.jplph.2011.12.011

Kobayashi, S., Ishimaru, M., Hiraoka, K., and Honda, C. (2002). Myb-related genes of the Kyoho grape (*Vitis labruscana*) regulate anthocyanin biosynthesis. *Planta* 215, 924–933. doi: 10.1007/s00425-002-0830-5

Krikorian, R., Shidler, M. D., Nash, T. A., Kalt, W., Vinqvist-Tymchuk, M. R., Shukitt-Hale, B., et al. (2010). Blueberry supplementation improves memory in older adults. *J. Agric. Food Chem.* 58, 3996–4000. doi: 10.1021/jf9029332

Lai, B., Du, L., Liu, R., Hu, B., Su, W., Qin, Y., et al. (2016). Two LcbHLH transcription factors interacting with *LcMYB1* in regulating late structural genes of anthocyanin biosynthesis in *Nicotiana* and *Litchi chinensis* during anthocyanin accumulation. *Front. Plant Sci.* 7:166. doi: 10.3389/fpls.2016.00166

Landa, P., Skalova, L., Bousova, I., Kutil, Z., Langhansova, L., Lou, J.-D., et al. (2014). *In vitro* anti-proliferative and anti-inflammatory activity of leaf and fruit extracts from *Vaccinium bracteatum* Thunb. *Pak. J. Pharm. Sci.* 27, 103–106.

Langmead, B., Trapnell, C., Pop, M., and Salzberg, S. L. (2009). Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol.* 10, 1–10. doi: 10.1186/gb-2009-10-3-r25

Lee, S., Jung, E. S., Do, S.-G., Jung, G.-Y., Song, G., Song, J.-M., et al. (2014). Correlation between species-specific metabolite profiles and bioactivities of blueberries (*Vaccinium* spp.). *J. Agric. Food Chem.* 62, 2126–2133. doi: 10.1021/jf405272b

Li, B., and Dewey, C. N. (2011). RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinformatics* 12:323. doi: 10.1186/1471-2105-12-323

Li, D., Li, B., Ma, Y., Sun, X., Lin, Y., and Meng, X. (2017). Polyphenols, anthocyanins, and flavonoids contents and the antioxidant capacity of various cultivars of highbush and half-high blueberries. *J. Food Comp. Anal.* 62, 84–93. doi: 10.1016/j.jfca.2017.03.006

Lin-Wang, K., Bolitho, K., Grafton, K., Kortstee, A., Karunairetnam, S., McGhie, T., et al. (2010). An R2R3 MYB transcription factor associated with regulation

of the anthocyanin biosynthetic pathway in Rosaceae. *BMC Plant Biol.* 10:50. doi: 10.1186/1471-2229-10-50

Liu, Y., Lin-Wang, K., Espley, R. V., Wang, L., Yang, H., Yu, B., et al. (2016). Functional diversification of the potato R2R3 MYB anthocyanin activators AN1, MYBA1, and MYB113 and their interaction with basic helix-loop-helix cofactors. *J. Exp. Bot.* 67, 2159–2176. doi: 10.1093/jxb/erw014

Lohachoompol, V., Mulholland, M., Srzednicki, G., and Craske, J. (2008). Determination of anthocyanins in various cultivars of highbush and rabbiteye blueberries. *Food Chem.* 111, 249–254. doi: 10.1016/j.foodchem.2008.03.067

Montefiori, M., Brendolise, C., Dare, A. P., Lin-Wang, K., Davies, K. M., Hellens, R. P., et al. (2015). In the Solanaceae, a hierarchy of bHLHs confer distinct target specificity to the anthocyanin regulatory complex. *J. Exp. Bot.* 66, 1427–1436. doi: 10.1093/jxb/eru494

Nesi, N., Jond, C., Debeaujon, I., Caboche, M., and Lepiniec, L. (2001). The Arabidopsis *TT2* gene encodes an R2R3 MYB domain protein that acts as a key determinant for proanthocyanidin accumulation in developing seed. *Plant Cell* 13, 2099–2114.

Nguyen, C. T. T., Lim, S., Lee, J. G., and Lee, E. J. (2017). VcBBX, VcMYB21, and VcR2R3MYB transcription factors are involved in UV–B-induced anthocyanin biosynthesis in the peel of harvested blueberry fruit. *J. Agric. Food Chem.* 65, 2066–2073. doi: 10.1021/acs.jafc.6b05253

Niu, S., Xu, C., Zhang, W., Zhang, B., Li, X., Lin-Wang, K., et al. (2010). Coordinated regulation of anthocyanin biosynthesis in Chinese bayberry (*Myrica rubra*) fruit by a R2R3 MYB transcription factor. *Planta* 231, 887–899. doi: 10.1007/s00425-009-1095-z

Norberto, S., Silva, S., Meireles, M., Faria, A., Pintado, M., and Calhau, C. (2013). Blueberry anthocyanins in health promotion: a metabolic overview. *J. Funct. Foods* 5, 1518–1528. doi: 10.1016/j.jff.2013.08.015

Oh, D. R., Kim, Y., Choi, E. J., Jung, M. A., Oh, K. N., Hong, J. A., et al. (2018a). Antidepressant-Like effects of *Vaccinium bracteatum* in chronic restraint stress mice: functional actions and mechanism explorations. *Am. J. Chin. Med.* 46, 357–387. doi: 10.1142/s0192415x18500180

Oh, D.-R., Yoo, J.-S., Kim, Y., Kang, H., Lee, H., Lm, S. J., et al. (2018b). *Vaccinium bracteatum* leaf extract reverses chronic restraint stress-induced depression-like behavior in mice: regulation of hypothalamic-pituitary-adrenal axis, serotonin turnover systems, and ERK/Akt phosphorylation. *Front. Pharmacol.* 9:604. doi: 10.3389/fphar.2018.00604

Peng, Y., Lin-Wang, K., Cooney, J. M., Wang, T., Espley, R. V., and Allan, A. C. (2019). Differential regulation of the anthocyanin profile in purple kiwifruit (*Actinidia species*). *Hortic. Res.* 6:3. doi: 10.1038/s41438-018-0076-4

Plunkett, B. J., Espley, R. V., Dare, A. P., Warren, B. A. W., Grierson, E. R. P., Cordiner, S., et al. (2018). MYBA from blueberry (*Vaccinium Section Cyanococcus*) is a subgroup 6 type R2R3MYB transcription factor that activates anthocyanin production. *Front. Plant Sci.* 9:1300. doi: 10.3389/fpls.2018.01300

Primetta, A. K., Karppinen, K., Riihinen, K. R., and Jaakola, L. (2015). Metabolic and molecular analyses of white mutant Vaccinium berries show down-regulation of MYBPA1-type R2R3 MYB regulatory factor. *Planta* 242, 631–643. doi: 10.1007/s00425-015-2363-8

Robert, X., and Gouet, P. (2014). Deciphering key features in protein structures with the new ENDscript server. *Nucleic Acids Res.* 42, W320–W324.

Shi, M., Loftus, H., McAinch, A. J., and Su, X. Q. (2017). Blueberry as a source of bioactive compounds for the treatment of obesity, type 2 diabetes and chronic inflammation. *J. Funct. Foods* 30, 16–29. doi: 10.1016/j.jff.2016.12.036

Stevenson, D., and Scalzo, J. J. J. O. B. R (2012). Anthocyanin composition and content of blueberries from around the world. *J. Berry Res.* 2, 179–189.

Tuan, P. A., Bai, S., Yaegaki, H., Tamura, T., Hihara, S., Moriguchi, T., et al. (2015). The crucial role of PpMYB10.1 in anthocyanin accumulation in peach and relationships between its allelic type and skin color phenotype. *BMC Plant Biol.* 15:280. doi: 10.1186/s12870-015-0664-5

Wang, L., Dong, M., and Yao, H. (2004). The nutritive value of black berry fruit of *Vaccinium bracteatum* Thunb. and its exploitation and utilization. *Zhong Cao Yao* 35, 14–15.

Wang, L., Zhang, X. T., Zhang, H. Y., Yao, H. Y., and Zhang, H. (2010). Effect of *Vaccinium bracteatum* Thunb. leaves extract on blood glucose and plasma lipid levels in streptozotocin-induced diabetic mice. *J. Ethnopharmacol.* 130, 465–469. doi: 10.1016/j.jep.2010.05.031

Wang, L., Zhang, Y., Xu, M., Wang, Y., Cheng, S., Liebrecht, A., et al. (2013). Anti-diabetic activity of *Vaccinium bracteatum* Thunb. leaves' polysaccharide in STZ-induced diabetic mice. *Int. J. Biol. Macromol.* 61, 317–321. doi: 10.1016/j.ijbiomac.2013.07.028

Williams, C. E., and Grotewold, E. (1997). Differences between plant and animal Myb domains are fundamental for DNA binding activity, and chimeric Myb domains have novel DNA binding specificities. *J. Biol. Chem.* 272, 563–571. doi: 10.1074/jbc.272.1.563

Xu, Y., Fan, M., Zhou, S., Wang, L., Qian, H., Zhang, H., et al. (2017). Effect of *Vaccinium bracteatum* Thunb. leaf pigment on the thermal, pasting, and textural properties and microstructure characterization of rice starch. *Food Chem.* 228, 435–440. doi: 10.1016/j.foodchem.2017.02.041

Yang, Y., Cui, B., Tan, Z., Song, B., Cao, H., and Zong, C. (2018). RNA sequencing and anthocyanin synthesis-related genes expression analyses in white-fruited *Vaccinium uliginosum*. *BMC Genomics* 19:930. doi: 10.1186/s12864-018-5351-0

You, Q., Wang, B., Chen, F., Huang, Z., Wang, X., and Luo, P. G. (2011). Comparison of anthocyanins and phenolics in organically and conventionally grown blueberries in selected cultivars. *Food Chem.* 125, 201–208. doi: 10.1016/j.foodchem.2010.08.063

Zhang, J., Chu, C.-J., Li, X.-L., Yao, S., Yan, B., Ren, H.-L., et al. (2014). Isolation and identification of antioxidant compounds in *Vaccinium bracteatum* Thunb. by UHPLC-Q-TOF LC/MS and their kidney damage protection. *J. Funct. Foods* 11, 62–70. doi: 10.1016/j.jff.2014.09.005

Zhang, Y., Fang, Z., Ye, X., and Pan, S. (2018). Identification of candidate genes involved in anthocyanin accumulation in the peel of jaboticaba (*Myrciaria cauliflora*) fruits by transcriptomic analysis. *Gene* 676, 202–213. doi: 10.1016/j.gene.2018.07.039

Zheng, Y., Chen, L., Liu, Y., Shi, L., Wan, S., and Wang, L. (2019). Evaluation of antimicrobial activity of water-soluble flavonoids extract from *Vaccinium bracteatum* Thunb. leaves. *Food Sci. Biotechnol.* 28, 1853–1859. doi: 10.1007/s10068-019-00634-4

Zhou, H., Liao, L., Xu, S., Ren, F., Zhao, J., Ogutu, C., et al. (2018). Two amino acid changes in the R3 repeat cause functional divergence of two clustered *MYB10* genes in peach. *Plant Mol. Biol.* 98, 169–183. doi: 10.1007/s11103-018-0773-2

Zhou, H., Lin-Wang, K., Wang, H., Gu, C., Dare, A. P., Espley, R. V., et al. (2015). Molecular genetics of blood-fleshed peach reveals activation of anthocyanin biosynthesis by NAC transcription factors. *Plant J.* 82, 105–121. doi: 10.1111/tpj.12792

Zhou, L., Xie, M., Yang, F., and Liu, J. (2020). Antioxidant activity of high purity blueberry anthocyanins and the effects on human intestinal microbiota. *LWT* 117:108621. doi: 10.1016/j.lwt.2019.108621

Zhou, Y., Zhou, H., Lin-Wang, K., Vimolmangkang, S., Espley, R. V., Wang, L., et al. (2014). Transcriptome analysis and transient transformation suggest an ancient duplicated MYB transcription factor as a candidate gene for leaf red coloration in peach. *BMC Plant Biol.* 14:1–13. doi: 10.1186/s12870-014-0388-y

Zifkin, M., Jin, A., Ozga, J. A., Zaharia, L. I., Schernthaner, J. P., Gesell, A., et al. (2011). Gene expression and metabolite profiling of developing highbush blueberry fruit indicates transcriptional regulation of flavonoid metabolism and activation of abscisic acid metabolism. *Plant Physiol.* 158, 200–224.

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.

frontiers in Plant Science

REVIEW
published: 10 December 2021
doi: 10.3389/fpls.2021.787049

Exploring the Diversity and Regulation of Apocarotenoid Metabolic Pathways in Plants

*Xiongjie Zheng, Yu Yang and Salim Al-Babili**

The BioActives Lab, Center for Desert Agriculture (CDA), Biological and Environment Science and Engineering (BESE), King Abdullah University of Science and Technology, Thuwal, Saudi Arabia

OPEN ACCESS

Edited by:
Lourdes Gómez-Gómez, University of Castilla-La Mancha, Spain

Reviewed by:
Maria Jesus Rodrigo, Institute of Agrochemistry and Food Technology, Spanish National Research Council (CSIC), Spain
Akemi Ohmiya, National Agriculture and Food Research Organization (NARO), Japan

***Correspondence:**
Salim Al-Babili
salim.babili@kaust.edu.sa

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

Received: *30 September 2021*
Accepted: *17 November 2021*
Published: *10 December 2021*

Citation:
Zheng X, Yang Y and Al-Babili S (2021) Exploring the Diversity and Regulation of Apocarotenoid Metabolic Pathways in Plants. Front. Plant Sci. 12:787049. doi: 10.3389/fpls.2021.787049

In plants, carotenoids are subjected to enzyme-catalyzed oxidative cleavage reactions as well as to non-enzymatic degradation processes, which produce various carbonyl products called apocarotenoids. These conversions control carotenoid content in different tissues and give rise to apocarotenoid hormones and signaling molecules, which play important roles in plant growth and development, response to environmental stimuli, and in interactions with surrounding organisms. In addition, carotenoid cleavage gives rise to apocarotenoid pigments and volatiles that contribute to the color and flavor of many flowers and several fruits. Some apocarotenoid pigments, such as crocins and bixin, are widely utilized as colorants and additives in food and cosmetic industry and also have health-promoting properties. Considering the importance of this class of metabolites, investigation of apocarotenoid diversity and regulation has increasingly attracted the attention of plant biologists. Here, we provide an update on the plant apocarotenoid biosynthetic pathway, especially highlighting the diversity of the enzyme carotenoid cleavage dioxygenase 4 (CCD4) from different plant species with respect to substrate specificity and regioselectivity, which contribute to the formation of diverse apocarotenoid volatiles and pigments. In addition, we summarize the regulation of apocarotenoid metabolic pathway at transcriptional, post-translational, and epigenetic levels. Finally, we describe inter- and intraspecies variation in apocarotenoid production observed in many important horticulture crops and depict recent progress in elucidating the genetic basis of the natural variation in the composition and amount of apocarotenoids. We propose that the illustration of biochemical, genetic, and evolutionary background of apocarotenoid diversity would not only accelerate the discovery of unknown biosynthetic and regulatory genes of bioactive apocarotenoids but also enable the identification of genetic variation of causal genes for marker-assisted improvement of aroma and color of fruits and vegetables and CRISPR-based next-generation metabolic engineering of high-value apocarotenoids.

Keywords: carotenoids, apocarotenoids, pigments, volatiles, natural variation, molecular regulation

INTRODUCTION

Carotenoids are lipophilic isoprenoid molecules consisting of a polyene backbone that generally contains 3–11 conjugated double bounds. They can be synthesized *de novo* by plants, algae, photosynthetic bacteria, and many nonphotosynthetic microorganisms (Hirschberg, 2001; Fraser and Bramley, 2004; DellaPenna and Pogson, 2006; Moise et al., 2014; Rodriguez-Concepcion et al., 2018; Zheng et al., 2020a). Humans cannot build carotenoids *de novo*, but need them as important antioxidants and as essential source of provitamin A, particularly if animal-derived food is not available (Fraser and Bramley, 2004; DellaPenna and Pogson, 2006; Giuliano, 2014; Nisar et al., 2015; Zheng et al., 2020b).

In many crops and other plant species, carotenoids confer their vivid yellow to red colors to fruits, flowers, tubers, and seeds (Yuan et al., 2015; Giuliano, 2017), which attract animal pollinators and distributors and are important feature for consumers, which indicate the quality of plant products and, hence, promote their marketability. More importantly, carotenoids are essential for plant photosynthesis, as they protect the photosynthetic apparatus from photooxidative damage and contribute to light-harvesting (Hashimoto et al., 2016; Rodriguez-Concepcion et al., 2018; Moreno et al., 2021). The conjugated double bonds make carotenoids themselves susceptible to photooxidation and other unspecific oxidative breakdown processes, such as lipoxygenase mediated co-oxidation, as well as to targeted, enzyme-catalyzed cleavage of certain double bonds (Zheng et al., 2020a; Moreno et al., 2021). The term "apocarotenoid" is used to define the cleavage products arising from these processes. The large number of double bonds in carotenoid backbone and of different carotenoids leads to a plentitude of apocarotenoids with various physico-chemical properties and biological functions. The compound family of apocarotenoids includes precursors of important phytohormones, that is, abscisic acid (ABA) and strigolactones (SLs), signaling molecules involved in plant growth, development, and resistance against pathogens and herbivores (Jia et al., 2018; Felemban et al., 2019; Watkins and Pogson, 2020; Wang et al., 2020b), as well as apocarotenoid pigments and volatiles responsible for color and aroma of fruits, flowers, or vegetables (Liang et al., 2021; Zheng et al., 2021).

Plant apocarotenoids and their derivatives have also important functions in promoting human health (Rodriguez-Concepcion et al., 2018). Humans contain two carotenoid cleavage enzymes, i.e., β-carotene cleavage oxygenase 1 and 2 (BCO1 and 2), which convert some carotenoids into a restricted set of apocarotenoids, including vitamin A (Harrison, 2005; Von Lintig, 2010); however, they cannot produce many health-promoting apocarotenoids, such as crocins and safranal (Frusciante et al., 2014; Ahrazem et al., 2016b). Given the importance of apocarotenoids for plants and humans, numerous studies have been performed on investigating the biological functions of apocarotenoids, elucidating their biosynthetic pathways and the regulation of their formation. Naturally occurring inter- or intraspecies apocarotenoid variation was observed in some important crops, such as citrus, tomato, and *Capsicum* species (Gao et al., 2019; Zheng et al., 2019; Zoccali et al., 2021), which would help to decipher apocarotenoid biosynthesis and its regulation, and genetic control of apocarotenoid diversity. Nowadays, the advancements in global profiling strategies, such as genomics, transcriptomics, and metabolomics, and multi-omics approaches have been very helpful for investigating the variation of apocarotenoid levels as well as the underlying genetic mechanisms. Several recent reviews cover the biosynthesis of apocarotenoid-derived hormones and signaling molecules and their roles in plant development and growth (Felemban et al., 2019; Fiorilli et al., 2019; Wang et al., 2020b; Moreno et al., 2021). In this review, we will mainly focus on the advances that have recently been made in understanding the regulation, metabolism, and genetic mechanism underlying apocarotenoid diversity, particularly apocarotenoid pigments and volatiles. We also depict the possible impact of understanding the mechanisms behind apocarotenoids diversity on accelerating molecular breeding and CRISPR-based next-generation metabolic engineering of apocarotenoids for crop improvement.

THE SIGNIFICANCE AND FUNCTION OF APOCAROTENOIDS

Non-enzymatic and enzymatic oxidative cleavage of carotenoids and further metabolism of thereby arising apocarotenoids lead to various biologically important plant metabolites, including phytohormones, pigments, volatiles, and signaling molecules (**Table 1**). Due to their important roles in many physiological, developmental processes, and plant-biotic environment interactions, the phytohormones ABA and SLs are best-known examples for biologically important plant apocarotenoids. ABA is widely distributed in nature and is common in cyanobacteria, algae, fungi, plants as well as in animals (Wang et al., 2020b). It plays important roles in many physiological and developmental processes, including the establishment of seed dormancy, root, and shoot growth and regulation of stomatal closure, and is a major component in plant abiotic and biotic stress response (Chen et al., 2020; Moreno et al., 2021). SLs were originally discovered as the host-derived rhizospheric chemical signal that induces seed germination in root parasitic plants, such as *Striga hermonthica*, which represent a major threat to global food security (Xie et al., 2010; Jamil et al., 2021). Later studies on arbuscular mycorrhization revealed the role of SLs as the plant-derived signal that induces hyphal branching in AM-fungi, which paves the way for host colonization (Cook et al., 1966; Lanfranco et al., 2018; Wang et al., 2020b). Within plants, SLs inhibit shoot branching and the formation of adventitious roots, stimulate internodes elongation, and increase stem thickness (Al-Babili and Bouwmeester, 2015; Moreno et al., 2021). Furthermore, SLs are involved in leaf and floral organ senescence, as well as in biotic and abiotic stress response (Al-Babili and Bouwmeester, 2015; Decker et al., 2017; Wang et al., 2020b). Besides these two apocarotenoid-derived phytohormones, recent studies on the formation and biological functions of apocarotenoids unraveled new carotenoid-derived plant signaling molecules and growth regulators. For example, Wang et al. (2019a) identified a new growth regulator named zaxinone, which promotes rice growth and modulates its hormone homeostasis. Exogenous application

TABLE 1 | Biosynthesis and biological functions of plant apocarotenoids and their derivatives.

Apocarotenoid	CCD enzymes involved	Biological functions	References
Strigolactones	CCD7; CCD8	Phytohormones involved in different developmental processes and rhizospheric signaling molecules inducing seed germination of root parasitic plants and hyphal branching of arbuscular mycorrhizal (AM) fungi. The latter is needed for establishing the AM symbiosis.	Alder et al., 2012; Zhang et al., 2014
Carlactone	CCD7; CCD8	The central intermediate in of strigolactone biosynthesis.	Alder et al., 2012; Bruno et al., 2014
3-Hydroxy-carlactone	CCD7; CCD8	A precursor of yet-unidentified strigolactones.	Baz et al., 2018; Yoneyama et al., 2020
Abscisic acids	NCEDs	A phytohormone involved in plant abiotic and biotic stress response as well as in many developmental processes including seed dormancy.	Tan et al., 2003; Moreno et al., 2021
Zaxinone	ZAS (CCD10)	A natural growth regulator that promotes rice root growth and is required for normal rice growth and development. It is involved in mycorrhization and regulates SL and ABA biosynthesis in *Arabidopsis*	Wang et al., 2019a, 2021; Ablazov et al., 2020; Zhong et al., 2020
Anchorene	Unknown	An apocarotenoid dialdehyde and natural plant metabolite involved in the formation of Arabidopsis anchor roots.	Jia et al., 2019
Iso-anchorene	Unknown	An isomer of anchorene, which inhibits primary root growth in Arabidopsis.	Jia et al., 2021a
β-Cyclocitral	Citrus CCD4b	A cyclic volatile apocarotenoid with a tobacco-like or grassy flavor. It acts as a signaling molecule regulating oxidative stress response, promotes Arabidopsis root growth, and induces plant resistance against *Spodoptera littoralis* as an herbivore-triggered signal;	Ramel et al., 2012; Dickinson et al., 2019; Gao et al., 2019; Mitra et al., 2021; Zheng et al., 2021
β-Ionone	CCD1, CCD4, or CCD7	A cyclic volatile apocarotenoid with a violet-like or fruity aroma. It attracts pollinators and seed dispersers and shows a strong repellent effect toward both the spider mite and flea beetle.	Simkin et al., 2004a; Cáceres et al., 2016; Liang et al., 2021
α-Ionone	CCD1/CCD4	A cyclic volatile apocarotenoid with a violet-like or fruity aroma. It attracts pollinators and seed dispersers and induces tomato resistance against western flower thrips;	Murata et al., 2020; Liang et al., 2021
β-Damascenone	Unknown	A cyclic volatile apocarotenoid with a honey-like, fruity aroma, and a super low odor threshold for perception. It contributes to the aroma of fruits and flowers.	Shi et al., 2020; Liang et al., 2021
6-Methyl-5-hepten-2-one	CCD1/4	A linear volatile apocarotenoid and an important aroma component of fruits and flowers	Shi et al., 2020; Liang et al., 2021
Geranylace-Tone	CCD1/4	A linear volatile apocarotenoid and an important aroma component of fruits and flowers	Simkin et al., 2004a; Lashbrooke et al., 2013
Loliolide or (–)-Loliolide	Unknown	An endogenous regulatory metabolite that mediates plant defense response to herbivores and enhances production of allelochemicals in the barnyardgrass-rice allelopathic interactions.	Murata et al., 2019; Li et al., 2020
Bixin	*Bixa* CCD4	A di-carboxylic monomethyl ester apocarotenoid that confers red color to *Bixa orellana* seeds. It is used in food and cosmetic industry and shows anti-inflammatory and antinociceptive activities.	Bouvier et al., 2003; Pacheco et al., 2019
Crocetin	*Crocus* CCD2, *Gardenia* CCD4a *Buddleja* CCD4.1/3	A natural apocarotenoid dicarboxylic acid mainly found in red stigmas of crocus flowers and gardenia fruits. It shows significant antitumorigenic effects in cell culture systems and animal models.	Gutheil et al., 2012; Frusciante et al., 2014; Ahrazem et al., 2016b, 2017
Crocins	*Crocus* CCD2, *Gardenia* CCD4a *Buddleja* CCD4.1/3	Natural water-soluble apocarotenoids that consist of a group of crocetin glycosides. They are mainly found in red stigmas of crocus flowers and gardenia fruit. They have different pharmacological effects, such as anti-inflammatory, antiaging, analgesic, and neuroprotective;	Frusciante et al., 2014; Ahrazem et al., 2016b, 2017; Bukhari et al., 2018
Picrocrocin	*Crocus* CCD2	A β-D-glucoside of 3-OH-β-cyclocitral. It is the precursor of safranal and responsible for the bitter taste of stigmas of crocus. It has pharmacological effects, such as reduction of the proliferation of human malignant melanoma, and is used in medical, food, and cosmetics industry.	Diretto et al., 2019; Martí et al., 2020
Safranal	*Crocus* CCD2	An apocarotenoid with pungent aroma and a major volatile component of crocus stigma.	Rezaee and Hosseinzadeh, 2013; Diretto et al., 2019
β-Citraurin	*Citrus* CCD4b	A C_{30} red apocarotenoid pigment responsible the red peel of citrus fruit.	Ma et al., 2013; Rodrigo et al., 2013; Zheng et al., 2019
β-Citraurinene	*Citrus* CCD4b	An C_{30} apocarotenoid pigment mainly found in red peel of citrus fruit.	Zheng et al., 2019

of zaxinone as well as of zaxinone mimics increased the growth of crown roots of rice seedlings and alleviated *Striga*-infestation in a *Striga*-susceptible rice cultivar, by reducing SL content and release (Wang et al., 2019a, 2020a), demonstrating its application potential for improving rice growth and combating root parasitic weeds. Very recently, Wang et al. (2021) also demonstrated that increased sugar uptake and metabolism is likely the major reason for growth-promoting activities of zaxinone in rice and that zaxinone effect on root cell division activity in root meristem and on the number of cortex cell layers is likely caused by modulating cytokinin content. The enzyme involved in zaxinone formation in rice, the carotenoid cleavage dioxygenase (CCD) Zaxinone synthase, is common in mycorrhizal plants, but absent in those, such as *Arabidopsis thaliana*, that do not build this symbiosis (Wang et al., 2019a). In contrast to rice, a recent study showed that zaxinone application to hydroponically grown *Arabidopsis* seedlings increased SL and ABA content in roots and did not improve growth. This result suggests that the activities and functions of zaxinone may differ between plant species and are likely coupled to their ability to establish arbuscular mycorrhizal fungi (AMF) symbiosis (Wang et al., 2019a; Ablazov et al., 2020; Moreno et al., 2021).

Oxidative cleavage of carotenoids at more than one double bond results in the production of various dialdehyde products. Jia et al. (2021a) identified a carotenoid-derived C_{10}-dialdehyde that was named anchorene, which specifically promoted the growth and development of Arabidopsis anchor roots. This type of roots emerges from the collet region, the root–hypocotyl junction, through modulating auxin homeostasis. The identity of anchorene as a natural carotenoid-derived metabolite was confirmed using mutants affected in carotenoid biosynthesis and chemical inhibitors of carotenoid biosynthesis and by establishing a LC-MS-based system for isolating and detecting of carotenoid-derived dialdehydes (Jia et al., 2019; Mi et al., 2019). The discovery of anchorene unraveled, for the first time, a biological function of carotenoid-derived dialdehydes in plant development.

β-Cyclocitral (β-CC) is a volatile apocarotenoid derived from the oxidation of carotenoids (e.g., β-carotene), which is widely present in nature and is common in various organisms ranging from cyanobacteria and fungi to plants. It was shown that β-CC is a retrograde signaling molecule regulating the expression of oxidative stress-responsive genes, resulting in acclimation of plants to high-light conditions (Ramel et al., 2012). β-cyclocitric acid (β-CCA), a direct oxidation product of β-CC, is also a stress signal that enhanced drought tolerance likely through a signaling pathway different from that of β-CC (D'alessandro et al., 2019). Because of its water-soluble character, β-CCA is more suitable for agricultural application and can be more easily used to increase drought tolerance of crops through irrigation or by spraying, compared with its precursor β-CC (Havaux, 2020). In addition to its role in oxidative stress response, β-CC is a root growth regulator. Exogenous application of β-CC promoted the growth of primary roots in Arabidopsis, tomato, and rice. It was also shown that β-CC alleviates the inhibitory effect of salt on the growth of rice roots (Dickinson et al., 2019). A recent study showed that β-CC induces plant resistance against *Spodoptera littoralis*, as an herbivore-triggered signal, and downregulates the methylerythritol-4-phosphate (MEP) pathway flux, which is responsible for plastid isoprenoid biosynthesis, through direct inhibition of 1-deoxy-*D*-xylulose-5-phosphate synthase (DXS), a rate-limiting enzyme of the MEP pathway (Mitra et al., 2021). Other volatile apocarotenoids are also involved in herbivore resistance or response. For example, β-ionone is released as herbivores-induced defense volatile in canola, *Brassica napus*, and shows a strong repellent effect toward both the spider mite and flea beetle (Cáceres et al., 2016). α-Ionone, a structural isomer of β-ionone, induces jasmonic acid (JA) independent resistance to western flower thrips in tomato and Arabidopsis, by reducing the survival rate of western flower thrips without exhibiting direct insecticidal activity (Murata et al., 2020). Loliolide, a C_{11}-terpene lactone, functions as an endogenous signaling metabolite that induces resistance to multiple herbivores, such as the two-spotted spider mite, common cutworm, and western flower thrips, possibly independent of JA signaling (Murata et al., 2019). Application of loliolide to tomato leaves reduced the survival rate of the two-spotted spider mite and larvae of the common cutworm, without exhibiting direct toxic activity against these herbivores (Murata et al., 2019). Moreover, loliolide [also referred to as (–)-loliolide] was found to act as a soil-borne signaling molecule that enhances the production of allelochemical, such as momilactone B and tricin, to coordinate barnyardgrass-rice allelopathic interactions (Kong et al., 2018; Li et al., 2020).

Besides their biological functions within plants, many apocarotenoid volatiles are recognized as important aroma components of different flowers and fruits of horticultural plants (Shi et al., 2020). For instance, linear volatile apocarotenoids, such as geranyl acetone, 6-methyl-5-hepten-2-one (MHO), and geranial, were identified as important aroma components in fruits of papaya, citrus, and blackberry (González-Mas et al., 2011; Jing et al., 2015; Reidel et al., 2016). Cyclic apocarotenoids, such as β-CC, β-ionone, and β-damascenone, have lower odor thresholds and stronger impacts on people's aroma perception, compared with linear volatiles (Goff and Klee, 2006). Therefore, cyclic apocarotenoids are recognized as more important volatiles with respect to the aromatic odor of fruits and flowers (Shi et al., 2020). The C_{13} cyclic β-damascenone has an attractive honey-like, floral, and fruity flavor and shows a very low odor threshold of 2 ng/L on people's aroma perception, hence, it has been frequently used as flavoring ingredient in fragrance industry (Liang et al., 2021). The cyclic volatiles α- and β-ionone have violet-like, woody, and fruity odor, while β-CC has a grassy, woody, tobacco-like odor. Both of them are widely used as flavor molecules in fragrances. In addition, these apocarotenoids volatiles also function as effective attractants for pollinators and seed dispersers (Liang et al., 2021).

Apart from being aroma compounds, apocarotenoids can contribute to the color of flowers and fruits and act as antioxidants. Indeed, some long-chain apocarotenoids and their derivatives are used as high-value food colorants, cosmetic agents, and antioxidants and have an important impact in promoting human health. β-citraurin (3-hydroxy-β-apo-8′-carotenal) and β-citraurinene are two major C_{30} apocarotenoid pigments found in *Citrus*, and their hyper-accumulation is responsible for the attractive red coloration of fruit peel in

some citrus species (Zheng et al., 2019). Bixin (C_{25}) is a further natural apocarotenoid pigment that confers red color to *B. orellana* seeds. It has been widely extracted from natural sources and used in food and cosmetic industry as a color additive (Giuliano et al., 2003b). Furthermore, Pacheco et al. (2019) demonstrated anti-inflammatory and antinociceptive activities of bixin, which were shown to be caused by reduction of neutrophil migration in simulated wound tests in rats. The proposed biosynthetic pathway of bixin is supposed to start with the cleavage of all-*trans*-lycopene into bixin aldehyde, followed by three sequential conversion steps catalyzed by the bixin aldehyde dehydrogenase and nor-bixin carboxyl methyltransferase (Bouvier et al., 2003). Saffron is one of the oldest natural food additives and worldwide the most expensive spice. It corresponds to the stigma of saffron flowers, which accumulates the 3-OH-β-cyclocitral glycoside picrocrocin, the C_{10} volatile apocarotenoid safranal, and the C_{20} pigment crocetin and its glycosides crocins. Crocetins and crocins are responsible for the red color, while picrocrocin and safranal give rise for the bitter taste and pungent aroma of saffron (Frusciante et al., 2014). These apocarotenoids are beneficial to health and are commonly used in food and cosmetics in many countries. Saffron apocarotenoids were also shown to have several pharmacological effects, such as anti-inflammatory, antiaging, analgesic, and neuroprotective (Bukhari et al., 2018).

THE METABOLIC BACKGROUND OF PLANT APOCAROTENOIDS STRUCTURAL DIVERSITY

Various Carotenoid Cleavage Dioxygenases Mediate Site-Specific Oxidative Tailoring of Carotenoids, Forming a Diverse Set of Apocarotenoids

In plants, the enzymatic cleavage of carotenoids at specific positions is generally catalyzed by CCDs, non-heme iron-dependent enzymes present in plants, animals, bacteria, and fungi (Giuliano et al., 2003a; Sui et al., 2013; Ahrazem et al., 2016a). In the last decades, different types of plant CCDs with distinct features have been identified and investigated. Arabidopsis contains five types of CCD enzymes, known as CCD1, CCD4, CCD7, CCD8, and 9-*cis*-epoxycarotenoid dioxygenase (NCED). There are five NCEDs in Arabidopsis, NCED2, NCED3, NCED5, NCED6, and NCED9, while the other types are represented by a single enzyme each (Auldridge et al., 2006; Walter and Strack, 2011). The five NCEDs catalyze the oxidative cleavage of 9-*cis*-violaxanthin and/or 9′-*cis*-neoxanthin to produce xanthoxin, the precursor of ABA (**Figure 1**; Schwartz et al., 1997; Tan et al., 2003; Al-Babili and Bouwmeester, 2015; Jia et al., 2018). This apocarotenoid product is exported from plastids, the site of carotenoid biosynthesis and most of plant cleavage enzymes, to the cytosol and converted to abscisic aldehyde by the enzyme short-chain dehydrogenase (SDR; González-Guzmán et al., 2002). The arising abscisic aldehyde is then oxidized by a molybdenum-dependent aldehyde oxidase (AAO) to produce ABA (Seo et al., 2000). Very recently, Jia et al. (2021b) found that β-apo-11-carotenoids (C_{15}), i.e., β-apo-11-carotenal, 9-*cis*-β-apo-11-carotenal, 3-OH-β-apo-11-carotenal, and 9-*cis*-3-OH-β-apo-11-carotenal, exert ABA-like biological functions in maintaining seed dormancy and inducing the expression of ABA-responsive genes. Further feeding experiments with labeled apocarotenoids, combined with analysis of physiological and transcriptional responses, showed that plants can synthesize ABA from C_{15} β-apo-11-carotenoids, which represents an alternative, zeaxanthin epoxidase-independent ABA biosynthetic pathway (Jia et al., 2021b). The genes involved in the hydroxylation, isomerization, and oxidation the C_{15} β-apo-11-carotenoids in this new ABA biosynthesis pathway remain elusive, which would be important for future exploration of the biological significance of this route.

The other four CCDs of Arabidopsis, CCD1, CCD4, CCD7, and CCD8 show different cleavage activities and biological functions (Zheng et al., 2020a). CCD7 and CCD8 are involved in SL biosynthesis. CCD7 cleaves 9-*cis*-β-carotene formed by the *cis-trans*-isomerase D27 (Alder et al., 2012; Bruno et al., 2014; Bruno and Al-Babili, 2016; Abuauf et al., 2018) to produce the apocarotenoids 9-*cis*-β-apo-10′-carotenal (C_{27}) and β-ionone (C_{13}). CCD8 catalyzes a combination of reactions converting 9-*cis*-β-apo-10′-carotenal into carlactone, the central intermediate in SL biosynthesis (**Figure 1**; Alder et al., 2012; Bruno et al., 2014; Bruno and Al-Babili, 2016; Abuauf et al., 2018). Thereby, it is assumed that CCD8 mediates isomerization, intramolecular rearrangement, and repeated oxygenation reactions (Alder et al., 2012; Bruno et al., 2017).

Plant CCD1 enzymes, characterized by their relaxed cleavage site and substrate specificity, cleave linear, monocyclic, and bicyclic carotenoid and apocarotenoid substrates at different double bonds (**Figure 2**), yielding various apocarotenoid aldehydes and ketones with different chain lengths (Schwartz et al., 2001; Vogel et al., 2008; Ilg et al., 2009, 2014). Some CCD1 cleavage products are volatile compounds that contribute to the aroma and flavor of fruits and flowers of important horticulture crops. For example, CCD1-catalyzed C9-C10 and/or C9′-C10′cleavage of carotenoids forms C_{14} dialdehyde and C_{13} ketone products (**Figure 2**), including β-ionone, β-ionophore, α-ionone, pseudoionone, and geranylacetone, which are flavor and fragrance volatiles in fruits or flowers of many different plant species, such as tomato (Simkin et al., 2004a), grape (Mathieu et al., 2005), melon (Ibdah et al., 2006), and petunia (Simkin et al., 2004b). CCD1 was also shown to cleave lycopene at the C5-C6 and/or C5′-C6′ position to yield MHO (**Figure 2**), an important aroma compound in tomato (Vogel et al., 2008). Additionally, the rice CCD1 enzyme showed an *in vitro* enzymatic activity that targets the C7-C8 bond of lycopene (**Figure 2**), generating a C_{10} flavor compound, geranial (Ilg et al., 2009). However, CCD1 enzymes are localized to cytoplasm, therefore they seem to cleave already destructed carotenoids transported to cytoplasm (i.e., apocarotenoids) rather than carotenoid substrates in plastid (Ilg et al., 2009, 2010). Consistent with this speculation, overexpression of rice CCD1 in Golden Rice did not lead to significant decrease of the carotenoid content in endosperm (Ilg et al., 2010).

CCD2, a specific CCD type that is restricted to the *Crocus* and *Freesia* genus of the *Iridaceae* family, is closely related to the cytoplasmic CCD1 subfamily, although it is localized in plastids (Frusciante et al., 2014; Ahrazem et al., 2016b; Fang et al., 2020). Crocus CCD2 was shown to cleave the C7-C8 and C7′-C8′ double bonds of zeaxanthin to form crocetin dialdehyde (C_{20}) and 3-OH-β-cyclocitral (C_{10}; Frusciante et al., 2014; Ahrazem et al., 2016b), which are subsequently converted to crocins and safranal (**Figure 3**), respectively, contributing to the color and aroma of saffron stigma (Demurtas et al., 2018; Diretto et al., 2019).

Carotenoid cleavage dioxygenase 4s are a further type of plastid-localized CCDs. They are involved in the production of apocarotenoid volatiles and their activity determines carotenoid content and apocarotenoid pigments accumulation (Ohmiya et al., 2006; Campbell et al., 2010; Brandi et al., 2011; Gonzalez-Jorge et al., 2013; Zhang et al., 2015). CCD4 enzymes from different plant species vary in their substrate specificity and regioselectivity (**Figure 3**; Mi and Al-Babili, 2019). In Arabidopsis and potato, CCD4 was shown to catalyze the cleavage of bicyclic carotenoids at the C9-C10 or C9′-C10′ double bond to produce C_{13} volatiles, and C_{27} apocarotenoids that are supposed to be further degraded to smaller metabolites (**Figure 3**; Bruno et al., 2016). This enzymatic activity of CCD4 decreases carotenoid content in various tissues of plant species, including potato tubers, Chrysanthemum and Brassica flowers, Japanese morning glory petals, peach fruits, and Arabidopsis seeds (Ohmiya et al., 2006; Campbell et al., 2010; Brandi et al., 2011; Falchi et al., 2013; Gonzalez-Jorge et al., 2013; Zhang et al., 2015; Watanabe et al., 2018; Han et al., 2019a). In contrast, the *Citrus* CCD4b targets at a different site, i.e., the C7-C8 or C7′-C8′ double bond, of β-carotene, β-cryptoxanthin, and zeaxanthin to form C_{30} apocarotenoid pigments, such as β-citraurin, together with C_{10} volatiles, such as β-CC in citrus peel (**Figure 3**; Ma et al., 2013; Rodrigo et al., 2013; Zheng et al., 2015, 2019, 2021). It is believed that *Citrus* CCD4b is also involved in the production of β-citraurinene and β-citraurol, two further *Citrus* fruit-specific C_{30} apocarotenoid pigments in citrus peels (Zheng et al., 2021). Indeed, C_{30} apocarotenoids are the main long-chain apocarotenoids in *CCD4b*-overexpressing citrus callus, reinforcing the major roles of CCD4b in the biosynthesis of C_{30} pigments in citrus peel (Zheng et al., 2019, 2021). The BdCCD4.1 and BdCCD4.3 from *Buddleja davidii*, showing the same enzymatic activity of crocus CCD2, cleave C7-C8, and C7′-C8′ double bonds of zeaxanthin *in vitro* and *in vivo*, to form crocetin dialdehyde and 3-hydroxy-β-cyclocitral (Ahrazem et al., 2017). A recent study reported another CCD4 type from *Gardenia jasminoides*

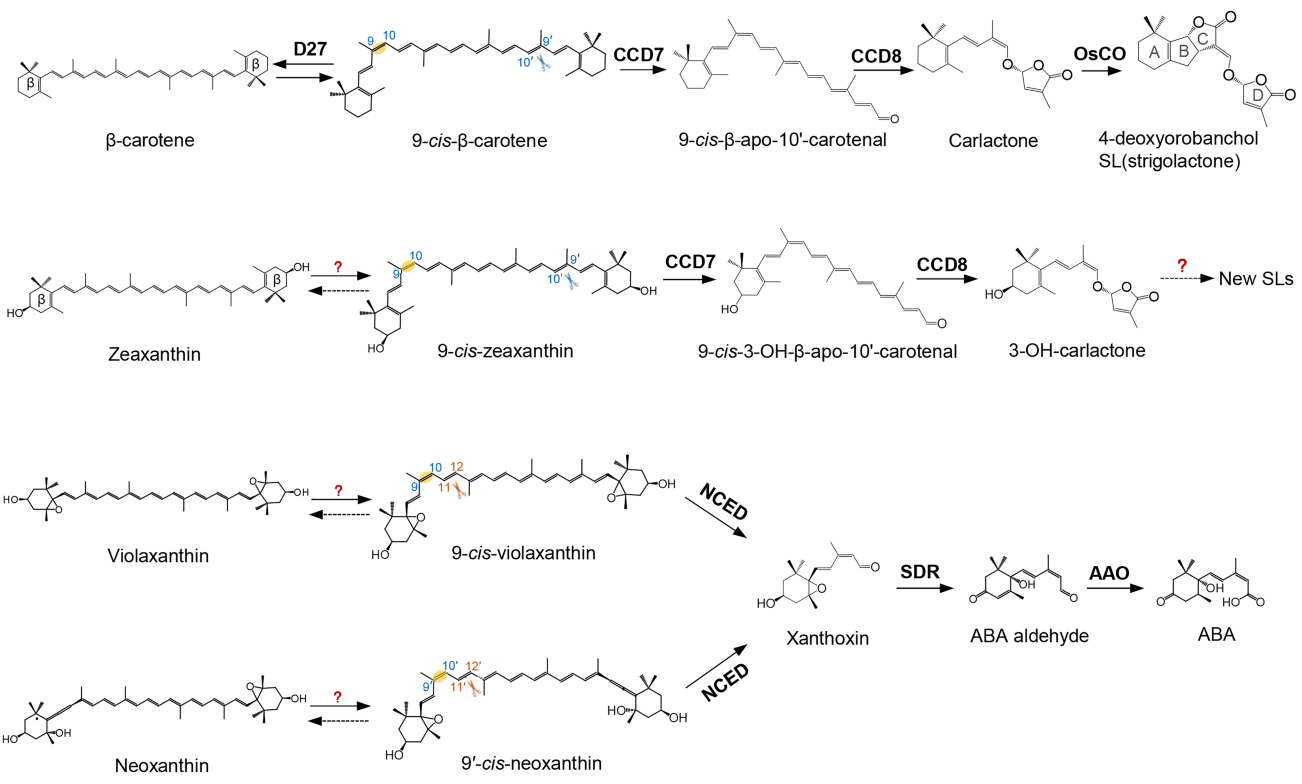


FIGURE 1 | Biosynthetic pathway of carotenoid-derived phytohormones. Names of carotenoids and apocarotenoids are shown in black. Names of carotenoid cleavage dioxygenases and apocarotenoid modifying enzymes are depicted in black boldface. D27 catalyzes the reversible isomerization of all-*trans*-β-carotene into 9-*cis*-β-carotene. 9-*cis*-β-carotene is then converted into strigolactones (SLs; 4-deoxyorobanchol is depicted as an example) by CCD7, CCD8, and CYP enzymes, such as the rice carlactone oxidase (OsCO), a homolog of the Arabidopsis MAX1. Violaxanthin and neoxanthin can be converted into 9-*cis*-violaxanthin and 9'-*cis*-neoxanthin, respectively. Nine-*cis*-Epoxycarotenoid Dioxygenases (NCEDs) cleave these two *cis* epoxy-xanthophylls to yield xanthoxin. Xanthoxin is then converted into abscisic acid (ABA) by short-chain dehydrogenase (SDR reductase) and abscisic aldehyde oxidase (AAO). Scissors indicate the double bond positions cleaved by carotenoid cleavage dioxygenase (CCD) enzymes.

that cleaves several substrates, including β-carotene, zeaxanthin, and lycopene, at both C7-C8 and C7′-C8′ double bonds, and is responsible for the production of crocetin and crocins in gardenia fruit (**Figure 3**; Xu et al., 2020b).

Zaxinone synthase (ZAS), also referred to as CCD10, is a new CCD-type that is missing in Arabidopsis and which has been recently unraveled in rice and maize (Wang et al., 2019a; Zhong et al., 2020). The rice ZAS was shown to cleave the 3-OH-apocarotenoids with different chain lengths, such as C_{30} all-*trans*-3-OH-apo-8′-carotenal, C_{27} all-*trans*-3-OH-apo-10′-carotenal, and C_{25} all-*trans*-3-OH-apo-12′-carotenal, at the C13-C14 double bond to yield a C_{18}-ketone product, zaxinone, that is characterized as a signaling molecule regulating plant growth and development (Wang et al., 2019a). It seems likely that the presence of this enzyme correlates with the ability to establish mycorrhizal symbiosis, as non-mycorrhizal plants, such as *Brassica* species, do not contain this type of CCDs.

Unspecific Oxidative Cleavage Processes in the Formation of Apocarotenoids

There are other enzymes mediating carotenoid cleavage, besides CCDs. For instance, lipoxygenases (LOXs) use carotenoids as co-substrates, however, resulting in unspecific oxidative cleavage of carotenoids. The fatty acid peroxyl radicals produced by LOXs enzymatic peroxidation of polyunsaturated fatty acids, such as linolenic, arachidonic, and linoleic acid, can attack carotenoids, leading to their degradation, as was shown *in vitro* (Jarén-Galán and Mínguez-Mosquera, 1999). Disruption and downregulation of *LOX* genes reduced carotenoid degradation during pasta processing and decreased the degradation of carotenoids in Golden Rice during storage, respectively (Carrera et al., 2007; Gayen et al., 2015), reinforcing the role of LOXs in carotenoid oxidative cleavage. Recently, by using pan-genome analysis, quantitative trait locus mapping and functional analysis, Gao et al. found that one tomato LOX (referred to as TomLoxC) is responsible for the production of apocarotenoid volatiles, including β-CC and β-ionone, in tomato fruits (Gao et al., 2019).

In addition to unspecific enzymatic degradation processes, carotenoids are subjected to unspecific non-enzymatic destructive oxidation triggered by reactive oxygen species (ROS; Ramel et al., 2012; Hou et al., 2016), which initiates metabolic pathways yielding various signaling molecules and/or volatile compounds. For example, singlet oxygen formed in chloroplasts, particularly at photosystem (PS) II under high-light condition, attacks all-*trans* β-carotene, giving rise to the formation of various

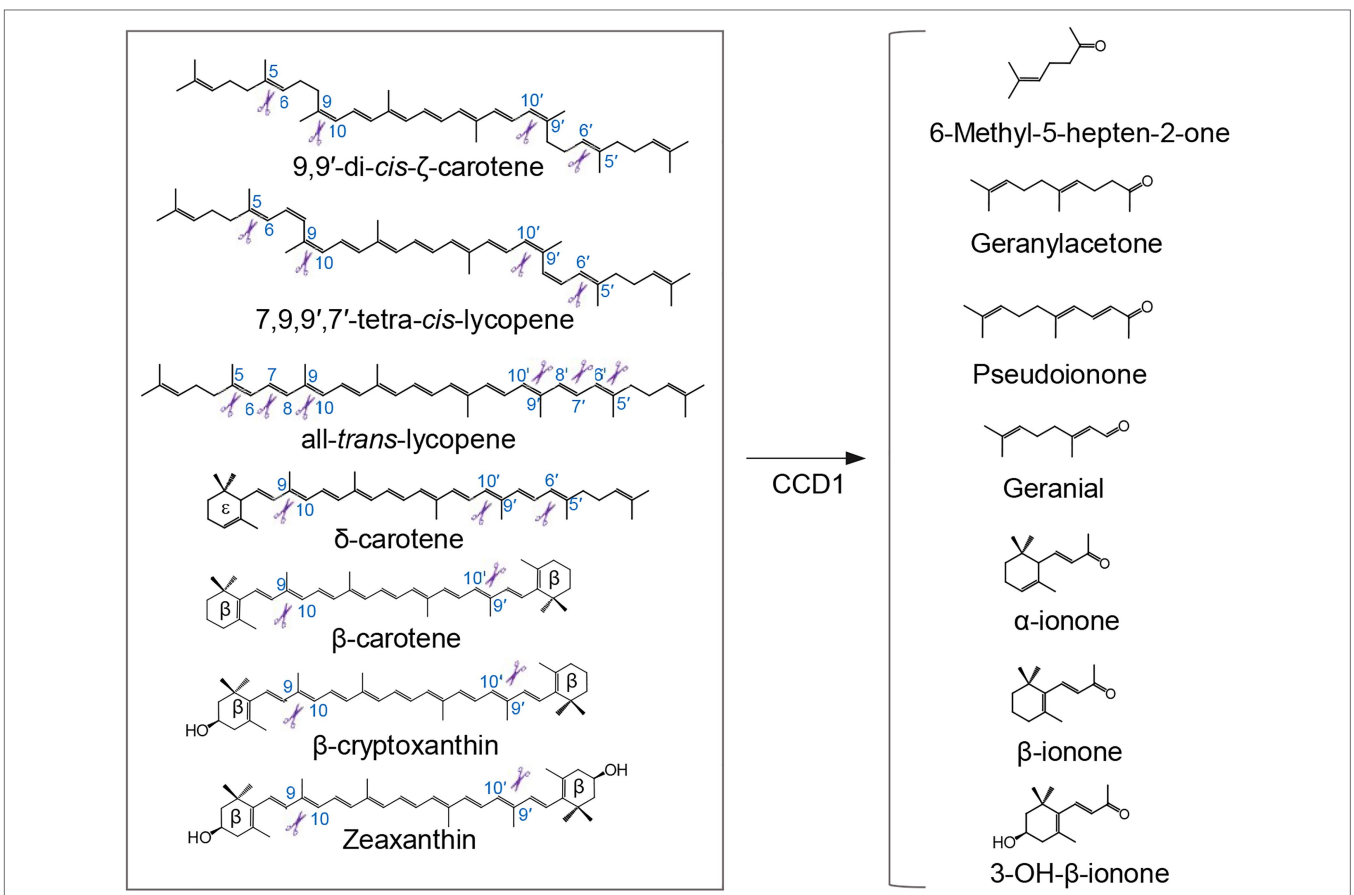


FIGURE 2 | Cleavage of carotenoid substrates by CCD1 enzyme and examples of generated volatiles. Plant CCD1 enzymes can cleave linear, monocyclic, and bicyclic carotenoid and apocarotenoid substrates at different double bonds in *E. coli* and/or *in vitro*, yielding various apocarotenoid volatiles, in addition to a plentitude of apocarotenoid dialdehydes (not shown). Scissors indicate the double bond positions cleaved by CCD1.

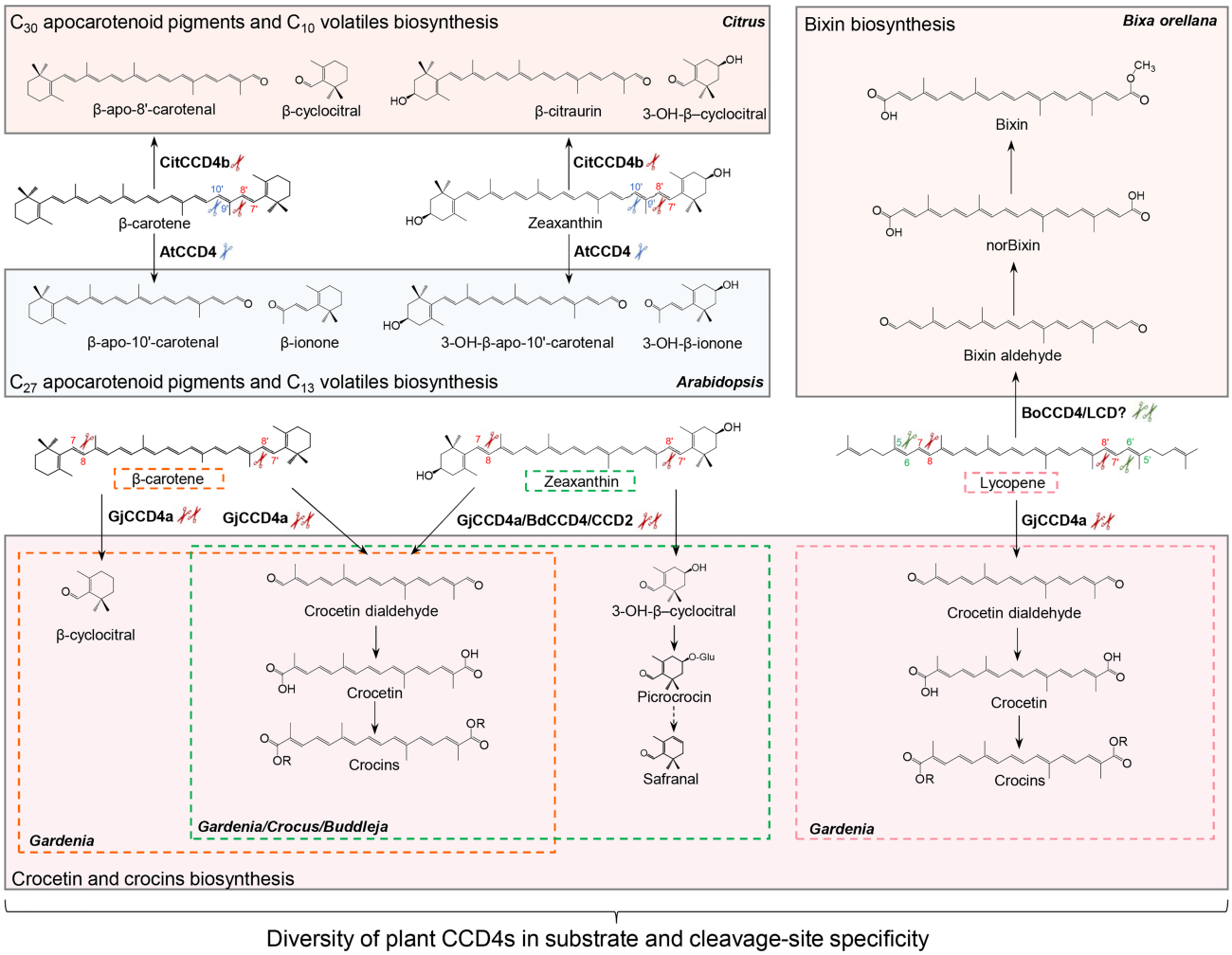


FIGURE 3 | Diversity of carotenoid cleavage dioxygenase 4 (CCD4) cleavage reactions. Names of apocarotenoids are shown in black. Names of carotenoid cleavage dioxygenases are in depicted black boldface. Scissors in different color indicate the positions cleaved by CCD4 or CCD2 enzymes from different plant species. Abbreviations: CitCCD4b, *Citrus* CCD4b; AtCCD4, *Arabidopsis thaliana* CCD4; GjCCD4a, *Gardenia jasminoides* CCD4a; BdCCD4, *Buddleja davidii* CCD4, CsCCD2, *Crocus sativus* CCD2; and BoCCD4/LCD, *Bixa orellana* CCD4/LCD.

apocarotenoid aldehydes, ketones, endoperoxides, epoxides, and lactones (Stratton et al., 1993; Hasegawa et al., 2004; Fiedor et al., 2005; Yamauchi et al., 2014), including the growth regulator and signaling molecule, β-CC (Ramel et al., 2012; Dickinson et al., 2019). It is assumed that linear volatile apocarotenoids, such as MHO, citral, geranylacetone, pseudoionone, and farnesylacetone, can also be generated through the unspecific oxidative degradation of carotenoid precursors, besides the possible involvement of CCD1 enzymes (Simkin et al., 2004a; Moreno et al., 2021).

MOLECULAR REGULATION OF APOCAROTENOID BIOSYNTHESIS

The production of plant apocarotenoids varies depending on developmental stage and tissue and is affected by environmental stimuli that lead to non-enzymatic cleavage of carotenoid. Due to their role as regulatory metabolites, the genesis of several apocarotenoids is tightly regulated at different levels, including transcriptional, post-translational, and epigenetic regulation.

Transcriptional Regulation

Many studies have reported that some CCD genes show tissue/organ-specific expression pattern that also depends on growth/developmental stages and environmental conditions (García-Limones et al., 2008; Chiou et al., 2010; Gonzalez-Jorge et al., 2013; Yuan et al., 2015; Zheng et al., 2019). Several transcriptional regulators are known to regulate the expression of genes related to apocarotenoid production in flowers and fruits of horticulture crops. For instance, sweet osmanthus (*Osmanthus fragrans*) OfWRKY3 and OfERF61 transcription factors were shown to directly bind to the CAACA and W-box elements in *OfCCD4* promoter, which results in an increase of *OfCCD4* expression

level (Han et al., 2016, 2019b). It is assumed that these two transcription factors are positive regulators of the biosynthesis of β-ionone, a key aromatic component in sweet osmanthus petals, which is formed through OfCCD4 cleavage reaction (Huang et al., 2009; Han et al., 2016, 2019b). In grapes, the expression level of *VvCCD4b* was shown to positively correlate with the content of several apocarotenoid volatiles, including MHO, β-ionone, and β-damascenone. *VvMADS4*, a MADS family transcription factor, directly binds to the promoter of *VvCCD4b* and represses its expression (Meng et al., 2020). Transcription factors that activate apocarotenoid biosynthetic genes, in some cases, also increase the expression of carotenoid biosynthetic genes responsible for the production of the carotenoid substrates. For instance, the citrus CsMADS6 was shown to directly bind to and to activate the promoters of citrus *CCD1*, as well as those of carotenoid biosynthetic genes, such as *LCYb1* (*Lycopene β-cyclase 1*), *PSY* (*phytoene synthase*), and *PDS* (*phytoene desaturase*; Lu et al., 2018). Very recently, Zhu et al. (2021) found that citrus CsERF061 enhances the expression levels of *CCD1*, *NCED3*, and *CCD4* by direct binding to their promoters and to those of the seven carotenoid biosynthetic genes, *PSY1*, *PDS*, *carotene isomerase* (*CRTISO*), *LCYB1*, *LCYB2*, *β-carotene hydroxylase* (*BCH*), and *zeaxanthin epoxidase* (*ZEP*), indicating the involvement of CsERF061 in a multi-target regulation of carotenoid and apocarotenoid metabolism.

The CsULT1 Ultrapetala transcription factor, identified in crocus, was found to positively regulate the gene expression of *CCD2* and *CCD4b*, as well as of *PSY*, *PDS*, and *BCH* genes (Ashraf et al., 2015).

Unspecific enzymatic cleavage of carotenoids is also regulated at transcriptional levels. In tomato, overexpression of *SlMYB75*, a MYB-type transcription factor, resulted in significant increase in the expression level of *LOXC* gene that is involved in unspecific enzymatic production of volatile apocarotenoids (e.g., β-CC). Yeast one-hybrid assay revealed that SlMYB75 can directly bind to MYBPZM and MYBPLANT *cis*-elements in *LOXC* promoter. Further dual-luciferase assays proved the role of SlMYB75 in activating the promoters of the *LOXC* (Jian et al., 2019).

It should be mentioned that although these transcription factors have been shown to regulate the expression levels of carotenoid and/or apocarotenoid biosynthetic genes, their effects on the apocarotenoid content and composition in target tissues are still poorly understood. Further quantification of apocarotenoid pigments and/or volatiles by LC-MS/GC-MS profiling should be performed in plants overexpressing or lacking these transcription factors, to confirm their contribution to apocarotenoid production.

As it can be expected, apocarotenoid-derived hormones biosynthesis is also regulated by transcriptional factors in various tissues/organs. For example, basic pentacysteine 1 (BPC1) can bind to the promoter of *CCD7*, and transient overexpression of *BPC1* repressed *CCD7's* promoter activity in roots of *Malus baccata* (Yue et al., 2015). The CDF4 transcription factor from Arabidopsis, a DOF family protein, is a positive regulator of ABA biosynthesis and leaf senescence, which increases the expression of *NCED2* and *NCED3* by directly binding to their promotors (Xu et al., 2020a). A NAC transcription factor, ATAF1, from Arabidopsis was shown to bind to the promoter of *NCED3 in vivo* and to activate its expression, which resulted in increased ABA levels. Moreover, a rice *NAC* transcription factor OsNAC2 also upregulates the expression of *OsNCED3* as well as *OsZEP1* and represses the expression of the ABA catabolic gene *OsABA8ox1*, which results in an increase of ABA levels and accelerated leaf senescence. In peach fruit, the ethylene response transcription factor PpERF3 was found to bind to the promoters of *PpNCED2/3* genes and to promote their expression, which results in an increased ABA content during fruit ripening (Wang et al., 2019b). In *Citrus reticulata* fruit, the R2R3-MYB transcriptional factor CrMYB68 is a negative regulator of *CrBCH2* and *CrNCED5* gene expression, contributing to the delay in the biosynthesis of ABA in "Green Ougan," a stay-green mutant of *C. reticulata* cv. Suavissima (Zhu et al., 2017). CrMYB68 also interacts with a novel NAC transcription factor, CrNAC036, during binding to the promoter of *CrNCED5*, causing a synergetic effect in inhibiting ABA biosynthesis in *C. reticulata* fruits (Zhu et al., 2020). Very recently, a citrus homeodomain leucine zipper I transcription factor called CsHB5 was found to positively regulate ABA accumulation and senescence, by directly binding to the promoters of carotenoid and ABA biosynthetic genes, including *BCH* and *NCED2*, and activating their transcription (Zhang et al., 2021). In Arabidopsis seeds, the basic helix-loop-helix transcription factor bHLH57 was shown to induce the expression of *NCED6* and *NCED9* through binding to the E-box (CANNTG)/G-box (CACGTG) motifs within their promoters, thereby leading to higher ABA levels that impose seed dormancy. In contrast, reversal of rdo5 (ODR1), a Homolog of Rice Seed Dormancy4, was shown to decrease seed dormancy by directly interacting with bHLH57 and inhibiting the bHLH57-mediated induction of *NCED6* and *NCED9* in the nucleus (Liu et al., 2020).

Post-translational Regulation

Post-translational regulation *via* protein-protein interactions provides an alternative mechanism to regulate CCD enzymatic activity in plants, thereby regulating the content and composition of apocarotenoids. PGM48, a plastoglobule (PG)-localized metallopeptidase identified in Arabidopsis, was found to positively regulate leaf senescence (Bhuiyan et al., 2016; Bhuiyan and van Wijk, 2017). PGM48 is the only protease present in PG proteomes. The protein level of PGM48 increased significantly during leaf senescence, whereas PG-localized CCD4 decreased. Indeed, overexpression of PGM48 in *Arabidopsis* dramatically reduced protein levels of CCD4. Moreover, it was shown that PGM48 can directly interact with CCD4, suggesting that CCD4 could be a substrate of PGM48 (Bhuiyan et al., 2016; Bhuiyan and van Wijk, 2017). The direct interaction and co-localization of PGM48 and CCD4 provided new insights into post-translational regulation of CCD enzymes; however, the role of PGM48 on the production of apocarotenoids is still elusive. Further apocarotenoids profiling and comprehensive characterization of the peptidase activity of PGM48 are still needed (Bhuiyan et al., 2016; Bhuiyan and van Wijk, 2017).

Plastid-localized ORANGE (OR), a DnaJ cysteine-rich protein, was found to act as a post-translational regulator of the protein level and enzymatic activity of PSY and a determinant of carotenoid accumulation (Zhou et al., 2015; Park et al., 2016; Chayut et al., 2017). Recent evidence showed that OR protein of sweet potato (IbOR) can interact not only with the known target PSY, but also with a new target, CCD4, suggesting that IbOR might also affect carotenoid cleavage as an important factor in carotenoid homeostasis (Park et al., 2020). However, the impact of this interaction on CCD4 protein level or activity is still elusive.

Epigenetic Modification

Epigenetic regulation is a further way deployed by plants to regulate apocarotenoid biosynthesis. For example, methylation analysis of *Oncidium OgCCD1* promoters of white-colored "Jade" and yellow-colored "Gower Ramsey" (GR) cultivars showed that a high level of DNA methylation in GR *OgCCD1* promoter, which is consistent with a silent *OgCCD1* and higher content of carotenoids in yellow floral tissues of GR (Chiou et al., 2010). In citrus callus, treatment with the DNA methyltransferase inhibitor 5-azacytidine resulted in a dramatic decrease of carotenoid content along with significant upregulation of *CpCCD1* gene (Xu et al., 2017). However, in this study, the tested promoter regions of *CpCCD1* were not demethylated by the 5-azacytidine effect, suggesting that methylation changes may indirectly regulate expression of *CpCCD1* (Xu et al., 2017). Chiou et al. (2010), found that the *Oncidium OgCCD1* promoter region of yellow-colored GR cultivar showed a higher level of DNA methylation than that in "Jade," which may result in the downregulation of *OgCCD1* in GR, thus leading to less conversion of carotenoids to colorless apocarotenoid in floral tissue. Moreover, a METHYLTRANSFERASE1 (SlMET1) acts as facilitator for *SlNCED* gene expression *via* directly altering methylation levels of *SlNCED* promoter region, thereby affecting ABA production in tomato epimutant *Colorless non-ripening* fruits (Yao et al., 2020).

GENETIC INSIGHTS INTO NATURAL VARIATION OF APOCAROTENOID

Many crop plants, especially horticulture crops, show large variation in quantity and types of apocarotenoids among different species and/or different natural accessions of a single species. For instance, different papaya accessions display big natural variation in accumulating MHO and β-ionone in fruits (Jing et al., 2015). Farnesyl acetone, α-ionone, and β-ionone show also high natural variation among orange-colored "Nairobi" and "Rothild," yellow-colored "Yellowstone," purple "Purple Haze," and white "Creme de Lite" carrot accessions (Yahyaa et al., 2013). As direct non-enzymatic degradation is supposed to be a major source of apocarotenoids, different concentration of carotenoids may be the reason for the diversity in apocarotenoid levels; however, more and more studies reported that some genetic factors can also determine the apocarotenoid variation, independent of carotenoid levels. For example, in citrus, α-ionone, geranial, and β-ionone are common in "Temple" hybrid mandarin fruit, but cannot be detected in "Murcott" hybrid that even contains significantly higher carotenoid amounts (Yu et al., 2015), In apricot fruit, β-ionone and dihydro-β-ionone are considered as two of the major aroma compounds for fruit quality by odor activity values (OVA) method. Contents of β-ionone and dihydro-β-ionone were found to be significantly higher in pulps of white apricot cultivars, such as "Luntaixiaobaixing" and "Baixing," than in those of the yellow "Hongyuxing" and "Danxing" cultivars that even accumulate higher carotenoid content (Xi et al., 2016). This finding suggests that the variation of apocarotenoid content in different apricot species is also genetically controlled, other than only determined by the concentration of the carotenoid precursors.

Technological advancements in high-throughput sequencing, metabolomics, and data mining have enabled the integration of various omics datasets, such as genomics, transcriptomics, and metabolomics (**Figure 4**). The multi-omics strategy coupled with genetic and functional analysis has facilitated exploiting the natural variation of apocarotenoid pigments and volatiles, which paves the way new discoveries in metabolism, function, and evolution evolutionary of plant apocarotenoids.

For instance, Gao et al. (2019) constructed a tomato pan-genome by using genome sequence datasets of 725 tomato accessions. Comparative analyses using this constructed tomato pan-genome identified a rare allele in cultivated tomatoes defined by promoter variation, i.e., a ~ 4-kb substitution in the promoter region, of *TomLoxC* gene encoding a 13-lipoxygenase. Further RNA-Seq data analysis showed that accessions containing both this rare allele captured by pan-genome and another allele present in the "Heinz 1706" genome exhibit significantly higher expression levels of *TomLoxC* gene than that in other accessions. Further metabolites, QTL and gene expression analysis provided strong evidence that *TomLoxC* variation is a genetic control for natural variation in the accumulation of nine apocarotenoid volatiles common in tomato fruits and might additionally play an important role in apocarotenoid volatiles biosynthesis. The levels of specific apocarotenoid volatiles were also significantly decreased in transgenic tomato fruits with reduced expression of *TomLoxC* and in two knock-out Arabidopsis mutants of *AtLOX2*, the closest *TomLoxC* homolog. These experiments further confirm that *LOX* is a genetic determinant for natural variation of apocarotenoid formation and also reinforce the significance of alternative apocarotenoid biosynthesis route *via* unspecific oxidation process, besides CCD-mediated cleavage reaction. Apocarotenoid volatiles are important quality trait for tomato fruit and flavor (Tieman et al., 2017; Gao et al., 2019). However, this *TomLoxC* promoter allele captured by pan-genome displays strong negative selection during domestications. This may be due to the primary focus of modern breeding on tomato yield and on resistance to stresses and long shelf-life, which did not consider aroma quality traits.

In addition to volatiles, several recent studies performed apocarotenoid pigments profiling in fruits and related natural variation in some important horticulture crops, such as *Citrus* and *Capsicum* species. Citrus fruits show peel color variation among different cultivars. Besides carotenoid content and composition, the production of C_{30} apocarotenoid pigments, such as β-citraurin, is also critical for peel pigmentation of citrus mature fruit (Ma et al., 2013; Rodrigo et al., 2013; Zheng et al., 2015, 2019, 2021). Indeed, the content of C_{30} apocarotenoid

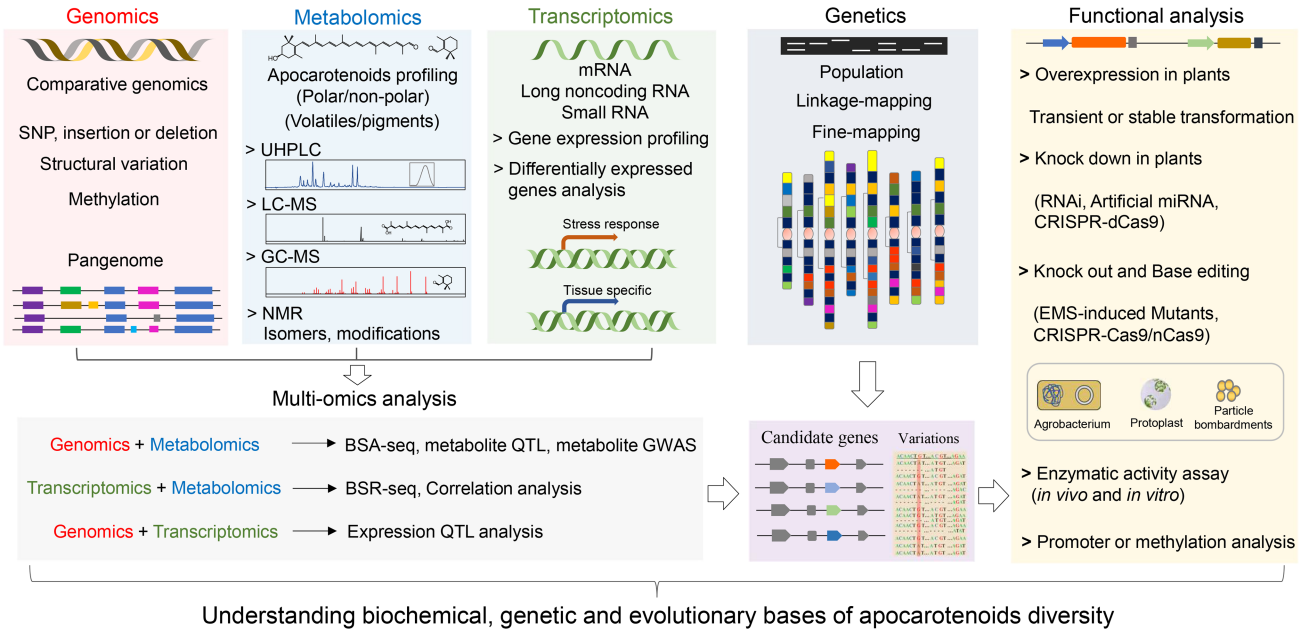


FIGURE 4 | A schematic workflow for multi-omics strategy coupled with genetic and functional analysis to elucidate the molecular mechanisms underlying apocarotenoid variation. The multi-omics strategy combines genomics, transcriptomics, and metabolomics approaches. Functional analysis includes overexpression, knockdown, and knockout of candidate genes, which can be performed by using Agrobacterium-, protoplast-, or particle bombardments-mediated transformation strategies. For structural genes, such as *CCD* genes, evaluation of the activity of encoded enzymes (*in vivo* and *in vitro*) can be performed to confirm the functional variation. Promoter and methylation analysis are usually conducted to explore genetic bases underlying expression variation of candidate genes. GWAS, genome-wide association study; QTL, quantitative trait locus; eQTL, expression QTL; and mQTL, metabolites QTL. BSA-Seq, bulked-segregant analysis (BSA) coupled to whole-genome sequencing. BSR-Seq, BSA coupled to RNA-seq (RNA-sequencing).

pigments, i.e., β-citraurin and β-citraurinene, and their ratio to total carotenoids in citrus peel showed the high correlation with the color index of red-peeled progenies in a F_1 pseudo-testcross population derived from a cross between a red-peeled *Citrus reticulata* and a yellow-peeled *Poncirus trifoliata* (Zheng et al., 2019). By using mQTL, eQTL, and BSR-seq analysis, this study showed that a 5′ *cis*-regulatory mutation of *Citrus CCD4b* (*CitCCD4b*) is a major genetic determinant of natural variation in the accumulation of the C_{30} pigments β-citraurin and β-citraurinene. This study demonstrated that CitCCD4b enzyme is involved in the biosynthesis of β-citraurinene, which indicates a novel reaction in modification of aldehyde group of β-citraurin. The presence of a specific SNP in the MITE of *CitCCD4b* promoter is strongly correlated with the high expression level of *CitCCD4b* among progenies of pseudo-testcross population as well as within 115 different citrus accessions, resulting in the high-accumulation of C_{30} apocarotenoid pigments responsible for the red coloration of citrus peel. A recent study also found that a red-peeled mutant of "Huyou" (*Citrus changshanensis*) contains much higher content of the C_{30} apocarotenoid β-citraurin, accompanied by 100 times higher expression level of *CitCCD4b*, compared to yellow-peeled ordinary "Huyou" (Luan et al., 2020b), which also correlated with the above-mentioned putative enhancer SNP in the *CitCCD4b* promoter (Zheng et al., 2019; Luan et al., 2020a). Very recently, it was found that the higher transcript level of *CitCCD4b* in red peel of F1 hybrids also accompanied the higher content of the two C_{10} apocarotenoid volatiles, β-CC and 3-OH-β-cyclocitral, in comparison to that in the yellow peel of three other progenies (Zheng et al., 2021). This result suggests a dual role of *CitCCD4b* expression variation in the natural variation of both apocarotenoid pigments and volatiles among different citrus species and in fruit pigmentation and aroma formation.

However, in other crops, such as peach and *Brassica* species, mutations in coding sequence of CCD4 enzyme enhanced the carotenoid pigments in tissues/organs, due to the capability of this enzyme in cleaving colorful carotenoids into C_{27} apocarotenoids that are catabolized into colorless smaller metabolites (Falchi et al., 2013; Mi and Al-Babili, 2019). Recent Studies demonstrated that CRISPR/Cas9-mediated disruption of CCD4 gene in *Ipomoea nil* and carrot resulted in a remarkable enhancement of carotenoids and led to pale-yellow phenotype in petal of *I. nil* and yellow phenotype in the taproot of a white-colored carrot variety, respectively (Watanabe et al., 2018; Li et al., 2021). The corresponding CCD4 enzymes have also been reported to be responsible for the formation of apocarotenoid volatiles in tissues/organs, such as in peach fruit (Brandi et al., 2011). The crucial step of future research is to determine whether the natural genetic variation of such *CCD4*s is also linked to the diversity of apocarotenoid volatiles among different accessions, which may be utilized in molecular breeding for improvement of fruit/vegetable flavor.

Apart from apocarotenoid pigments and volatiles, apocarotenoid-derived hormones also show natural variation among different accessions. For example, at low water potential, *Arabidopsis* accession "Shahdara" (Sha) showed less accumulation of ABA

than that in "Columbia" (Col) or "Landsberg *erecta*" (Ler). Genetic analysis of Ler × Sha recombinant inbred lines (RILs) and complementation experiments demonstrated that *NCED3* variation is responsible for the difference in ABA accumulation between Sha and Ler accession. Further sequence analysis, site-directed mutagenesis and structural observations revealed that nonsynonymous substitutions of *NCED3* coding sequence found in Sha are associated with decreased ABA accumulation and altered NCED3 post-translational processing, which may affect NCED3 activity (Kalladan et al., 2019). The rice cultivars "Bala" and "Azucena" showed large difference in SL biosynthesis that is responsible for their degree of tillering and distinct susceptibility to Striga infection (Cardoso et al., 2014). By using a "Bala" × "Azucena" F6 RIL population, Cardoso et al. (2014) identified a major QTL, a rearrangement of a 51- to 59-kbp stretch between 28.9 and 29.0 Mbp of chromosome 1, which explains most of the variation in the SL contents of rice exudates. This genomic rearrangement leads to deletion of two *AtMAX1* orthologs and is associated with low SL contents in 367 rice cultivars.

CONCLUDING REMARKS AND FUTURE OUTLOOK

Apocarotenoids are not just carotenoid degradation products but play important functions as antioxidants, pharmaceuticals, scents, and pigments, beside their role as precursors of plant hormones, signaling molecules, and growth regulators. In this review, we have briefly introduced the significance of apocarotenoid compounds for human health and their biological functions that are critical for plant development, biotic interactions, and stress response and provided an update on recently identified biosynthetic pathways and their regulation. More and more studies have shown that exogenous application of apocarotenoids is able to affect plant growth and development, as well as both abiotic and biotic stress resistance in many plants, which uncovered their biological function as regulatory metabolites. Nevertheless, the correlation between natural variation in the content of these apocarotenoids and plant phenotypes remains elusive. It is still unknown why apocarotenoids with slightly different structures have different activities and diverse functions in plants and some apocarotenoids, i.e., zaxinone, exerts contradictory effects depending on plant species. We proposed that the exploration of biochemical and genetic reasons of apocarotenoid metabolic diversity/variation could accelerate exploring currently unknown/unclear bioactive apocarotenoids pathway that could help understand their contributions to crops end-phenotypes, as well as their regulation and modifications that are still poorly understood. To date, the progress in determining natural genetic variation mechanisms and its molecular background is limited to only few apocarotenoid volatiles and pigments, while that of other important apocarotenoids, such as crocins, bixin, zaxinone, and anchorene, and apocarotenoid-derived phytohormones, is still poorly characterized. The rapid development of next-generation sequencing strategies provides reliable reference genomic sequences that are now available for most crop plants. We anticipate that mass spectrometry apocarotenoids profiling coupled with multi-omics strategies, including quantitative trait locus (QTL) mapping, genome-wide association (GWA), and RNA-seq analysis, will pave the way for uncovering new QTLs for apocarotenoid metabolite traits and identifying their underlying genes in various important crops (**Figure 4**). Advancement of knowledge in this area will enable to identify DNA marker-apocarotenoid traits associations for marker-assisted improvement of crops agronomic traits, such as aroma and color of fruits and flowers, related to apocarotenoid content, but also provide new metabolic engineering and/or CRISPR editing target for improvements of high-value beneficial apocarotenoids in crops in a low-cost, efficient, stable, and environment-friendly manner, to satisfy increasing demands for better crop quality in the future.

AUTHOR CONTRIBUTIONS

XZ organized and drafted the manuscript with substantial input from YY and designed and prepared the figures with input from YY. YY and XZ prepared the table. SA-B supervised the writing of the review and was involved in its further editing and revision. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by baseline funding and Competitive Research Grants (CRG 2017 and CRG 2020) given to SA-B from King Abdullah University of Science and Technology (KAUST).

ACKNOWLEDGMENTS

We thank Jianing Mi for reading and correcting the manuscript.

REFERENCES

Ablazov, A., Mi, J., Jamil, M., Jia, K.-P., Wang, J. Y., Feng, Q., et al. (2020). The apocarotenoid zaxinone is a positive regulator of strigolactone and abscisic acid biosynthesis in *Arabidopsis* roots. *Front. Plant Sci.* 11:578. doi: 10.3389/fpls.2020.00578

Abuauf, H., Haider, I., Jia, K.-P., Ablazov, A., Mi, J., Blilou, I., et al. (2018). The *Arabidopsis* DWARF27 gene encodes an all-*trans*–/9-*cis*-β-carotene isomerase and is induced by auxin, abscisic acid and phosphate deficiency. *Plant Sci.* 277, 33–42. doi: 10.1016/j.plantsci.2018.06.024

Ahrazem, O., Diretto, G., Argandoña, J., Rubio-Moraga, Á., Julve, J. M., Orzáez, D., et al. (2017). Evolutionarily distinct carotenoid cleavage dioxygenases are responsible for crocetin production in *Buddleja davidii*. *J. Exp. Biol.* 68, 4663–4677. doi: 10.1093/jxb/erx277

Ahrazem, O., Gómez-Gómez, L., Rodrigo, M. J., Avalos, J., and Limón, M. C. (2016a). Carotenoid cleavage oxygenases from microbes and photosynthetic organisms: features and functions. *Int. J. Mol. Sci.* 17:1781. doi: 10.3390/ijms17111781

Ahrazem, O., Rubio-Moraga, A., Berman, J., Capell, T., Christou, P., Zhu, C., et al. (2016b). The carotenoid cleavage dioxygenase CCD2 catalysing the

synthesis of crocetin in spring crocuses and saffron is a plastidial enzyme. *New Phytol.* 209, 650–663. doi: 10.1111/nph.13609

Al-Babili, S., and Bouwmeester, H. J. (2015). Strigolactones, a novel carotenoid-derived plant hormone. *Annu. Rev. Plant Biol.* 66, 161–186. doi: 10.1146/annurev-arplant-043014-114759

Alder, A., Jamil, M., Marzorati, M., Bruno, M., Vermathen, M., Bigler, P., et al. (2012). The path from β-carotene to carlactone, a strigolactone-like plant hormone. *Science* 335, 1348–1351. doi: 10.1126/science.1218094

Ashraf, N., Jain, D., and Vishwakarma, R. A. (2015). Identification, cloning and characterization of an ultrapetala transcription factor CsULT1 from crocus: a novel regulator of apocarotenoid biosynthesis. *BMC Plant Biol.* 15:25. doi: 10.1186/s12870-015-0423-7

Auldridge, M. E., Block, A., Vogel, J. T., Dabney-Smith, C., Mila, I., Bouzayen, M., et al. (2006). Characterization of three members of the *Arabidopsis* carotenoid cleavage dioxygenase family demonstrates the divergent roles of this multifunctional enzyme family. *Plant J.* 45, 982–993. doi: 10.1111/j.1365-313X.2006.02666.x

Baz, L., Mori, N., Mi, J., Jamil, M., Kountche, B. A., Guo, X., et al. (2018). 3-Hydroxycarlactone, a novel product of the strigolactone biosynthesis core pathway. *Mol. Plant* 11, 1312–1314. doi: 10.1016/j.molp.2018.06.008

Bhuiyan, N. H., Friso, G., Rowland, E., Majsec, K., and Van Wijk, K. J. (2016). The plastoglobule-localized metallopeptidase PGM48 is a positive regulator of senescence in *Arabidopsis thaliana*. *Plant Cell* 28, 3020–3037. doi: 10.1105/tpc.16.00745

Bhuiyan, N. H., and Van Wijk, K. J. (2017). Functions and substrates of plastoglobule-localized metallopeptidase PGM48. *Plant Signal. Behav.* 12, 3020–3037. doi: 10.1080/15592324.2017.1331197

Bouvier, F., Dogbo, O., and Camara, B. (2003). Biosynthesis of the food and cosmetic plant pigment bixin (annatto). *Science* 300, 2089–2091. doi: 10.1126/science.1085162

Brandi, F., Bar, E., Mourgues, F., Horváth, G., Turcsi, E., Giuliano, G., et al. (2011). Study of 'Redhaven' peach and its white-fleshed mutant suggests a key role of CCD4 carotenoid dioxygenase in carotenoid and norisoprenoid volatile metabolism. *BMC Plant Biol.* 11, 1–14. doi: 10.1186/1471-2229-11-24

Bruno, M., and Al-Babili, S. (2016). On the substrate specificity of the rice strigolactone biosynthesis enzyme DWARF27. *Planta* 243, 1429–1440. doi: 10.1007/s00425-016-2487-5

Bruno, M., Hofmann, M., Vermathen, M., Alder, A., Beyer, P., and Al-Babili, S. (2014). On the substrate-and stereospecificity of the plant carotenoid cleavage dioxygenase 7. *FEBS Lett.* 588, 1802–1807. doi: 10.1016/j.febslet.2014.03.041

Bruno, M., Koschmieder, J., Wuest, F., Schaub, P., Fehling-Kaschek, M., Timmer, J., et al. (2016). Enzymatic study on AtCCD4 and AtCCD7 and their potential to form acyclic regulatory metabolites. *J. Exp. Biol.* 67, 5993–6005. doi: 10.1093/jxb/erw356

Bruno, M., Vermathen, M., Alder, A., Wüst, F., Schaub, P., Van Der Steen, R., et al. (2017). Insights into the formation of carlactone from in-depth analysis of the CCD 8-catalyzed reactions. *FEBS Lett.* 591, 792–800. doi: 10.1002/1873-3468.12593

Bukhari, S. I., Manzoor, M., and Dhar, M. (2018). A comprehensive review of the pharmacological potential of *Crocus sativus* and its bioactive apocarotenoids. *Biomed. Pharmacother.* 98, 733–745. doi: 10.1016/j.biopha.2017.12.090

Cáceres, L., Lakshminarayan, S., Yeung, K.-C., Mcgarvey, B., Hannoufa, A., Sumarah, M., et al. (2016). Repellent and attractive effects of α-, β-, and dihydro-β-ionone to generalist and specialist herbivores. *J. Chem. Ecol.* 42, 107–117. doi: 10.1007/s10886-016-0669-z

Campbell, R., Ducreux, L. J., Morris, W. L., Morris, J. A., Suttle, J. C., Ramsay, G., et al. (2010). The metabolic and developmental roles of carotenoid cleavage dioxygenase 4 from potato. *Plant Physiol.* 154, 656–664. doi: 10.1104/pp.110.158733

Cardoso, C., Zhang, Y., Jamil, M., Hepworth, J., Charnikhova, T., Dimkpa, S. O., et al. (2014). Natural variation of rice strigolactone biosynthesis is associated with the deletion of two MAX1 orthologs. *Proc. Natl. Acad. Sci. U. S. A.* 111, 2379–2384. doi: 10.1073/pnas.1317360111

Carrera, A., Echenique, V., Zhang, W., Helguera, M., Manthey, F., Schrager, A., et al. (2007). A deletion at the Lpx-B1 locus is associated with low lipoxygenase activity and improved pasta color in durum wheat (*Triticum turgidum ssp. durum*). *J. Cereal Sci.* 45, 67–77. doi: 10.1016/j.jcs.2006.07.001

Chayut, N., Yuan, H., Ohali, S., Meir, A., Sa'ar, U., Tzuri, G., et al. (2017). Distinct mechanisms of the ORANGE protein in controlling carotenoid flux. *Plant Physiol.* 173, 376–389. doi: 10.1104/pp.16.01256

Chen, K., Li, G. J., Bressan, R. A., Song, C. P., Zhu, J. K., and Zhao, Y. (2020). Abscisic acid dynamics, signaling, and functions in plants. *J. Integr. Plant Biol.* 62, 25–54. doi: 10.1111/jipb.12899

Chiou, C.-Y., Pan, H.-A., Chuang, Y.-N., and Yeh, K.-W. (2010). Differential expression of carotenoid-related genes determines diversified carotenoid coloration in floral tissues of *Oncidium* cultivars. *Planta* 232, 937–948. doi: 10.1007/s00425-010-1222-x

Cook, C., Whichard, L. P., Turner, B., Wall, M. E., and Egley, G. H. (1966). Germination of witchweed (*Striga lutea* lour.): isolation and properties of a potent stimulant. *Science* 154, 1189–1190. doi: 10.1126/science.154.3753.1189

D'alessandro, S., Mizokami, Y., Legeret, B., and Havaux, M. (2019). The apocarotenoid β-cyclocitric acid elicits drought tolerance in plants. *Iscience* 19, 461–473. doi: 10.1016/j.isci.2019.08.003

Decker, E. L., Alder, A., Hunn, S., Ferguson, J., Lehtonen, M. T., Scheler, B., et al. (2017). Strigolactone biosynthesis is evolutionarily conserved, regulated by phosphate starvation and contributes to resistance against phytopathogenic fungi in a moss, *Physcomitrella patens*. *New Phytol.* 216, 455–468. doi: 10.1111/nph.14506

DellaPenna, D., and Pogson, B. J. (2006). Vitamin synthesis in plants: tocopherols and carotenoids. *Annu. Rev. Plant Biol.* 57, 711–738. doi: 10.1146/annurev.arplant.56.032604.144301

Demurtas, O. C., Frusciante, S., Ferrante, P., Diretto, G., Azad, N. H., Pietrella, M., et al. (2018). Candidate enzymes for saffron crocin biosynthesis are localized in multiple cellular compartments. *Plant Physiol.* 177, 990–1006. doi: 10.1104/pp.17.01815

Dickinson, A. J., Lehner, K., Mi, J., Jia, K.-P., Mijar, M., Dinneny, J., et al. (2019). β-Cyclocitral is a conserved root growth regulator. *Proc. Natl. Acad. Sci. U. S. A.* 116, 10563–10567. doi: 10.1073/pnas.1821445116

Diretto, G., Ahrazem, O., Rubio-Moraga, Á., Fiore, A., Sevi, F., Argandoña, J., et al. (2019). UGT709G1: a novel uridine diphosphate glycosyltransferase involved in the biosynthesis of picrocrocin, the precursor of safranal in saffron (*Crocus sativus*). *New Phytol.* 224, 725–740. doi: 10.1111/nph.16079

Falchi, R., Vendramin, E., Zanon, L., Scalabrin, S., Cipriani, G., Verde, I., et al. (2013). Three distinct mutational mechanisms acting on a single gene underpin the origin of yellow flesh in peach. *Plant J.* 76, 175–187. doi: 10.1111/tpj.12283

Fang, Q., Li, Y., Liu, B., Meng, X., Yang, Z., Yang, S., et al. (2020). Cloning and functional characterization of a carotenoid cleavage dioxygenase 2 gene in safranal and crocin biosynthesis from Freesia hybrida. *Plant Physiol. Biochem.* 154, 439–450. doi: 10.1016/j.plaphy.2020.06.035

Felemban, A., Braguy, J., Zurbriggen, M. D., and Al-Babili, S. (2019). Apocarotenoids involved in plant development and stress response. *Front. Plant Sci.* 10:1168. doi: 10.3389/fpls.2019.01168

Fiedor, J., Fiedor, L., Haeßner, R., and Scheer, H. (2005). Cyclic endoperoxides of β-carotene, potential pro-oxidants, as products of chemical quenching of singlet oxygen. Biochim. *Biophys. Acta, Bioenerg.* 1709, 1–4. doi: 10.1016/j.bbabio.2005.05.008

Fiorilli, V., Wang, J. Y., Bonfante, P., Lanfranco, L., and Al-Babili, S. (2019). Apocarotenoids: old and new mediators of the arbuscular mycorrhizal symbiosis. *Front. Plant Sci.* 10:1186. doi: 10.3389/fpls.2019.01186

Fraser, P. D., and Bramley, P. M. (2004). The biosynthesis and nutritional uses of carotenoids. *Prog. Lipid Res.* 43, 228–265. doi: 10.1016/j.plipres.2003.10.002

Frusciante, S., Diretto, G., Bruno, M., Ferrante, P., Pietrella, M., Prado-Cabrero, A., et al. (2014). Novel carotenoid cleavage dioxygenase catalyzes the first dedicated step in saffron crocin biosynthesis. *Proc. Natl. Acad. Sci. U. S. A.* 111, 12246–12251. doi: 10.1073/pnas.1404629111

Gao, L., Gonda, I., Sun, H., Ma, Q., Bao, K., Tieman, D. M., et al. (2019). The tomato pan-genome uncovers new genes and a rare allele regulating fruit flavor. *Nat. Genet.* 51, 1044–1051. doi: 10.1038/s41588-019-0410-2

García-Limones, C., SchnäBele, K., Blanco-Portales, R., Luz Bellido, M., Caballero, J. L., Schwab, W., et al. (2008). Functional characterization of FaCCD1: a carotenoid cleavage dioxygenase from strawberry involved in lutein degradation during fruit ripening. *J. Agric. Food Chem.* 56, 9277–9285. doi: 10.1021/jf801096t

Gayen, D., Ali, N., Sarkar, S. N., Datta, S. K., and Datta, K. (2015). Down-regulation of lipoxygenase gene reduces degradation of carotenoids of golden rice during storage. *Planta* 242, 353–363. doi: 10.1007/s00425-015-2314-4

Giuliano, G. (2014). Plant carotenoids: genomics meets multi-gene engineering. *Curr. Opin. Plant Biol.* 19, 111–117. doi: 10.1016/j.pbi.2014.05.006

Giuliano, G. (2017). Provitamin A biofortification of crop plants: a gold rush with many miners. *Curr. Opin. Plant Biol.* 44, 169–180. doi: 10.1016/j.copbio.2017.02.001

Giuliano, G., Al-Babili, S., and Von Lintig, J. (2003a). Carotenoid oxygenases: cleave it or leave it. *Trends Plant Sci.* 8, 145–149. doi: 10.1016/S1360-1385(03)00053-0

Giuliano, G., Rosati, C., and Bramley, P. M. (2003b). To dye or not to dye: biochemistry of annatto unveiled. *Trends Biotechnol.* 21, 513–516. doi: 10.1016/j.tibtech.2003.10.001

Goff, S. A., and Klee, H. J. (2006). Plant volatile compounds: sensory cues for health and nutritional value? *Science* 311, 815–819. doi: 10.1126/science.1112614

González-Guzmán, M., Apostolova, N., Bellés, J. M., Barrero, J. M., Piqueras, P., Ponce, M. R., et al. (2002). The short-chain alcohol dehydrogenase ABA2 catalyzes the conversion of xanthoxin to abscisic aldehyde. *Plant Cell* 14, 1833–1846. doi: 10.1105/tpc.002477

Gonzalez-Jorge, S., Ha, S. H., Magallanes-Lundback, M., Gilliland, L. U., Zhou, A., Lipka, A. E., et al. (2013). Carotenoid cleavage dioxygenase 4 is a negative regulator of β-carotene content in *Arabidopsis* seeds. *Plant Cell* 25, 4812–4826. doi: 10.1105/tpc.113.119677

González-Mas, M. C., Rambla, J. L., Alamar, M. C., Gutiérrez, A., and Granell, A. (2011). Comparative analysis of the volatile fraction of fruit juice from different citrus species. *PLoS One* 6:e22016. doi: 10.1371/journal.pone.0022016

Gutheil, W. G., Reed, G., Ray, A., Anant, S., and Dhar, A. (2012). Crocetin: an agent derived from saffron for prevention and therapy for cancer. *Curr. Pharm. Biotechnol.* 13, 173–179. doi: 10.2174/138920112798868566

Han, F., Cui, H., Zhang, B., Liu, X., Yang, L., Zhuang, M., et al. (2019a). Map-based cloning and characterization of BoCCD4, a gene responsible for white/yellow petal color in *B. oleracea*. *BMC Genomics* 20:242. doi: 10.1186/s12864-019-5596-2

Han, Y., Wang, H., Wang, X., Li, K., Dong, M., Li, Y., et al. (2019b). Mechanism of floral scent production in *Osmanthus fragrans* and the production and regulation of its key floral constituents, β-ionone and linalool. *Hortic. Res.* 6:106. doi: 10.1038/s41438-019-0189-4

Han, Y., Wu, M., Cao, L., Yuan, W., Dong, M., Wang, X., et al. (2016). Characterization of OfWRKY3, a transcription factor that positively regulates the carotenoid cleavage dioxygenase gene OfCCD4 in *Osmanthus fragrans*. *Plant Mol. Biol.* 91, 485–496. doi: 10.1007/s11103-016-0483-6

Harrison, E. H. (2005). Mechanisms of digestion and absorption of dietary vitamin A. *Annu. Rev. Nutr.* 25, 87–103. doi: 10.1146/annurev.nutr.25.050304.092614

Hasegawa, T., Bando, A., Tsuchiya, K., Abe, S., Okamoto, M., Kirima, K., et al. (2004). Enzymatic and nonenzymatic formation of reactive oxygen species from 6-anilino-5, 8-quinolinequinone. *Biophys Acta Gen Subj.* 1670, 19–27. doi: 10.1016/j.bbagen.2003.10.008

Hashimoto, H., Uragami, C., and Cogdell, R. J. (2016). Carotenoids and photosynthesis. *Carotenoid. Nat.* 79, 111–139. doi: 10.1007/978-3-319-39126-7_4

Havaux, M. (2020). β-Cyclocitral and derivatives: emerging molecular signals serving multiple biological functions. *Plant Physiol. Biochem.* 155, 35–41. doi: 10.1016/j.plaphy.2020.07.032

Hirschberg, J. (2001). Carotenoid biosynthesis in flowering plants. *Curr. Opin. Plant Biol.* 4, 210–218. doi: 10.1016/S1369-5266(00)00163-1

Hou, X., Rivers, J., León, P., Mcquinn, R. P., and Pogson, B. J. (2016). Synthesis and function of apocarotenoid signals in plants. *Trends Plant Sci.* 21, 792–803. doi: 10.1016/j.tplants.2016.06.001

Huang, F.-C., Molnár, P., and Schwab, W. (2009). Cloning and functional characterization of carotenoid cleavage dioxygenase 4 genes. *J. Exp. Bot.* 60, 3011–3022. doi: 10.1093/jxb/erp137

Ibdah, M., Azulay, Y., Portnoy, V., Wasserman, B., Bar, E., Meir, A., et al. (2006). Functional characterization of CmCCD1, a carotenoid cleavage dioxygenase from melon. *Phytochemistry* 67, 1579–1589. doi: 10.1016/j.phytochem.2006.02.009

Ilg, A., Beyer, P., and Al-Babili, S. (2009). Characterization of the rice carotenoid cleavage dioxygenase 1 reveals a novel route for geranial biosynthesis. *FEBS J.* 276, 736–747. doi: 10.1111/j.1742-4658.2008.06820.x

Ilg, A., Bruno, M., Beyer, P., and Al-Babili, S. (2014). Tomato carotenoid cleavage dioxygenases 1A and 1B: relaxed double bond specificity leads to a plenitude of dialdehydes, mono-apocarotenoids and isoprenoid volatiles. *FEBS Open Bio.* 4, 584–593. doi: 10.1016/j.fob.2014.06.005

Ilg, A., Yu, Q., Schaub, P., Beyer, P., and Al-Babili, S. (2010). Overexpression of the rice carotenoid cleavage dioxygenase 1 gene in Golden Rice endosperm suggests apocarotenoids as substrates in planta. *Planta* 232, 691–699. doi: 10.1007/s00425-010-1205-y

Jamil, M., Kountche, B. A., and Al-Babili, S. (2021). Current progress in Striga management. *Plant Physiol.* 185, 1339–1352. doi: 10.1093/plphys/kiab040

Jarén-Galán, M., and Mínguez-Mosquera, M. I. (1999). Effect of pepper lipoxygenase activity and its linked reactions on pigments of the pepper fruit. *J. Agric. Food Chem.* 47, 4532–4536. doi: 10.1021/jf9900682

Jia, K.-P., Baz, L., and Al-Babili, S. (2018). From carotenoids to strigolactones. *J. Exp. Biol.* 69, 2189–2204. doi: 10.1093/jxb/erx476

Jia, K.-P., Dickinson, A. J., Mi, J., Cui, G., Xiao, T. T., Kharbatia, N. M., et al. (2019). Anchorene is a carotenoid-derived regulatory metabolite required for anchor root formation in *Arabidopsis*. *Sci. Adv.* 5:eaaw6787. doi: 10.1126/sciadv.aaw6787

Jia, K.-P., Mi, J., Ablazov, A., Ali, S., Yang, Y., Balakrishna, A., et al. (2021a). Iso-anchorene is an endogenous metabolite that inhibits primary root growth in *Arabidopsis*. *Plant J.* 107, 54–66. doi: 10.1111/tpj.15271

Jia, K.-P., Mi, J., Ali, S., Ohyanagi, H., Moreno, J.C., Ablazov, A., et al. (2021b). An alternative, zeaxanthin epoxidase-independent abscisic acid biosynthetic pathway in plants. *Mol. Plant* doi: 10.1016/j.molp.2021.09.008 [Epub ahead of print].

Jian, W., Cao, H., Yuan, S., Liu, Y., Lu, J., Lu, W., et al. (2019). SlMYB75, an MYB-type transcription factor, promotes anthocyanin accumulation and enhances volatile aroma production in tomato fruits. *Hortic. Res.* 6:22. doi: 10.1038/s41438-018-0098-y

Jing, G., Li, T., Qu, H., Yun, Z., Jia, Y., Zheng, X., et al. (2015). Carotenoids and volatile profiles of yellow-and red-fleshed papaya fruit in relation to the expression of carotenoid cleavage dioxygenase genes. *Postharvest Biol. Technol.* 109, 114–119. doi: 10.1016/j.postharvbio.2015.06.006

Kalladan, R., Lasky, J. R., Sharma, S., Kumar, M. N., Juenger, T. E., Des Marais, D. L., et al. (2019). Natural variation in 9-*cis*-epoxycartenoid dioxygenase 3 and ABA accumulation. *Plant Physiol.* 179, 1620–1631. doi: 10.1104/pp.18.01185

Kong, C.-H., Zhang, S.-Z., Li, Y.-H., Xia, Z.-C., Yang, X.-F., Meiners, S. J., et al. (2018). Plant neighbor detection and allelochemical response are driven by root-secreted signaling chemicals. *Nat. Commun.* 9:3867. doi: 10.1038/s41467-018-06429-1

Lanfranco, L., Fiorilli, V., Venice, F., and Bonfante, P. (2018). Strigolactones cross the kingdoms: plants, fungi, and bacteria in the arbuscular mycorrhizal symbiosis. *J. Exp. Biol.* 69, 2175–2188. doi: 10.1093/jxb/erx432

Lashbrooke, J. G., Young, P. R., Dockrall, S. J., Vasanth, K., and Vivier, M. A. (2013). Functional characterisation of three members of the *Vitis vinifera* L. carotenoid cleavage dioxygenase gene family. *BMC Plant Biol.* 13:156. doi: 10.1186/1471-2229-13-156

Li, T., Deng, Y.J., Liu, J.X., Duan, A.Q., Liu, H., and Xiong, A.S. (2021). DcCCD4 catalyzes the degradation of α-carotene and β-carotene to affect carotenoid accumulation and taproot color in carrot. *Plant J.* doi: 10.1111/tpj.15524 [Epub ahead of print].

Li, L.-L., Zhao, H.-H., and Kong, C.-H. (2020). (−)-Loliolide, the most ubiquitous lactone, is involved in barnyardgrass-induced rice allelopathy. *J. Exp. Bot.* 71, 1540–1550. doi: 10.1093/jxb/erz497

Liang, M.-H., He, Y.-J., Liu, D.-M., and Jiang, J.-G. (2021). Regulation of carotenoid degradation and production of apocarotenoids in natural and engineered organisms. *Crit. Rev. Biotechnol.* 41, 531–534. doi: 10.1080/07388551.2021.1873242

Liu, F., Zhang, H., Ding, L., Soppe, W. J., and Xiang, Y. (2020). Reversal of rdo51, a homolog of rice seed dormancy4, interacts with bHLH57 and controls ABA biosynthesis and seed dormancy in *Arabidopsis*. *Plant Cell* 32, 1933–1948. doi: 10.1105/tpc.20.00026

Lu, S., Zhang, Y., Zhu, K., Yang, W., Ye, J., Chai, L., et al. (2018). The citrus transcription factor CsMADS6 modulates carotenoid metabolism by directly regulating carotenogenic genes. *Plant Physiol.* 176, 2657–2676. doi: 10.1104/pp.17.01830

Luan, Y., Fu, X., Lu, P., Grierson, D., and Xu, C. (2020a). Molecular mechanisms determining the differential accumulation of carotenoids in plant species and varieties. *Crit. Rev. Plant Sci.* 39, 125–139. doi: 10.1080/07352689.2020.1768350

Luan, Y., Wang, S., Wang, R., and Xu, C. (2020b). Accumulation of red apocarotenoid β-citraurin in peel of a spontaneous mutant of huyou (*citrus changshanensis*) and the effects of storage temperature and ethylene application. *Food Chem.* 309:125705. doi: 10.1016/j.foodchem.2019.125705

Ma, G., Zhang, L., Matsuta, A., Matsutani, K., Yamawaki, K., Yahata, M., et al. (2013). Enzymatic formation of β-citraurin from β-cryptoxanthin and

zeaxanthin by carotenoid cleavage dioxygenase 4 in the flavedo of citrus fruit. *Plant Physiol.* 163, 682–695. doi: 10.1104/pp.113.223297

Martí, M., Diretto, G., Aragonés, V., Frusciante, S., Ahrazem, O., Gómez-Gómez, L., et al. (2020). Efficient production of saffron crocins and picrocrocin in *Nicotiana benthamiana* using a virus-driven system. *Metab. Eng.* 61, 238–250. doi: 10.1016/j.ymben.2020.06.009

Mathieu, S., Terrier, N., Procureur, J. M., Bigey, F., and Günata, Z. (2005). A carotenoid cleavage dioxygenase from *Vitis vinifera* L.: functional characterization and expression during grape berry development in relation to C13-norisoprenoid accumulation. *J. Exp. Bot.* 56, 2721–2731. doi: 10.1093/jxb/eri265

Meng, N., Wei, Y., Gao, Y., Yu, K., Cheng, J., Li, X.-Y., et al. (2020). Characterization of transcriptional expression and regulation of *carotenoid cleavage dioxygenase 4b* in grapes. *Front. Plant Sci.* 11:483. doi: 10.3389/fpls.2020.00483

Mi, J., and Al-Babili, S. (2019). To color or to decolor: that is the question. *Mol. Plant* 12, 1173–1175. doi: 10.1016/j.molp.2019.07.007

Mi, J., Jia, K.-P., Balakrishna, A., Feng, Q., and Al-Babili, S. (2019). A highly sensitive SPE derivatization–UHPLC–MS approach for quantitative profiling of carotenoid-derived dialdehydes from vegetables. *J. Agric. Food Chem.* 67, 5899–5907. doi: 10.1021/acs.jafc.9b01749

Mitra, S., Estrada-Tejedor, R., Volke, D. C., Phillips, M. A., Gershenzon, J., and Wright, L. P. (2021). Negative regulation of plastidial isoprenoid pathway by herbivore-induced β-cyclocitral in *Arabidopsis thaliana*. *Proc. Natl. Acad. Sci. U. S. A.* 118:e2008747118. doi: 10.1073/pnas.2008747118

Moise, A. R., Al-Babili, S., and Wurtzel, E. T. (2014). Mechanistic aspects of carotenoid biosynthesis. *Chem. Rev.* 114, 164–193. doi: 10.1021/cr400106y

Moreno, J. C., Mi, J., Alagoz, Y., and Al-Babili, S. (2021). Plant apocarotenoids: from retrograde signaling to interspecific communication. *Plant J.* 105, 351–375. doi: 10.1111/tpj.15102

Murata, M., Kobayashi, T., and Seo, S. (2020). α-Ionone, an apocarotenoid, induces plant resistance to western flower thrips, *Frankliniella occidentalis*, independently of jasmonic acid. *Molecules* 25:17. doi: 10.3390/molecules25010017

Murata, M., Nakai, Y., Kawazu, K., Ishizaka, M., Kajiwara, H., Abe, H., et al. (2019). Loliolide, a carotenoid metabolite, is a potential endogenous inducer of herbivore resistance. *Plant Physiol.* 179, 1822–1833. doi: 10.1104/pp.18.00837

Nisar, N., Li, L., Lu, S., Khin, N. C., and Pogson, B. J. (2015). Carotenoid metabolism in plants. *Mol. Plant* 8, 68–82. doi: 10.1016/j.molp.2014.12.007

Ohmiya, A., Kishimoto, S., Aida, R., Yoshioka, S., and Sumitomo, K. (2006). Carotenoid cleavage dioxygenase (CmCCD4a) contributes to white color formation in chrysanthemum petals. *Plant Physiol.* 142, 1193–1201. doi: 10.1104/pp.106.087130

Pacheco, S. D. G., Gasparin, A. T., Jesus, C. H. A., Sotomaior, B. B., Ventura, A. C. S. S. B., Redivo, D. D. B., et al. (2019). Antinociceptive and anti-inflammatory effects of bixin, a carotenoid extracted from the seeds of *Bixa orellana*. *Planta Med.* 85, 1216–1224. doi: 10.1055/a-1008-1238

Park, S.-C., Kang, L., Park, W. S., Ahn, M.-J., Kwak, S.-S., and Kim, H. S. (2020). Carotenoid cleavage dioxygenase 4 (CCD4) cleaves β-carotene and interacts with IbOr in sweetpotato. *Plant Biotechnol. Rep.* 14, 737–742. doi: 10.1007/s11816-020-00649-y

Park, S., Kim, H. S., Jung, Y. J., Kim, S. H., Ji, C. Y., Wang, Z., et al. (2016). Orange protein has a role in phytoene synthase stabilization in sweetpotato. *Sci. Rep.* 6, 1–12. doi: 10.1038/srep33563

Ramel, F., Birtic, S., Ginies, C., Soubigou-Taconnat, L., Triantaphylides, C., and Havaux, M. (2012). Carotenoid oxidation products are stress signals that mediate gene responses to singlet oxygen in plants. *Proc. Natl. Acad. Sci. U. S. A.* 109, 5535–5540. doi: 10.1073/pnas.1115982109

Reidel, R. V. B., Melai, B., Cioni, P., Flamini, G., and Pistelli, L. (2016). Aroma profile of Rubus ulmifolius flowers and fruits during different ontogenetic phases. *Chem. Biodivers.* 13, 1776–1784. doi: 10.1002/cbdv.201600170

Rezaee, R., and Hosseinzadeh, H. (2013). Safranal: from an aromatic natural product to a rewarding pharmacological agent. *Iran. J. Basic Med. Sci.* 16, 12–26.

Rodrigo, M. J., Alquezar, B., Alos, E., Medina, V., Carmona, L., Bruno, M., et al. (2013). A novel carotenoid cleavage activity involved in the biosynthesis of *citrus* fruit-specific apocarotenoid pigments. *J. Exp. Bot.* 64, 4461–4478. doi: 10.1093/jxb/ert260

Rodriguez-Concepcion, M., Avalos, J., Bonet, M. L., Boronat, A., Gomez-Gomez, L., Hornero-Mendez, D., et al. (2018). A global perspective on carotenoids: metabolism, biotechnology, and benefits for nutrition and health. *Prog. Lipid Res.* 70, 62–93. doi: 10.1016/j.plipres.2018.04.004

Schwartz, S. H., Qin, X., and Zeevaart, J. A. (2001). Characterization of a novel carotenoid cleavage dioxygenase from plants. *J. Biol. Chem.* 276, 25208–25211. doi: 10.1074/jbc.M102146200

Schwartz, S. H., Tan, B. C., Gage, D. A., Zeevaart, J. A., and Mccarty, D. R. (1997). Specific oxidative cleavage of carotenoids by VP14 of maize. *Science* 276, 1872–1874. doi: 10.1126/science.276.5320.1872

Seo, M., Peeters, A. J., Koiwai, H., Oritani, T., Marion-Poll, A., Zeevaart, J. A., et al. (2000). The *Arabidopsis* aldehyde oxidase 3 (AAO3) gene product catalyzes the final step in abscisic acid biosynthesis in leaves. *Proc. Natl. Acad. Sci. U. S. A.* 97, 12908–12913. doi: 10.1073/pnas.220426197

Shi, J., Cao, C., Xu, J., and Zhou, C. (2020). Research advances on biosynthesis, regulation, and biological activities of apocarotenoid aroma in horticultural plants. *J. Chemother.* 2020, 1–11. doi: 10.1155/2020/2526956

Simkin, A. J., Schwartz, S. H., Auldridge, M., Taylor, M. G., and Klee, H. J. (2004a). The tomato carotenoid cleavage dioxygenase 1 genes contribute to the formation of the flavor volatiles β-ionone, pseudoionone, and geranylacetone. *Plant J.* 40, 882–892. doi: 10.1111/j.1365-313X.2004.02263.x

Simkin, A. J., Underwood, B. A., Auldridge, M., Loucas, H. M., Shibuya, K., Schmelz, E., et al. (2004b). Circadian regulation of the PhCCD1 carotenoid cleavage dioxygenase controls emission of β-ionone, a fragrance volatile of petunia flowers. *Plant Physiol.* 136, 3504–3514. doi: 10.1104/pp.104.049718

Stratton, S. P., Schaefer, W. H., and Liebler, D. C. (1993). Isolation and identification of singlet oxygen oxidation products of β-carotene. *Chem. Res. Toxicol.* 6, 542–547. doi: 10.1021/tx00034a024

Sui, X., Kiser, P. D., Von Lintig, J., and Palczewski, K. (2013). Structural basis of carotenoid cleavage: from bacteria to mammals. *Arch. Biochem. Biophys.* 539, 203–213. doi: 10.1016/j.abb.2013.06.012

Tan, B.-C., Joseph, L. M., Deng, W.-T., Liu, L., Li, Q.-B., Cline, K., et al. (2003). Molecular characterization of the *Arabidopsis* 9-*cis* epoxycarotenoid dioxygenase gene family. *Plant J.* 35, 44–56. doi: 10.1046/j.1365-313X.2003.01786.x

Tieman, D., Zhu, G., Resende, M. F., Lin, T., Nguyen, C., Bies, D., et al. (2017). A chemical genetic roadmap to improved tomato flavor. *Science* 355, 391–394. doi: 10.1126/science.aal1556

Vogel, J. T., Tan, B.-C., Mccarty, D. R., and Klee, H. J. (2008). The carotenoid cleavage dioxygenase 1 enzyme has broad substrate specificity, cleaving multiple carotenoids at two different bond positions. *J. Biol. Chem.* 283, 11364–11373. doi: 10.1074/jbc.M710106200

Von Lintig, J. (2010). Colors with functions: elucidating the biochemical and molecular basis of carotenoid metabolism. *Annu. Rev. Nutr.* 30, 35–56. doi: 10.1146/annurev-nutr-080508-141027

Walter, M. H., and Strack, D. (2011). Carotenoids and their cleavage products: biosynthesis and functions. *Nat. Prod. Rep.* 28, 663–692. doi: 10.1039/c0np00036a

Wang, J. Y., Alseekh, S., Xiao, T., Ablazov, A., De Souza, L. P., Fiorilli, V., et al. (2021). Multi-omics approaches explain the growth-promoting effect of the apocarotenoid growth regulator zaxinone in rice. *Commun. Biol.* 4:1222. doi: 10.1038/s42003-021-02740-8

Wang, J. Y., Haider, I., Jamil, M., Fiorilli, V., Saito, Y., Mi, J., et al. (2019a). The apocarotenoid metabolite zaxinone regulates growth and strigolactone biosynthesis in rice. *Nat. Commun.* 10:810. doi: 10.1038/s41467-019-08461-1

Wang, J. Y., Jamil, M., Lin, P.-Y., Ota, T., Fiorilli, V., Novero, M., et al. (2020a). Efficient mimics for elucidating zaxinone biology and promoting agricultural applications. *Mol. Plant* 13, 1654–1661. doi: 10.1016/j.molp.2020.08.009

Wang, J. Y., Lin, P.-Y., and Al-Babili, S. (2020b). On the biosynthesis and evolution of apocarotenoid plant growth regulators. *Semin. Cell Dev. Biol.* 109, 3–11. doi: 10.1016/j.semcdb.2020.07.007

Wang, X., Zeng, W., Ding, Y., Wang, Y., Niu, L., Yao, J.-L., et al. (2019b). PpERF3 positively regulates ABA biosynthesis by activating PpNCED2/3 transcription during fruit ripening in peach. *Hortic. Res.* 6:19. doi: 10.1038/s41438-018-0094-2

Watanabe, K., Oda-Yamamizo, C., Sage-Ono, K., Ohmiya, A., and Ono, M. (2018). Alteration of flower colour in Ipomoea nil through CRISPR/Cas9-mediated mutagenesis of carotenoid cleavage dioxygenase 4. *Transgenic Res.* 27, 25–38. doi: 10.1007/s11248-017-0051-0

Watkins, J. L., and Pogson, B. J. (2020). Prospects for carotenoid biofortification targeting retention and catabolism. *Trends Plant Sci.* 25, 501–512. doi: 10.1016/j.tplants.2019.12.021

Xi, W., Zheng, H., Zhang, Q., and Li, W. (2016). Profiling taste and aroma compound metabolism during apricot fruit development and ripening. *Int. J. Mol. Sci.* 17:998. doi: 10.3390/ijms17070998

Xie, X., Yoneyama, K., and Yoneyama, K. (2010). The strigolactone story. *Annu. Rev. Phytopathol.* 48, 93–117. doi: 10.1146/annurev-phyto-073009-114453

Xu, P., Chen, H., and Cai, W. (2020a). Transcription factor CDF4 promotes leaf senescence and floral organ abscission by regulating abscisic acid and reactive oxygen species pathways in *Arabidopsis*. *EMBO Rep.* 21:e48967. doi: 10.15252/embr.201948967

Xu, Z., Pu, X., Gao, R., Demurtas, O. C., Fleck, S. J., Richter, M., et al. (2020b). Tandem gene duplications drive divergent evolution of caffeine and crocin biosynthetic pathways in plants. *BMC Biol.* 18:63. doi: 10.1186/s12915-020-00795-3

Xu, J., Wang, X., Cao, H., Xu, H., Xu, Q., and Deng, X. (2017). Dynamic changes in methylome and transcriptome patterns in response to methyltransferase inhibitor 5-azacytidine treatment in citrus. *DNA Res.* 24, 509–522. doi: 10.1093/dnares/dsx021

Yahyaa, M., Bar, E., Dubey, N. K., Meir, A., Davidovich-Rikanati, R., Hirschberg, J., et al. (2013). Formation of norisoprenoid flavor compounds in carrot (*Daucus carota* L.) roots: characterization of a cyclic-specific carotenoid cleavage dioxygenase 1 gene. *J. Agric. Food Chem.* 61, 12244–12252. doi: 10.1021/jf404085k

Yamauchi, T., Watanabe, K., Fukazawa, A., Mori, H., Abe, F., Kawaguchi, K., et al. (2014). Ethylene and reactive oxygen species are involved in root aerenchyma formation and adaptation of wheat seedlings to oxygen-deficient conditions. *J. Exp. Bot.* 65, 261–273. doi: 10.1093/jxb/ert371

Yao, M., Chen, W., Kong, J., Zhang, X., Shi, N., Zhong, S., et al. (2020). Methyltransferase1 and ripening modulate vivipary during tomato fruit development. *Plant Physiol.* 183, 1883–1897. doi: 10.1104/pp.20.00499

Yoneyama, K., Akiyama, K., Brewer, P. B., Mori, N., Kawano-Kawada, M., Haruta, S., et al. (2020). Hydroxyl carlactone derivatives are predominant strigolactones in *Arabidopsis*. *Plant Direct* 4:e00219. doi: 10.1002/pld3.219

Yu, Q., Plotto, A., Baldwin, E. A., Bai, J., Huang, M., Yu, Y., et al. (2015). Proteomic and metabolomic analyses provide insight into production of volatile and non-volatile flavor components in mandarin hybrid fruit. *BMC Plant Biol.* 15:76. doi: 10.1186/s12870-015-0466-9

Yuan, H., Zhang, J., Nageswaran, D., and Li, L. (2015). Carotenoid metabolism and regulation in horticultural crops. *Hortic. Res.* 2:15036. doi: 10.1038/hortres.2015.36

Yue, Z., Liu, H., and Ma, F. (2015). The malus carotenoid cleavage dioxygenase 7 is involved in stress response and regulated by basic pentacysteine 1. *Sci. Hortic.* 192, 264–270. doi: 10.1016/j.scienta.2015.06.027

Zhang, B., Liu, C., Wang, Y., Yao, X., Wang, F., Wu, J., et al. (2015). Disruption of a CAROTENOID CLEAVAGE DIOXYGENASE 4 gene converts flower colour from white to yellow in brassica species. *New Phytol.* 206, 1513–1526. doi: 10.1111/nph.13335

Zhang, Y., Van Dijk, A. D., Scaffidi, A., Flematti, G. R., Hofmann, M., Charnikhova, T., et al. (2014). Rice cytochrome P450 MAX1 homologs catalyze distinct steps in strigolactone biosynthesis. *Nat. Chem. Biol.* 10:1028. doi: 10.1038/nchembio.1660

Zhang, Y., Zhang, Y., Sun, Q., Lu, S., Chai, L., Ye, J., et al. (2021). Citrus transcription factor CsHB5 regulates abscisic acid biosynthetic genes and promotes senescence. *Plant J.* 108, 151–168. doi: 10.1111/tpj.15431

Zheng, X., Giuliano, G., and Al-Babili, S. (2020a). Carotenoid biofortification in crop plants: *citius, altius, fortius*. *Biochim. Biophys. Acta Mol. Cell Biol. Lipids* 1865:158664. doi: 10.1016/j.bbalip.2020.158664

Zheng, X., Kuijer, H. N., and Al-Babili, S. (2020b). Carotenoid biofortification of crops in the Crispr era. *Trends Biotechnol.* 39, 857–860. doi: 10.1016/j.tibtech.2020.12.003

Zheng, X., Mi, J., Deng, X., and Al-Babili, S. (2021). LC–MS-based profiling provides new insights into apocarotenoid biosynthesis and modifications in citrus fruits. *J. Agric. Food Chem.* 69, 1842–1851. doi: 10.1021/acs.jafc.0c06893

Zheng, X., Xie, Z., Zhu, K., Xu, Q., Deng, X., and Pan, Z. (2015). Isolation and characterization of carotenoid cleavage dioxygenase 4 genes from different citrus species. *Mol. Gen. Genomics.* 290, 1589–1603. doi: 10.1007/s00438-015-1016-8

Zheng, X., Zhu, K., Sun, Q., Zhang, W., Wang, X., Cao, H., et al. (2019). Natural variation in CCD4 promoter underpins species-specific evolution of red coloration in citrus peel. *Mol. Plant* 12, 1294–1307. doi: 10.1016/j.molp.2019.04.014

Zhong, Y., Pan, X., Wang, R., Xu, J., Guo, J., Yang, T., et al. (2020). ZmCCD10a encodes a distinct type of carotenoid cleavage dioxygenase and enhances plant tolerance to low phosphate. *Plant Physiol.* 184, 374–392. doi: 10.1104/pp.20.00378

Zhou, X., Welsch, R., Yang, Y., Álvarez, D., Riediger, M., Yuan, H., et al. (2015). *Arabidopsis* OR proteins are the major posttranscriptional regulators of phytoene synthase in controlling carotenoid biosynthesis. *Proc. Natl. Acad. Sci. U. S. A.* 112, 3558–3563. doi: 10.1073/pnas.1420831112

Zhu, F., Luo, T., Liu, C., Wang, Y., Yang, H., Yang, W., et al. (2017). An R2R3-MYB transcription factor represses the transformation of α-and β-branch carotenoids by negatively regulating expression of CrBCH2 and CrNCED5 in flavedo of *citrus* reticulate. *New Phytol.* 216, 178–192. doi: 10.1111/nph.14684

Zhu, F., Luo, T., Liu, C., Wang, Y., Zheng, L., Xiao, X., et al. (2020). A NAC transcription factor and its interaction protein hinder abscisic acid biosynthesis by synergistically repressing NCED5 in Citrus reticulata. *J. Exp. Bot.* 71, 3613–3625. doi: 10.1093/jxb/eraa118

Zhu, K., Sun, Q., Chen, H., Mei, X., Lu, S., Ye, J., et al. (2021). Ethylene activation of carotenoid biosynthesis by a novel transcription factor CsERF061. *J. Exp. Bot.* 72, 3137–3154. doi: 10.1093/jxb/erab047

Zoccali, M., Giuffrida, D., Salafia, F., Rigano, F., Dugo, P., Casale, M., et al. (2021). Apocarotenoids profiling in different capsicum species. *Food Chem.* 334:127595. doi: 10.1016/j.foodchem.2020.127595

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.

The handling editor declared a past co-authorship with one of the authors SA-B.

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.

frontiers in Plant Science

ORIGINAL RESEARCH
published: 14 December 2021
doi: 10.3389/fpls.2021.767724

A Comparative Metabolomic Analysis Reveals Difference Manufacture Suitability in "Yinghong 9" and "Huangyu" Teas (*Camellia sinensis*)

Xin Mei[1,2], Chuyuan Lin[1], Shihua Wan[1], Baoyi Chen[1], Hualing Wu[3] and Lingyun Zhang[1]**

[1] College of Horticulture, South China Agricultural University, Guangzhou, China, [2] College of Biological Science and Agriculture, Qiannan Normal University for Nationalities, Duyun, China, [3] Tea Research Institute, Guangdong Academy of Agricultural Sciences, Guangdong Provincial Key Laboratory of Tea Plant Resources Innovation and Utilization, Guangzhou, China

OPEN ACCESS

Edited by:
Renata Rivera-Madrid, Scientific Research Center of Yucatán (CICY), Mexico

Reviewed by:
Kang Wei, Tea Research Institute, Chinese Academy of Agricultural Sciences (CAAS), China
Chuankui Song, Anhui Agricultural University, China
Yuhua Wang, Nanjing Agricultural University, China

***Correspondence:**
Hualing Wu
wuhualing@163.com
Lingyun Zhang
zhanglingyun@scau.edu.cn

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

Received: *31 August 2021*
Accepted: *16 November 2021*
Published: *14 December 2021*

Citation:
Mei X, Lin C, Wan S, Chen B, Wu H and Zhang L (2021) A Comparative Metabolomic Analysis Reveals Difference Manufacture Suitability in "Yinghong 9" and "Huangyu" Teas (Camellia sinensis). Front. Plant Sci. 12:767724. doi: 10.3389/fpls.2021.767724

"Yinghong 9" is a widely cultivated large-leaf variety in South China, and the black tea made from it has a high aroma and strong sweet flavor. "Huangyu" is a light-sensitive tea variety with yellow leaves. It was cultivated from the bud-mutation of "Yinghong 9" and has a very low level of chlorophyll during young shoot development. Due to chlorophyll being involved in carbon fixation and assimilation, the changes in photosynthesis might potentially affect the accumulation of flavor metabolites, as well as the quality of "Huangyu" tea. Although "Huangyu" has a golden yellow color and high amino acid content, the mechanism underlying the formation of leaf color and drinking value remains unclear. The widely targeted metabolomics and GC-MS analysis were performed to reveal the differences of key metabolites in fresh and fermented leaves between "Yinghong 9" and "Huangyu." The results showed that tea polyphenols, total chlorophyll, and carotenoids were more abundant in "Yinghong 9." Targeted metabolomics analysis indicated that kaempferol-3-glycoside was more abundant in "Yinghong 9," while "Huangyu" had a higher ratio of kaempferol-3-glucoside to kaempferol-3-galactoside. Compared with "Yinghong 9" fresh leaves, the contents of zeaxanthin and zeaxanthin palmitate were significantly higher in "Huangyu." The contents of α-farnesene, β-cyclocitral, nerolidol, and *trans*-geranylacetone, which were from carotenoid degradation and involved in flowery-fruity-like flavor in "Huangyu" fermented leaves, were higher than those of "Yinghong 9." Our results indicated that "Huangyu" was suitable for manufacturing non-fermented tea because of its yellow leaf and flowery-fruity-like compounds from carotenoid degradation.

Keywords: ***Camellia sinensis*, yellow-leaf mutant, metabolites analysis, flavonol glycosides, carotenoids, aroma compounds**

INTRODUCTION

Tea (*Camellia sinensis*) is an important economic crop in south China. The sensory quality of tea products depends on their chemical compounds, although it is still challenging to illustrate. Color and flavor are two important sensory quality indicators for tea products. The leaf colors of tea cultivars often have severely effects on appearance and flavor which involved in manufacturing suitability of tea leaf (Zhang et al., 2017;

Cheng et al., 2019; Hao et al., 2020). Many developers believe that the high-chlorophyll tea leaf is suitable for making green tea, while the low-chlorophyll leaf is suitable for making black or oolong tea. The main reason is that the low-chlorophyll tea leaf more easily turns yellowish or reddish during black tea or oolong tea processing. Recently, some white and yellow nature mutation varieties were cultured because of its yellow-green appearance of finished tea. Admittedly, yellow or white leaf tea mutations have high contents of amino acids, accompanied by low contents of tea polyphenols and total chlorophylls, which contribute to the qualities of white and yellow leaf tea (Feng et al., 2014). The discovery and revelation of the mechanism of leaf color mutants is a hot topic in tea plant breeding.

Previous studies showed that abnormal chlorophyll metabolism in nature mutation tea led to white or yellow bud and leaf. Many researches have been conducted on leaf color formation mechanisms in albino-cultivar during young shoot development using transcriptomics, metabolomics, and proteomics analysis (Li et al., 2011; Wang et al., 2013; Liu et al., 2017; Lu et al., 2019). Some research results showed that the yellow degree of leaves had a close relationship with the content of amino acids (Feng et al., 2014; Song et al., 2017). In nature, some natural bud mutants have been found to cope with the complex environment by changing their leaf color; this potential capability has been found in almost all higher plants (Huot et al., 2014; Li et al., 2020).

"Yinghong 9" (*Camellia sinensis var.* Yinghong No. 9) is a large leaf tea cultivar selected from Yunnandaye tea, it is a suitable cultivar for the manufacture of black tea because of its higher tea polyphenol content and flowery-aroma flavor. "Huangyu" is a yellow leaf mutant that bred from bud mutation of "Yinghong 9," and it has a relatively low level of chlorophyll than the parent plant (Ma et al., 2016; Cheng et al., 2019). The changes of photosynthesis and metabolite accumulation in "Huangyu" might potentially affect the quality of the finished tea (Ma et al., 2016). So, it is interesting for the researchers to study the chemical compounds, differences, and manufacture suitability of "Yinghong 9" and "Huangyu" with the same genetic background.

Carotenoids and chlorophyll are the dominant pigments involved in the formation of leaf color of plants (Clément et al., 2015). The content of total chlorophylls and carotenoids, and especially their ratio (carotenoid/chlorophyll) in leaf, affect the environmental adaptability of plants during their development (Solovchenko, 2010; Gitelson, 2020). Long's (Long et al., 2006) research showed that although carotenoid biosynthesis pathway would be changed in transgenic lines of tomato, which did not generally alter the content of phenolic or flavonoids. On the contrary, low chlorophyll and high carotenoids lead to a lower flavonoid in albino and yellow leaf tea, e.g., Anji baicha and Huangjinya. Some literature have shown that carotenoids were involved in the formation of aroma quality during tea processing. The products of carotenoid degradation, such as β-ionone, geranylacetone, β-cyclocitral, nerolidol and β-damascenone, are important aroma compounds of tea (Sanderson and Grahamm, 1973; Kawakami, 2002; Hazarika and Mahanta, 2010; Ho et al., 2015).

Currently, many aspects of coloration of the leaf and its influence on tea products and aroma formation are still poorly understood. "Huangyu," a natural mutant variety of golden-yellow leaf tea, has a higher carotenoid content compared to its parent, with the same genetic background, offers us a chance on revealing the relationship of aroma quality with the yellow leaf color in tea. The purpose of our study is to elucidate whether the golden-yellow-leaf tea variety "Huangyu" has a better quality on the basis of its higher carotenoids and carotenoid-derived aroma compounds. This study also provides new insights into the mechanisms of the difference of quality and the manufacture suitability of "Huangyu" and "Yinghong 9." The results of our study can provide theoretical evidence for tea breeding and utilization of unique wild tea resources.

MATERIALS AND METHODS

Plant Materials and Sample Preparation

One bud and two leaves of *Camellia sinensis* cv. "Yinghong 9" were collected from the experimental tea garden of Tea Research Institute, Guangdong Academy of Agricultural Sciences (Yingde City, Guangdong Province, China) in August 2020. The mutant materials ("Huangyu") were collected from the mutant plant, which was grafted on the parent plant ("Yinghong 9"). The tea leaves were equally divided into three portions. The samples were prepared by black tea processing described as follows: the fresh tea leaves were subjected to indoor natural withering at 25–27°C by spreading on bamboo sieves for 15 h. After withering, the leaves were rolled by a mini rolling machine for 60 min. Then the leaves were put into a bamboo-basket to ferment for 6 h at 25–27°C with a relative humidity of about 85%. During the black tea processing, fresh and fermented leaves were collected and immediately fixed with liquid nitrogen. And then the samples were ground to powder and stored in a freezer at −80°C before the following analysis. The water content of these leaves was measured at different stages during the black tea processing.

Determination of Tea Polyphenol, Chlorophyll, and Carotenoids

Content Determination of Tea Polyphenols

The tea polyphenols content was determined by spectrophotometer method with minor modifications (Liang et al., 2003). The absorbance values at 540 nm were determined by a spectrophotometer (UNICO Instrument Co., Ltd., Shanghai, China). A control solution (5 ml distilled water, 5 ml reagent and 15 ml phosphate buffer) was determined on the same step. The calculation formula is as follows: tea polyphenols concentration (%) = (A × 1.957 × 2)/100 × L1/(L2 × M × m). A represents absorbance values. L1 represents the total volume of extracting solution, in ml. L2 represents the tested volume of extracting solution. M represents the dry weight of the sample, in g. The "m" represents dry-weight percentage of the sample in %.

Determination of Total Chlorophyll and Carotenoids

The determination of total chlorophyll and carotenoids contents were extracted and determined as previously described with some

modifications (Song et al., 2017; Tian et al., 2019); The ground samples (500 mg) were extracted by 25 ml 95% ethanol (v/v) in darkness for 24 h at room temperature. After that, the extract was filtered and brought up to 25 ml using 95% ethanol (v/v), and then added 300 μl extracting into a 96 well plate, the absorbance values were determined using an automatic microplate reader (Sunrise, TECAN, Switzerland), and test wavelength is 663, 645, and 470 nm, respectively. The blank control used was a 95% ethanol (v/v). The content calculation equations of chlorophyll a, chlorophyll b, and carotenoids were as follows:

$$\text{Chlorophyll a} = (12.21 \times A_{663} - 2.81 \times A_{646})/(1000 \times W) \times L;$$

$$\text{Chlorophyll b} = (20.13 \times A_{646} - 5.03 \times A_{663})/(1000 \times W) \times L;$$

$$\text{Carotenoids} = (1,000 \times A_{470} - 3.27 \times C_{\text{chlorophylla}} - 104 \times C_{\text{chlorophyll b}})/(229 \times 1000 \times W) \times L.$$

where A663, A646, and A470 represent the absorbance at 663, 646, and 470 nm, respectively; where V represents the total volume of extracting solution (ml), and W represents the fresh weight (FW) of the samples (g); The contents of chlorophyll and carotenoids were estimated in mg/g (FW).

Metabolite Compounds Analysis in "Yinghong 9" and "Huangyu"

All above frozen samples were freeze-dried by LGJ-10 experimental vacuum freeze dryer (Songyuan Huaxing Technology Development Co., Ltd., Beijing, China). The freeze-dried samples were ground using a MM 400 grinder (Retsch, VERDER Group, Germany) with a zirconia bead for 90 s at 30 Hz. The lyophilized sample (100 mg) was extracted with 1.2 mL 70% methanol solution for 24 h at 4°C. During this period, the samples were vortexed (vortex-genie2, Scientific Industries, United States) 30 s once every 30 minutes for a total of six times to facilitate the extraction. Following centrifugation at 12,000 *g* for 15 min, the supernatant was filtered with SCAA-104 (0.22 μm microporous filter membrane, ANPEL, Shanghai, China) and analyzed by UPLC-MS/MS.

The samples were analyzed using Ultra Performance Liquid Chromatography Tandem Mass Spectrometry (UPLC-MS/MS) according to the procedure described by Zhou C. et al. (2020). The chromatographic separations were conducted on an Agilent UPLC column SB-C18 (1.8 μm, 2.1 mm × 100 mm) at 40°C. 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 flow rate of the mobile phase was 0.35 ml/min; The injection volume was 4 μl. The metabolites were identified by using the primary and secondary spectral data annotated on public databases (MassBank[1]) and proprietary database (Metware Database, Metware Biotechnology Co., Ltd., Wuhan, China) following the previously operating procedures (Rothenberg et al., 2019). Metabolite quantification was conducted using MRM. In order to monitor the stabilization of analyzing equipment, two mixture samples were run for quality control after every 10 samples processing. Unsupervised PCA (principal component analysis) was performed by statistics function prcomp within R[2]. Significantly differential metabolites between groups were determined when its variable importance in the project (VIP) ≥ 1 and fold change ≥ 2 or ≤ 0.5. VIP values were extracted from orthogonal partial least squares discriminant analysis (OPLS-DA) results, which were conducted using R package (Chong and Xia, 2018).

Isolation and Analysis of Carotenoid Compounds

The freeze-dried samples were ground (30 Hz, 1 min) using a MM 400 grinder (Retsch, VERDER Group, Germany). Fifty milligrams of ground samples were mixed with an appropriate volume of the extraction solution (ethanol: acetone: n-hexane, 2:1:1,v/v/v) with 0.01% butylated hydroxytoluene (g/mL) and internal standard. The mixture was adequately vortexed (Jingmei, Shanghai, China) at room temperature for 20 min. The mixture solution was then centrifuged (5424R, Eppendorf, Germany), and the supernatant was collected. The samples were repeatedly extracted two times using the same method, and the supernatants were combined and concentrated. The concentration of extracts redissolved in mixture solution (methanol: methyl tertiary-butyl ether, 3:1, v/v). The solutions were filtrated by SCAA-104 filter (0.22 μm, Anpel laboratory technologies Inc., Shanghai, China) before UPLC-MS/MS analysis. Carotenoid identification and quantification analysis were performed by UPLC-MS/MS system at Metware Biotechnology Company (Wuhan, China). UHPLC Chromatographic and APCI-Q Trap-MS/MS conditions were conducted following the previously stated procedure (Yan et al., 2019). Carotenoids were quantitatively analyzed as described in Zhou W. et al. (2020).

Volatile Compounds Determination by GC–MS

Volatile constituents were extracted according to the previous study (Mu et al., 2018; Shi, 2018) with minor changes. Briefly, ground tea powder (1.0 g) was accurately weighed into a homemade extraction bottle (50 mL), and 3 ml water and 1 50 μl decanoic acid ethyl ester (100 μM, internal standard) was added. The extraction bottle was placed in an air bath (45°C) to equilibrate for 5 min, volatile constituent was extracted using a headspace solid-phase microextraction (HS-SPME) device (57310U, 65 μm PDMS/DVB, Supelco, Bellefonte,

[1]http://www.massbank.jp/

[2]www.r-project.org

PA, United States) in a water bath (45°C) for 37 min. After extraction, the SPME fiber was immediately inserted into the GC injector for desorption at 250°C for 5.5 min. Volatile constituents were analyzed using GC–MS (Agilent GC 6890N-5973, Agilent, Santa Clara, CA, United States) as previously described (Mu et al., 2018). GC/MSD MassHunter workstation software (Agilent, Santa Clara, CA, United States) was used for GC–MS data analyses including peaks extracting, data baselines filtering, calibration of the baseline, peak alignment, deconvolution analysis, and integration of the peak area. Identification was performed by comparing mass spectra with NIST14.L libraries databases. To compare the differences of "Huangyu" and "Yinghong 9," the critical volatiles was selected by retention time as an approximate guide (Wu et al., 2019).

Determination of Enzyme Activities

The carotenoid cleavage crude enzyme extracts were carried out as a method described before (Ningrum and Schreiner, 2014). The polyphenol oxidase enzyme extracts were carried out as follows: 1,000 mg frozen samples were weighed into the mortar, add 5 mL precooling phosphate buffer (pH 7.0) with 300 mg polyvinylpolypyrrolidone (PVPP), and the mixture was sufficiently ground on the icebox. Then the mixture was washed into a 20 ml centrifugal tube using 5 ml buffer. The homogenate was centrifuged at 15,000 *g* for 20 min at 4°C, and then the supernatant was immediately used for enzyme assays. The enzyme activities of samples were determined according to the protocol of the assay kits, respectively (Polyphenol Oxidase assay kit, Suzhou Keming Biotechnology Co., Ltd. β-Glucosidase assay kit and Plant Carotenoid Cleavage Dioxygenase Elisa kit, Beijing Solarbio Science & Technology Co., Ltd.).

Data Analysis

Statistical analysis was carried out with SPSS software (IBM version 24.0 for Windows, SPSS Inc., Chicago, IL, United States). Significant differences between different groups were determined using Tukey's *post hoc* test. *P* value less than 0.05 was considered statistically significant. Excel 2010 (Microsoft, United States) was applied to drawbar graphs of the experimental data. Flavonol metabolites heat map was prepared using pheatmap in R package (**??**). The carotenoids heat map was prepared on major bio cloud platform (Major Biotechnology Co., Shanghai, China[3]).

RESULTS

Pigments Content in "Yinghong 9" and "Huangyu"

The leaves of "Yinghong 9" show normal green leaf and bud, while its mutant "Huangyu" has yellow leaf and bud (**Figure 1A**). The fermented leaves from "Huangyu" were bright red compared with "Yinghong 9" of which leaves show dark brown color (**Figure 1B**). The leaf color of "Yinghong 9" and "Huangyu" was affected by pigment composition and proportion, including the contents of tea polyphenols, chlorophyll a, chlorophyll b, carotenoids, and their ratio. The current results showed that the contents of tea polyphenols and total chlorophyll in "Yinghong 9" leaves were higher than that of "Huangyu," In contrast, the content of carotenoids in "Huangyu" leaves was still lower than those in normal green leaf parent, "Yinghong 9" (**Figures 1C,D**). The contents of all three pigments in "Yinghong 9" and "Huangyu" fermented leaves were significantly lower ($p < 0.05$) than those in fresh leaves (**Figures 1C–E**).

Differential Metabolites in "Yinghong 9" and "Huangyu"

A total of 609 metabolites were identified in "Yinghong 9" and "Huangyu" fresh and fermented leaves by UPLC-MS/MS analysis, and representative differential metabolites, including 26 up-regulated and 20 down-regulated metabolites were identified in "Huangyu" and "Yinghong 9" fresh leaves (fold change (FC) $\geq$ 2 or $\leq$ 0.5) (**Supplementary Table 1**). Interestingly, among the top 20 down-regulated metabolites, 9,12,13-Trihyroxy-10,15-octadecadienoic acid and δ-tridecalactone showed markedly different in "Yinghong 9" and "Huangyu" fresh leaves (**Figure 2A**). The reason might be that the two compounds were almost undetectable in "Huangyu." Among the 26 up-regulated metabolites, protocatechuic acid-4-glucoside showed the highest increased level in "Huangyu" fresh leaves, increasing about 3.37-fold (**Figure 2B**). It was noteworthy that most of the top 10 up-regulated metabolites were phenolic compounds (phenolic acids or tannins). In comparison, most of the top 10 down-regulated metabolites were lipids (high content in "Yinghong 9") (**Figure 2C**). The difference was possibly due to metabolic flux in the two cultivars because of their different leaf color.

Compared with "Yinghong 9" fresh leaves, there were 408 differential metabolites in fermented leaves (225 up-regulated and 183 down-regulated), while 439 differential metabolites in "Huangyu" fresh leaves were different from fermented leaves (255 up-regulated and 184 down-regulated) (**Supplementary Tables 2**, **3**). The results show that the highest increase in up-regulated metabolites was D-glucurono-6,3-lactone, while the most significant change in down-regulated metabolites was glutathione during "Yinghong 9" fermented processing. Most of them were organic acids or phenolic acids, which implied that organic acids or phenolic acids were converted to new metabolites during black tea fermentation. The top ten up-regulated and down-regulated metabolites during "Yinghong 9" fermented processing were very similar.

Comparison of Differences in Flavonol Glycosides in "Yinghong 9" and "Huangyu"

Altogether, 20 flavonol glycosides were identified in the "Yinghong 9" and "Huangyu" leaves as shown in **Supplementary Table 4**, including eight kaempferol-3-glycosides, ten quercetin-3-glycosides, and two myricetin-3-glycosides. We observed that the content of flavonol glycosides was slightly different in "Yinghong 9" and "Huangyu" fresh leaves, but the changes

[3]https://cloud.majorbio.com/

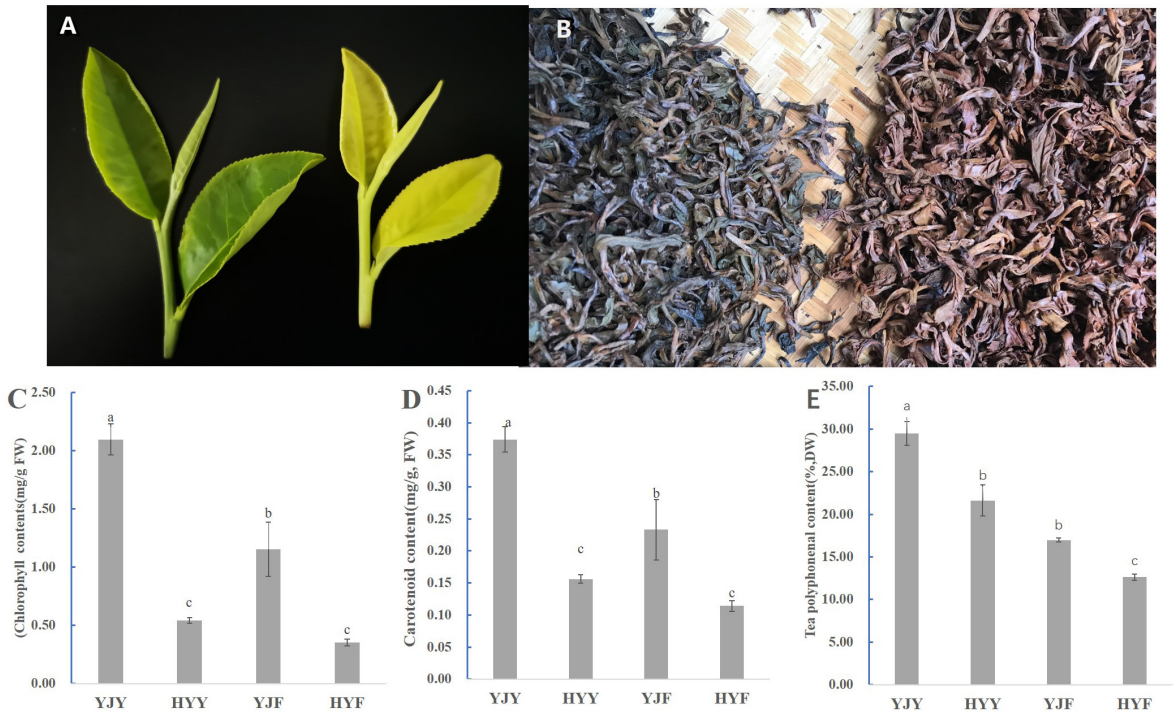


FIGURE 1 | Phenotypes and pigment content in "Yinghong 9" and "Huangyu" leaves. **(A)** Leaf phenotypes of "Yinghong 9" and "Huangyu." **(B)** Appearance of "Yinghong 9" and "Huangyu" fermented leaves. **(C)** Total chlorophyll content of "Yinghong 9" and "Huangyu" leaves. **(D)** Carotenoid content of "Yinghong 9" and "Huangyu" leaves. **(E)** Total tea polyphenols content of "Yinghong 9" and "Huangyu" leaves. The data are presented as the mean standard deviation ($n = 3$). Means with different letters at each treatment represent a significant difference at $p \leq 0.05$. YJY was "Yinghong 9" fresh leaves, YJF was "Yinghong 9" fermented leaves; HYY was "Huangyu" fresh leaves, HYF was "Huangyu" fermented leaves.

of flavonol glycosides were different during the two cultivars fermentation. Quercetin-3-*O*-xyloside was the most abundant detected compound in "Yinghong 9" and "Huangyu" fermented leaves (**Supplementary Table 4**). But the highest content of the top 15 changes was myricetin-3-*O*-arabinoside in "Yinghong 9" fermented leaves (**Figure 3A** and **Supplementary Table 5A**), while it was quercetin-3-*O*-α-L-arabinopyranoside in "Huangyu" fermented leaves (**Figure 3B** and **Supplementary Table 5B**). In addition, the content of kaempferol-3-glucoside, kaempferol-3-*O*-galactoside and kaempferol-3-glucoside to kaempferol-3-galactoside ratio were significantly different. The content of kaempferol-3-glucoside in "Yinghong 9" was 1.5 fold more than that of "Huangyu" (**Supplementary Table 4**). The significant difference in kaempferol-3-glucoside and its ratio to kaempferol-3-galactoside indicates the potential for the discrimination of the manufacture suitability of tea varieties.

Comparison of Carotenoids Metabolites in "Yinghong 6" and "Huangyu"

A total of 23 carotenoids were identified in "Yinghong 9" and "Huangyu" fresh leaves, including a high content of α-carotene, β-carotene, lutein; zeaxanthin, and minor compounds, such as antheraxanthin, violaxanthin, neoxanthin, apocarote, α-cryptoxanthin, β-cryptoxanthin, capsanthin, (E/Z)-phytoene, apocarotenal, zeaxanthin, lycopene, violaxanthin and neoxanthin (**Supplementary Table 6**). The types and concentrations of carotenoids were significantly different between "Yinghong 9" and "Huangyu." Compared with "Huangyu" fresh leaves, the contents of lutein, α-carotene, (E/Z)-phytoene, β-cryptoxanthin, β-carotene, and neoxanthin were significantly higher in "Yinghong 9." While the content of zeaxanthin was significantly higher in "Huangyu" leaves (**Figure 4** and **Supplementary Table 6**).

Most of the carotenoids significantly decreased in "Yinghong 9" fermented and "Hongyun" fermented leaves, e.g., β-carotene, lutein, zeaxanthin. On the contrary, some carotenoids significantly increased in fermented leaves, such as antheraxanthin, violaxanthin, and neoxanthin in "Yinghong 9" fermented leaves; lutein caprate and violaxanthin in "Huangyu" fermented leaves. Interestingly, lutein caprate even was not detected in "Yinghong 9" fresh leaves, which markedly increased in "Huangyu" fermented leaves. The reason may be that the lutein caprate is the product of β-carotene degradation in fermented leaves.

Volatile Compounds Profiling of "Yinghong 9" and "Huangyu" Leaves by SPME-GC–MS

In order to study the differences of volatile compounds, volatile compounds analyses were carried out in "Yinghong 9" and

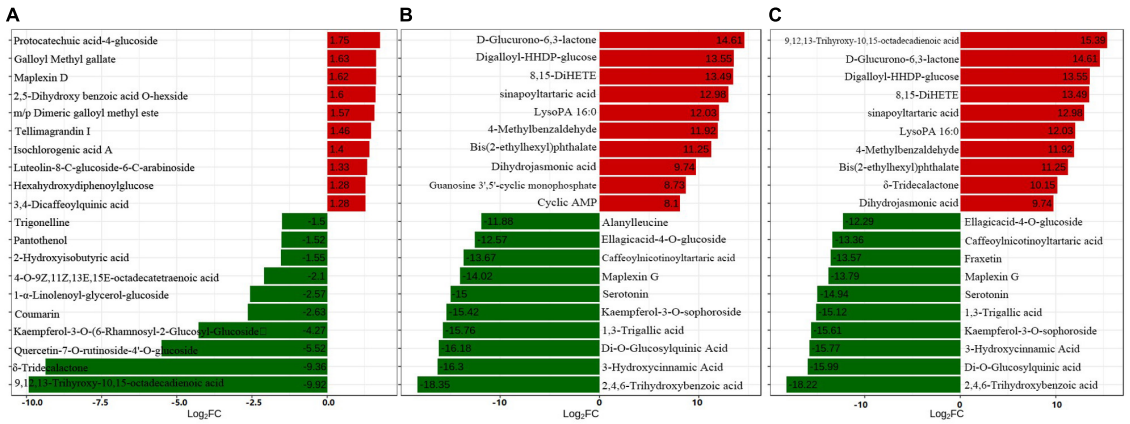


FIGURE 2 | The top 20 differential metabolites between "Yinghong 9" (YJ) and "Huangyu" (HY). **(A)** Top 20 FC differential metabolites between "Yinghong 9" (YJ) and "Huangyu" (HY), fold change (YJY/HYY). **(B)** Top 20 FC differential metabolites between fresh leaves (YJY) and fermented leaves (YJF)of "Yinghong 9," fold change (YJY/YJF). **(C)** Top 20 FC change metabolites between Fresh leaves (HYY) and Fermented leaves (HYF)of "Huangyu," fold change (HYY/HYF).

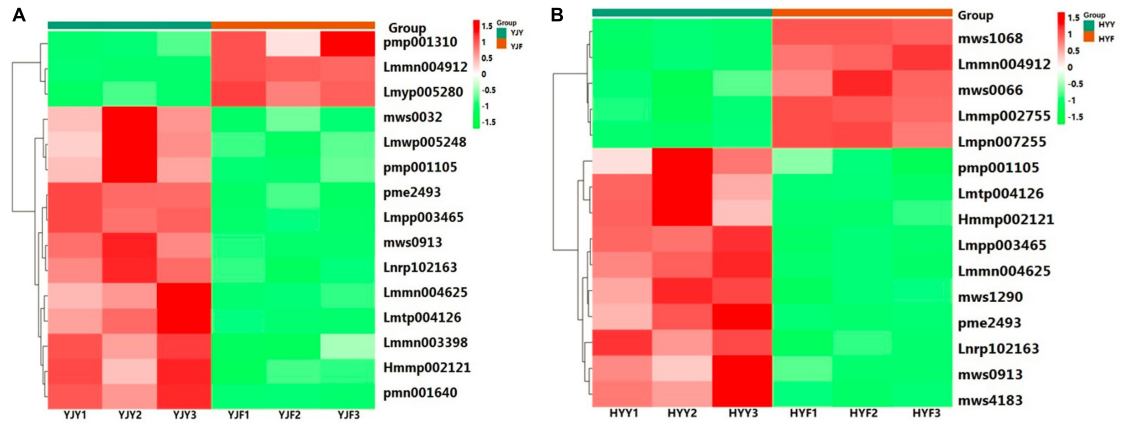


FIGURE 3 | Heat map of differential flavonol glycosides clusters in "Yinghong 9" and "Huangyu." **(A)** Top 15 FC change flavonol glycosides between fresh leaves (YJY) and fermented leaves (YJF)of "Yinghong 9." **(B)** Top 15 FC change flavonol glycosides between fresh leaves (YJY) and fermented leaves (YJF) of "Huangyu." In the heatmap, differential flavonol glycosides marked with green and red were down- and upregulated in "Yinghong 9" or "Huangyu," respectively. pmp001310,6-Hydroxykaempferol-3,6-O-Diglucoside; Lmmn004912,Quercetin-3-O-methyl ether; Lmyp005280, Isorhamnetin-3-O-[6″-p-coumarylglucosyl-O-rhamnoside; Lmmn003398, Kaempferol-3-O-(6″-acetyl)glucoside; Lmmn004625, Dihydrokaempferol-7-O-glucoside; Lmwp005248, Kaempferol-3-O-(6‴-p-Coumaroylglucosyl)-Glucoside-7-Rhamnoside; mws0913, Kaempferol-3-O-galactoside; mws0032, Myricetin; pmp001105, Kaempferol-3-O-neohesperidoside-7-O-glucoside; pme2493, Kaempferol-3,7-O-dirhamnoside; Lnrp102163, Quercetin-3-O-rutinoside-7-O-rhamnoside; Hmmp002121, Isorhamnetin-3-O-gallate Lmpp003465, Myricetin-3-O-glucoside; Lmtp004126, Myricetin-3-O-(6″-malony)glucoside; pmn001640, Myricetin-3-O-arabinoside. Lmmp002755, Quercetin-7-O-rutinoside-4'-O-glucoside; Lmpn007255, 6-Methoxyquercetin; mws0066, Isorhamnetin; mws1068, Kaempferol; mws1290, Kaempferol-3-O-(6″-p-coumaroyl)glucoside; mws4183, Quercetin-3-*O*-α-L-arabinopyranoside.

"Huangyu" fresh leaves and fermented leaves (**Figure 5**). Twenty-five volatile compounds were identified in "Yinghong 9" fresh leaves, while 21 volatile compounds were identified in "Huangyu" fresh leaves (**Supplementary Table 7**). Higher contents of hexanal, (E)-2-hexenal were detected in "Yinghong 9" – both of which were involved in conferring the grassy-green smell of fresh leaves or finished tea (**Supplementary Table S7**). Meanwhile, higher contents of 3-Hexen-1-ol, acetate, decanal and geraniol were detected in "Yinghong 9." All of them were involved in the flowery-fruity-like flavor of finished tea. But correspondingly, the contents of linalool, methyl salicylate, nerolidol were higher in "Huangyu" than that of "Yinghong 9" fresh leaves.

As confirmed by previous research, new volatile compounds were generated as the fermentation process proceeds. In the present study, the numbers of volatile compounds increased from 22 (fresh leaves) to 35 (fermented leaves) in "Yinghong 9" (**Supplementary Table 7**). Similar increase took place in "Huangyu" cultivar, which was from 22 (fresh leaves) to 35 (fermented leaves). The content of linalool, 1-octen-3-ol, β-ocimene, *cis*-allo-ocimene and 3-Hexen-1-ol, acetate significantly decreased during the fermentation of "Yinghong 9" as well as "Hongyun" cultivars. Some new volatile compounds, e.g., benzaldehyde, benzeneacetaldehyde, linalool oxideI, linalool oxide

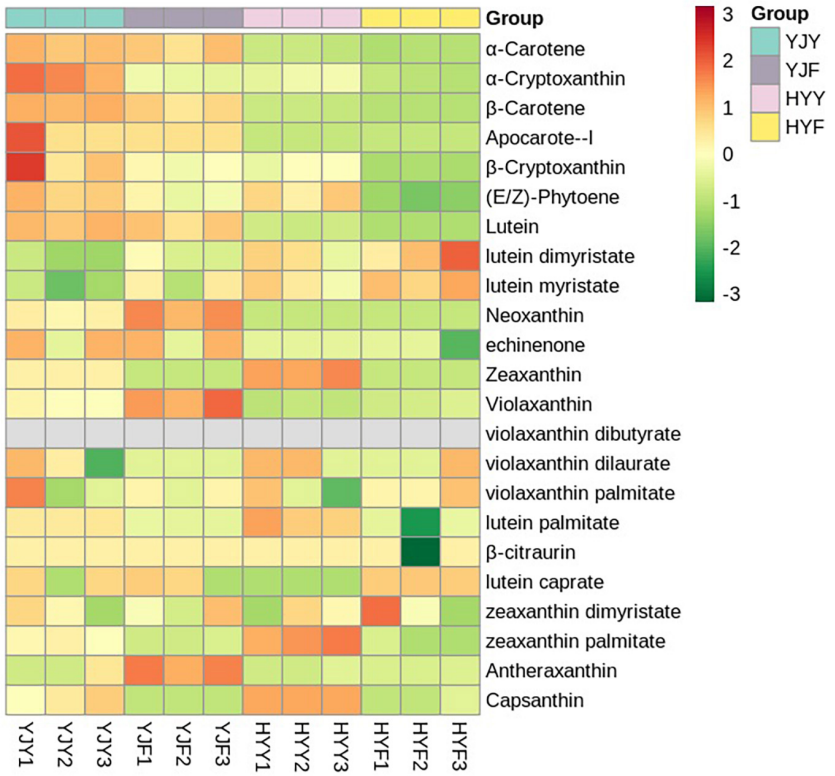


FIGURE 4 | Heat map of volatile compounds clusters in "Yinghong 9" and "Huangyu" fresh and their fermented leaves. Red and green indicate higher and lower abundances, respectively. YJY was "Yinghong 9" fresh leaves, YJF was "Yinghong 9" fermented leaves; HYY was "Huangyu" fresh leaves, HYF was "Huangyu" fermented leaves.

II,3-hexenyl-α-methylbutyrate, 2-hexenyl isovalerate, geraniol, *cis*-3-hexenyl hexanoate, *cis*-hexenyl hexanoate, *cis*-2-hexenyl hexanoate, α-bergamotene, 6,10-dimethyl-5,9-undecadien-2-one (*trans*-geranylacetone) and nerolidol were detected in "Yinghong 9" and "Huangyu" fermented leaves. At the same time, the contents of α-farnesene, β-ionone, β-cyclocitral and nerolidol in fermented leaves were significantly higher than those of fresh leaves.

Difference of Enzyme Activities in Fresh and Fermented Leaves in "Yinghong 9" and "Huangyu"

To reveal the manufacture suitability of "Yinghong 9" and "Huangyu," the enzyme activity of polyphenol oxidase β-glucosidase and carotenoid cleavage dioxygenase were measured in their fresh leaves and fermented leaves. The results showed that the activities of polyphenol oxidase, β-glucosidase in "Yinghong 9" were higher than that of "Huangyu," Similarly, the activity of carotenoid cleavage dioxygenase (CCD) showed also had the same change tendency in "Yinghong 9" and "Huangyu" fresh leaves (**Figure 6**). The difference was that the activity of polyphenol oxidase in fresh taste was higher than that of fermented leaves, but the activities of the other two enzymes, (β-glucosidase and carotenoid cleavage dioxygenase) decreased significantly in fermented leaves.

B-Glucosidase is one of the important enzymes involved in tea aroma compounds formation (Zeng et al., 2019), which can hydrolyze hydrolysis of β-1,4- glycosidic linkages and release free volatile compounds such as linalool, benzaldehyde, and methyl salicylate during tea processing (Yang, 2013). these compounds had flowery fruity fragrances and were beneficial to the aroma quality of black tea. As shown in **Figure 6**, the activity of the β-glucosidase in "Huangyu" was significantly higher than that of "Yinghong 9." The activity of the β-glucosidase decreased significantly in "Yinghong 9" and "Huangyu" fermented leaves. The decrease in enzyme activity in the fermented stage might be due to the interaction between the enzyme and polyphenol compounds, as well as the decrease of substrate concentration (Supriyadi et al., 2021). So, after the withered tea leaf was subjected to rolling, the increase of tea polyphenols and substrates decreases led to a decrease in the activity of the β-glucosidase during fermentation processing. This mechanism has been verified by some research findings (Wang et al., 2001a,b).

Wang et al. (2020) had confirmed that both carotenoid cleavage dioxygenase1 (CsCCD1) and carotenoid cleavage dioxygenase4 (CsCCD4) are involved in the formation of β-ionone during tea manufacture. In this study, the carotenoid cleavage dioxygenase activity was significantly higher in

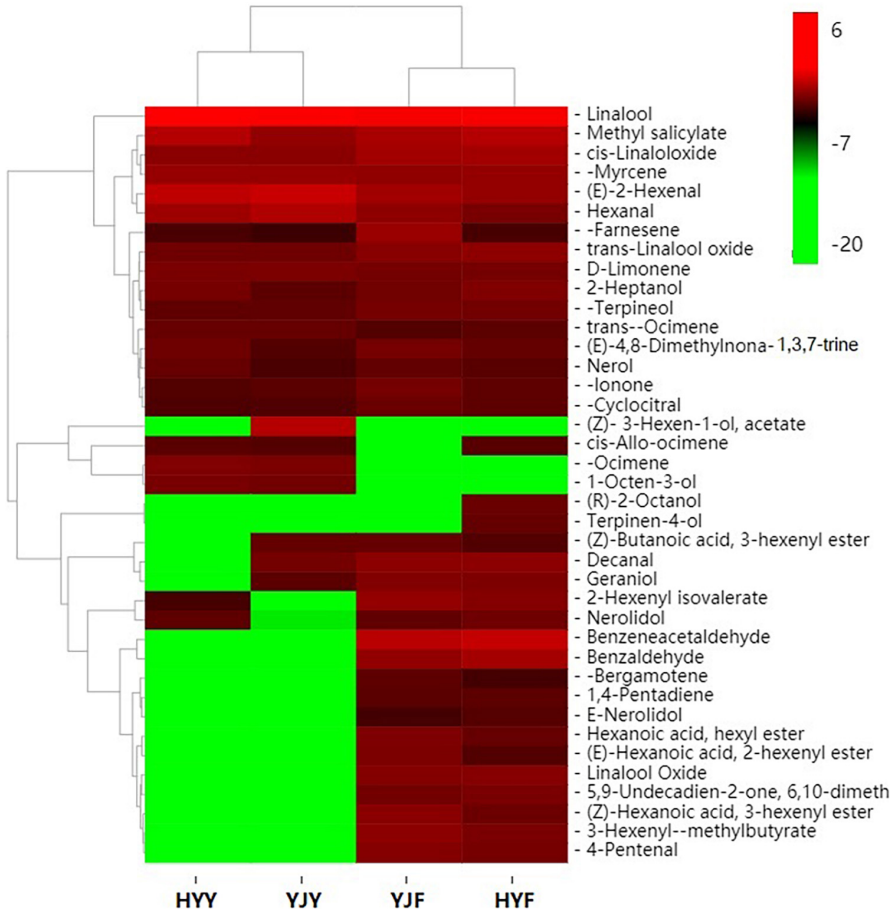


FIGURE 5 | Heat map of carotenoids metabolite clusters in "Yinghong 9" and "Huangyu" fresh and their fermented leaves. Red and green indicate higher and lower abundances in respective. YJY was "Yinghong 9" fresh leaves, YJF was "Yinghong 9" fermented leaves; HYY was "Huangyu" fresh leaves, HYF was "Huangyu" fermented leaves.

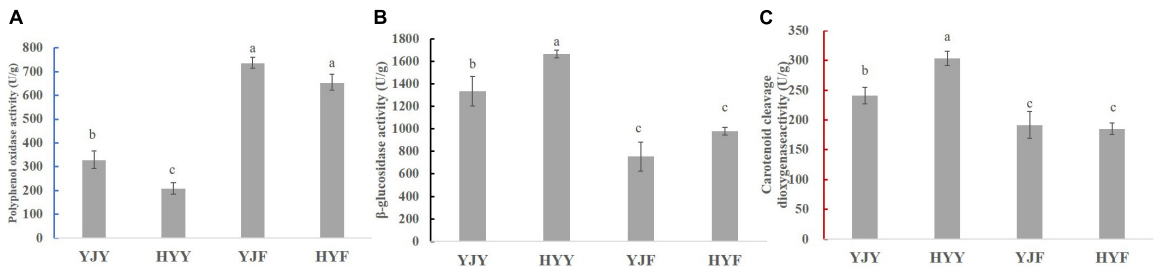


FIGURE 6 | Comparison of enzyme activities in fresh and fermented leaves of "Yinghong 9" and "Huangyu." **(A)** Polyphenol oxidase activity of "Yinghong 9" and "Huangyu" leaves. **(B)** β-glucosidase activity of "Yinghong 9" and "Huangyu" leaves. **(C)** Carotenoid cleavage dioxygenase activity of "Yinghong 9" and "Huangyu" leaves. All the data are presented as the mean standard deviation (n = 3). Means with different letters at each treatment represented a significant difference at $p \leq 0.05$. YJY was "Yinghong 9" fresh leaves, YJF was "Yinghong 9" fermented leaves; HYY was "Huangyu" fresh leaves, HYF was "Huangyu" fermented leaves.

"Huangyu" than in "Yinghong 9" fresh leaves (**Figure 6**). That might be the reason why the content of aroma components was still high in "Huangyu" fresh leaves, although the total amount of carotenoids was lower than "Yinghong 9." The results might be partially due to aroma compounds coming from glycoside hydrolysis and carotenoid degradation pathways (**Figure 7**). Now that there was a higher content of zeaxanthin and E-nerolidol, 6,10-dimethyl-5,9-undecadien-2-one, nerolidol in "Huangyu," which were the degradation product of carotenoids, the degradation mechanism of carotenoids needs to be studied further. The results also suggested that "Huangyu" was still a variety which had good manufacture suitability. Furthermore,

now that the decrease of carotenoid cleavage dioxygenase activity in fermented leaves indicates that the "Huangyu" variety is more suitable for manufacturing non-fermented tea.

DISCUSSION

The Yellow Leaf Phenotype Is Closely Associated With Chlorophyll and Carotenoid Metabolism

Tea, produced from fresh tea (*Camellia sinensis*) leaves, is an important, widely popular non-alcohol beverage due to its health benefit functions, as well as its pleasant flavor. Unlike most other agricultural products, the color of infused leaves is involved in the sensory quality formation of tea. Thus, recently the researchers paid more attention to the study on the relationship of leaf color and characteristic compounds, and its influence on health benefits. In fact, compared to typical green leaves, several albino tea varieties have been successfully developed into new products because of their precious and tasteful quality, e.g., "Anji baicha," "Yujingxiang," and "Huangjinya." Unluckily, tea breeding goals have been limited to a brighter color or better taste. Other valuable traits such as yield, flavor value, and manufacture suitability have long been ignored.

Previous studies indicated that the formation of tea leaf color from white to purple might be due to changes in different pigments and their ratios. In general, the greening of leaves was mainly due to the higher concentrations of chlorophylls, and the albino-induced white and yellow leaves were mainly due to lower concentrations of chlorophyll, while certain concentrations of carotenoids dominate in leaf formation (Zhu et al., 2015; Tsai et al., 2017). Similar research results had been proven in special light-sensitive albino yellow leaves cultivars such as "Yujinxiang," "Huangjinya" (Song et al., 2017). The current study results showed that total chlorophyll was significantly higher contents in "Yinghong 9" normal green leaves than in the "Huangyu" yellow leaves (**Figure 2A**). The content of tea polyphenols and carotenoids changed in a manner similar to that observed for chlorophyll in two cultivars. The content of tea polyphenols, total chlorophyll, and carotenoid decreased significantly during the fermentation of "Yinghong 9" and "Huangyu" leaves.

As previously reported, carotenoids as an important pigment in green leaves could maintain efficient photosynthesis, by scavenging various reactive oxygen free radicals and protecting chlorophylls from photooxidation. The low chlorophyll content in yellow leaves was due to the blocked chlorophyll synthesis and enhanced degradation of intermediate products. It was speculated that the yellow leaves under the strong light of "Huangjinya" were due to the color presentation of carotenoids and flavonoids after chlorophyll content was significantly reduced. In this study, the types and contents of carotenoids were significant differences between "Yinghong 9" and "Huangyu." Compared with "Huangyu" fresh leaves, the content of lutein, α-carotene, (E/Z)-phytoene, β-cryptoxanthin, β-carotene, and neoxanthin were significantly higher in "Yinghong 9." In comparison, the content of zeaxanthin, and zeaxanthin palmitate were significantly higher in "Huangyu" leaves. Previous studies showed that the changes in the abundances of carotenoids were closely relevant to the expression of carotenoid biosynthesis genes (Zhang et al., 2012).

In plants, zeaxanthin is an essential metabolite of β-carotene. β-carotene generates zeaxanthin through the catalysis of β-carotene hydroxylase (BCH) (Kim et al., 2009). Silencing the BCH genes leads to an increase in the levels of β-carotene and carotenoids but reduces that of zeaxanthin in potato tubers (Diretto et al., 2007). Similarly, the expression levels of the β-arotene hydroxylase gene were up-regulated in yellow leaf wheat, which enhanced the metabolic flux from β-carotene to zeaxanthin, and resulted in increasing the content of zeaxanthin and decreasing the content of β-carotene (Wu et al., 2018). Similar results had been reported in light-sensitive tea "Huangjinya" and "Yujinxiang" cultivars (Liu et al., 2017). The difference was that the content of lutein, β-carotene and zeaxanthin was higher in "Huangjinya," while zeaxanthin, carotene, violaxanthin, cryptoxanthin, and lutein content were higher in "Yujinxiang." Interestingly, as downstream products of zeaxanthin, the content of violaxanthin, neoxanthin, and antheraxanthin was higher in "Yinghong 9" than in "Huangyu." The difference could be due to the synthesis of the downstream product resulting in reducing substrate content. Altogether, despite the content of total carotenoid being higher in "Yinghong 9," chlorophyll might conceal the coloration of carotenoids and result in the formation of green in "Yinghong 9" normal leaf. But for "Huangyu," the absence of chlorophyll in leaf, leaf coloration is predominated by carotenoids.

The Type and Content of Flavonol Glycosides in "Yinghong 9" and "Huangyu" Lead to Their Different Manufacture Suitability

Flavonol glycosides were one of the most important groups of polyphenols besides catechins in tea. Previous studies showed that flavonol glycosides were the critical ingredient of green tea soup, in which they color the tea liquor as yellow or yellow-green pigments (Dai et al., 2017). Like catechins, flavonol-glycosides play an important role in forming the tea taste (Scharbert et al., 2004; Scharbert and Hofmann, 2005). Not only can they engender dry and smooth-astringency sensations in the mouth, but also, they are involved in enhancing the bitterness of tea soup. The previous researchers confirmed that kaempferol-3-glucoside and kaempferol-3-galactoside were both astringent components, of which the taste threshold of kaempferol-3-glucoside and kaempferol-3-galactoside were 0.67 and 6.7 μmol/L, respectively (Scharbert et al., 2004; Scharbert and Hofmann, 2005). The difference of taste threshold of the two flavonol-glycosides was up to ten times. In contrast, the taste thresholds of quercetin-3-glucoside and quercetin-3-galactoside were 0.65 and 0.43 μmol/L, respectively, with a difference of 1.5 times. Therefore, the ratio of kaempferol-3- glucoside to kaempferol-3-galactoside had an important impact on tea taste.

Glucosylation and galactosylation were two competing glycosylation metabolic ways. Simultaneously the content

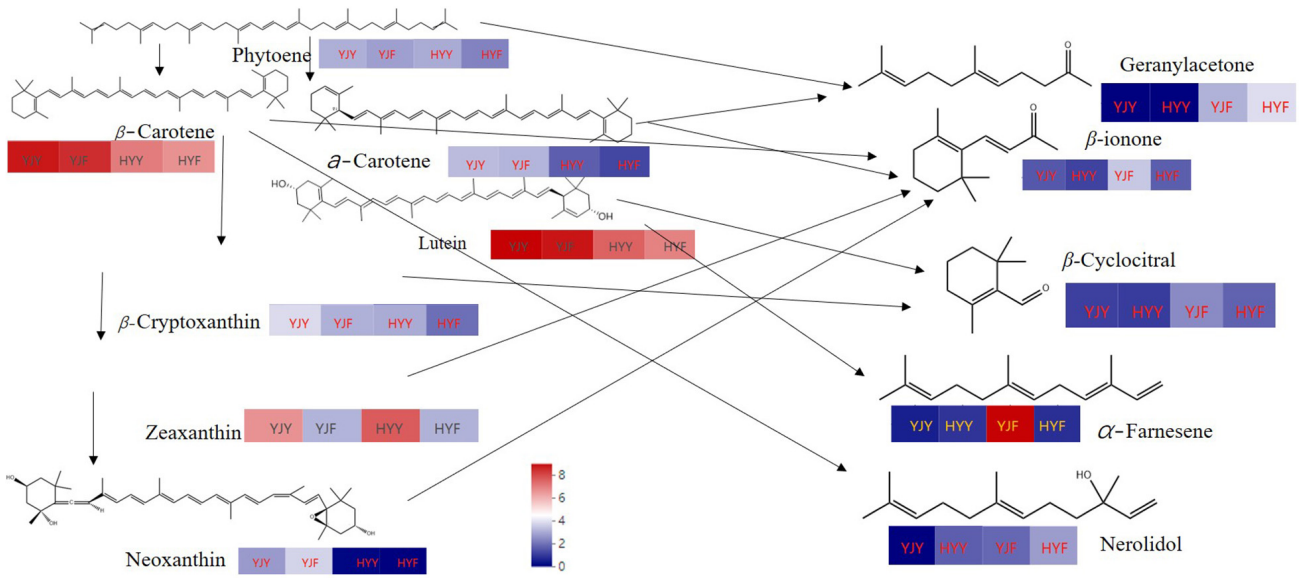


FIGURE 7 | The content of main types of carotenoids metabolism schematic and degradation products (main aroma compounds). Red and blue indicate higher and lower abundances, respectively. YJY was "Yinghong 9" fresh leaves, YJF was "Yinghong 9" fermented leaves; HYY was "Huangyu" fresh leaves, HYF was "Huangyu" fermented leaves.

of them varies among different tea varieties. The content of galactosylated flavonol in tea shoots leads to better manufacturing-suitability of green tea. In contrast, the content of glucosylated flavonol was significantly higher in black tea manufacturing-suitability varieties (Dai et al., 2017). This implies that the variety with more galactosylation of flavonol is suitable for manufacturing green tea and white tea (non-fermented tea), while the variety with more glucosylation of flavonol may be suitable for manufacturing black tea and oolong tea (fermented tea) (Jiang et al., 2015).

In this study, the content of kaempferol-3- glucoside in "Yinghong 9" was 1.5 times more than that of "Huangyu." And the ratio kaempferol-3-glucoside to kaempferol-3-galactoside was lower, which indicates that the glycosylation metabolism of kaempferol in tea plants flowed toward galactosylation, which was beneficial for reducing the content of kaempferol-3-glucoside. So, "Huangyu" is suitable for manufacturing non-fermented tea, e.g., green tea and white tea. Correspondingly, "Yinghong 9" was suitable for manufacturing black tea since it had a high concentration of kaempferol-3-glucoside as well as a lower ratio of kaempferol-3-glucoside to kaempferol-3-galactoside (**Supplementary Table 5**). Liu et al. (2017) found that compared to the green variety, metabolic flux convert to flavonoid and carotenoids pathways to a certain extent in "Yujingxiang" yellow leaf variety. The change of metabolism flux might enhance the production of the antioxidant quercetin or quercetin glycosides rather than catechin biosynthesis. However, our present results were not consistent in this conclusion. It seemed that the yellow leaf "Huangyu" as a rare variety was worth further study because of its complex metabolic flux.

The Differences of Glycosides Hydrolysis and Carotenoids Degradation Might Contribute to "Huangyu"xs Fresh Leaves Aroma Quality

Tea aroma quality, which is composed of the volatile compounds that originated during tea processing, is an important quality factor and determines dry tea's economic benefit. A large number of literatures was concerned about analysis technology, non-carotenoid volatile compounds formation during tea processing, as well as the difference of varieties. The relationship between the carotenoids and their volatile components in yellow leaf tea varieties has received little attention. The formation of volatile compounds from carotenoids comes from the various degradation pathways. Evidence showed that β-ionone, a critical flowery component was the product of β-carotene oxidation degradation, in addition, linalool, nerolidol, β-cyclocitral and geranylacetone were degradation products of other carotenes present in tea (Ravichandran, 2002). In fact, as critical yellow pigments, carotenoids were present in nearly all green or yellow leaf tea, and their degradation promotes the formation of flowery aroma components during black tea processing (Hazarika and Mahanta, 2010). It is interesting to note that Ravichandran has reported that compared with Assam varieties, China varieties had more carotenoids in fresh leaves (Ravichandran, 2002), which implied that China varieties were better manufacture suitability for high aroma black tea. We really should pay our attention to China's large leaf and yellow leaf mutant variety such as "Yinghong 9" and "Huangyu."

As stated above, β-Glucosidase was involved in tea aroma compounds formation, which could hydrolyze hydrolysis of β-1,4- glycosidic linkages and release free volatile compounds

such as linalool, benzaldehyde and methyl salicylate during tea processing (Yang, 2013). Our study results showed that the relatively high concentration of linalool, methyl salicylate, nerolidol was detected in "Huangyu" fresh leaves. Accordingly, the activity of the β-glucosidase in "Huangyu" was significantly higher than that of "Yinghong 9" (**Figure 6B**).

Similarly, the content of zeaxanthin, and zeaxanthin palmitate were significantly higher in "Huangyu" leaves, while the contents of nerolidol, α-farnesene and geranylacetone were significantly low in "Yinghong 9" fresh leaves, but the contents of these compounds were higher in 'Huangyu' fresh leaves. Correspondingly, the carotenoid cleavage dioxygenase activity was significantly higher in "Huangyu" than in "Yinghong 9" fresh leaves (**Figure 6C**). Due to all three of the compounds were carotenoids degradation products, it was reasonable to presume that the differences of carotenoids might influence aroma compounds in "Yinghong 9" and "Huangyu" fresh leaves. Despite some new volatile compounds such as benzaldehyde, benzeneacetaldehyde, linalool oxideI, linalool oxideII, and so on were detected in "Yinghong 9" and "Huangyu" fermented leaves. Particularly, the contents of α-farnesene, β-ionone, β-cyclocitral, and geranylacetone had a fruity floral odor and contribute significantly to black tea aroma quality were increased in "Yinghong 9" fermented leaves. However, the activities of β-glucosidase and carotenoid cleavage dioxygenase decreased greatly in "Huangyu" fermented leaves (**Figures 6B,C**). All these results suggested that the aroma quality of "Huangyu" came from hydrolyzing hydrolysis and carotenoid degradation in fresh leaves, and it was suitable for manufacturing non-fermented tea, e.g., green tea, white tea, or even yellow tea. While the typical aroma components of "Yinghong 9" came from fermentation processing, it was suitable for manufacturing fermented tea.

CONCLUSION

In this study, analysis of non-volatiles and volatiles was conducted to identify the major metabolite compounds in "Yinghong 9" and its yellow-leaf mutant "Huangyu." The contents of total carotenoids and chlorophyll revealed that chlorophyll concealed the coloration of carotenoids and led to the formation of normal green leaf in "Yinghong 9." But for "Huangyu," being the absence of chlorophyll in leaf, leaf coloration was predominated by carotenoids. The content of kaempferol-3- glucoside in "Yinghong 9" was 1.5 fold more than that of "Huangyu," and the ratio of kaempferol-3-glucoside to kaempferol-3-galactoside was lower, indicating that "Huangyu" was suitable for manufacturing non-fermented tea, e.g., green tea and white tea, while "Yinghong 9" was suitable for manufacturing black tea. Analysis of carotenoid metabolites revealed that the contents of lutein and α-carotene were significantly higher in "Yinghong 9," while zeaxanthin was the dominant compound in "Huangyu" leaves. In addition, the contents of nerolidol, α-farnesene and geranylacetone were significantly higher. And the higher activity of β-glucosidase and carotenoid cleavage dioxygenase in "Huangyu" fresh leaves also implied that "Huangyu" was suitable for manufacturing green tea. At the same time, nerolidol and geranyl acetone, the carotenoid degradation products, increased in "Yinghong 9" fermented leaves, which meant "Yinghong 9" was suitable for manufacturing black tea. Our current results provided a basis for understanding the mechanisms of yellow-leaf formation in "Huangyu" and facilitated the effective utilization for particular natural mutants.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

AUTHOR CONTRIBUTIONS

XM designed the experiment. CL and XM carried out most of the experiments. LZ drafted the manuscript. BC and SW did the data analyses. LZ and HW revised and finalized the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported financially by the National Key Research & Development Program (2018YFD1000601); the National Natural Science Foundation of China (32072628); Fund for Less Developed Regions of the National Natural Science Foundation of China (32160727); Qiannan Normal University for Nationalities (2020qnsyrc08); Key-Area Research and Development Program of Guangdong Province (2020B020220004); and China Agriculture Research System of MOF and MARA.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2021.767724/full#supplementary-material

Supplementary Table 1 | Representative differential metabolites between "Yinghong 9" and "Huangyu" fresh leaves.

Supplementary Table 2 | Differential metabolites between "Yinghong 9" fresh leaves and fermented leaves.

Supplementary Table 3 | Differential metabolites between "Huangyu" fresh leaves and fermented leaves.

Supplementary Table 4 | Representative flavonol glycosides between "Huangyu" fresh leaves and fermented leaves.

Supplementary Table 5A | Differential flavonol glycosides between "Yinghong 9" fresh leaves and fermented leaves.

Supplementary Table 5B | Representative flavonol glycosides between "Huangyu" fresh leaves and fermented leaves.

Supplementary Table 6 | The quantification results of all detected carotenoids.

Supplementary Table 7 | The results of volatile compounds in all samples.

REFERENCES

Cheng, S., Fu, X., Liao, Y., Xu, X., Zeng, L., Tang, J., et al. (2019). Differential accumulation of specialized metabolite l-theanine in green and albino-induced yellow tea (Camellia sinensis) leaves. *Food Chem.* 276, 93–100. doi: 10.1016/j.foodchem.2018.10.010

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

Clément, A., Verfaille, T., Lormel, C., and Jaloux, B. (2015). A new colour vision system to quantify automatically foliar discolouration caused by insect pests feeding on leaf cells. *Biosyst. Enginee.* 133, 128–140.

Dai, W., Xie, D., Meiling, L., Tan, J., Pengliang, L. I., Haipeng, L., et al. (2017). Relationship of Flavonol Glycoside and Processing Suitability of Tea Varieties. *Food Sci.* 96, 40–45.

Diretto, G., Welsch, R., Tavazza, R., Mourgues, F., Pizzichini, D., Beyer, P., et al. (2007). Silencing of beta-carotene hydroxylase increases total carotenoid and beta-carotene levels in potato tubers. *BMC Plant Biol.* 7:11. doi: 10.1186/1471-2229-7-11

Feng, L., Gao, M. J., Hou, R. Y., Hu, X. Y., Zhang, L., Wan, X. C., et al. (2014). Determination of quality constituents in the young leaves of albino tea cultivars. *Food Chem.* 155, 98–104. doi: 10.1016/j.foodchem.2014.01.044

Gitelson, A. (2020). Towards a generic approach to remote non-invasive estimation of foliar carotenoid-to-chlorophyll ratio. *J. Plant Physiol.* 252:153227. doi: 10.1016/j.jplph.2020.153227

Hao, X., Zhang, W., Liu, Y., Zhang, H., Ren, H., Chen, Y., et al. (2020). Pale green mutant analyses reveal the importance of CsGLKs in chloroplast developmental regulation and their effects on flavonoid biosynthesis in tea plant. *Plant Physiol. Biochem.* 146, 392–402. doi: 10.1016/j.plaphy.2019.11.036

Hazarika, M., and Mahanta, P. K. (2010). Some studies on carotenoids and their degradation in black tea manufacture. *J. Sci. Food Agricult.* 34, 1390–1396. doi: 10.1002/jsfa.2740341212

Ho, C. T., Zheng, X., and Li, S. (2015). Tea aroma formation. *Food Sci. Hum. Wellness* 4, 9–27. doi: 10.1016/j.fshw.2015.04.001

Huot, B., Yao, J., Montgomery, B. L., and He, S. Y. (2014). Growth-defense tradeoffs in plants: a balancing act to optimize fitness. *Mol. Plant* 7, 1267–1287. doi: 10.1093/mp/ssu049

Jiang, H., Engelhardt, U. H., Thraene, C., Maiwald, B., and Stark, J. (2015). Determination of flavonol glycosides in green tea, oolong tea and black tea by UHPLC compared to HPLC. *Food Chem.* 183, 30–35. doi: 10.1016/j.foodchem.2015.03.024

Kawakami, M. (2002). *Carotenoid-Derived Aroma Compounds in Tea*. Washington, DC: Oxford University Press Inc.

Kim, J., Smith, J. J., Tian, L., and Dellapenna, D. (2009). The evolution and function of carotenoid hydroxylases in Arabidopsis. *Plant Cell Physiol.* 50, 463–479. doi: 10.1093/pcp/pcp005

Li, Q., Huang, J., Liu, S., Li, J., Yang, X., Liu, Y., et al. (2011). Proteomic analysis of young leaves at three developmental stages in an albino tea cultivar. *Proteome Sci.* 9:44. doi: 10.1186/1477-5956-9-44

Li, S., Wang, S., Wang, P., Gao, L., Yang, R., and Li, Y. (2020). Label-free comparative proteomic and physiological analysis provides insight into leaf color variation of the golden-yellow leaf mutant of Lagerstroemia indica. *J. Proteomics* 228:103942. doi: 10.1016/j.jprot.2020.103942

Liang, Y., Lu, J., Zhang, L., Wu, S., and Wu, Y. (2003). Estimation of black tea quality by analysis of chemical composition and colour difference of tea infusions. *Food Chem.* 80, 283–290. doi: 10.1016/s0308-8146(02)00415-6

Liu, G. F., Han, Z. X., Feng, L., Gao, L. P., Gao, M. J., Gruber, M. Y., et al. (2017). Metabolic Flux Redirection and Transcriptomic Reprogramming in the Albino Tea Cultivar 'Yu-Jin-Xiang' with an Emphasis on Catechin Production. *Sci. Rep.* 7:45062. doi: 10.1038/srep45062

Long, M., Millar, D. J., Kimura, Y., Donovan, G., Rees, J., Fraser, P. D., et al. (2006). Metabolite profiling of carotenoid and phenolic pathways in mutant and transgenic lines of tomato: identification of a high antioxidant fruit line. *Phytochemistry* 67, 1750–1757. doi: 10.1016/j.phytochem.2006.02.022

Lu, M., Han, J., Zhu, B., Jia, H., Yang, T., Wang, R., et al. (2019). Significantly increased amino acid accumulation in a novel albino branch of the tea plant (Camellia sinensis). *Planta* 249, 363–376. doi: 10.1007/s00425-018-3007-6

Ma, C., Cao, J., Li, J., Zhou, B., Tang, J., and Miao, A. (2016). Phenotypic, histological and proteomic analyses reveal multiple differences associated with chloroplast development in yellow and variegated variants from Camellia sinensis. *Sci. Rep.* 6:33369. doi: 10.1038/srep33369

Mu, B., Zhu, Y., Lv, H. P., Yan, H., Peng, Q. H., and Lin, Z. (2018). The enantiomeric distributions of volatile constituents in different tea cultivars. *Food Chem.* 265, 329–336. doi: 10.1016/j.foodchem.2018.05.094

Ningrum, A., and Schreiner, M. (2014). Carotenoid-cleavage activities of crude enzymes from Pandanous amryllifolius. *Chem. Biodivers.* 11, 1871–1881. doi: 10.1002/cbdv.201400029

Ravichandran, R. (2002). Carotenoid composition, distribution and degradation to flavour volatiles during black tea manufacture and the effect of carotenoid supplementation on tea quality and aroma. *Food Chem.* 78, 23–28.

Rothenberg, D. O., Yang, H., Chen, M., Zhang, W., and Zhang, L. (2019). Metabolome and Transcriptome Sequencing Analysis Reveals Anthocyanin Metabolism in Pink Flowers of Anthocyanin-Rich Tea (Camellia sinensis). *Molecules* 24:1064. doi: 10.3390/molecules24061064

Sanderson, G. W., and Grahamm, H. N. (1973). Formation of black tea aroma. *J. Agricult. Food Chem.* 21, 576–585. doi: 10.1021/jf60188a007

Scharbert, S., and Hofmann, T. (2005). Molecular definition of black tea taste by means of quantitative studies, taste reconstitution, and omission experiments. *J. Agric. Food Chem.* 53, 5377–5384. doi: 10.1021/jf050294d

Scharbert, S., Holzmann, N., and Hofmann, T. (2004). Identification of the astringent taste compounds in black tea infusions by combining instrumental analysis and human bioresponse. *J. Agric. Food Chem.* 52, 3498–3508. doi: 10.1021/jf049802u

Shi, J. (2018). Characterization of Volatile Components of Eight FengHuang Dancong Manufactured Teas and Fresh Leaves by HS-SPME Coupled with GC-MS. *Int. J. Nutrit. Food Sci.* 7:160.

Solovchenko, A. (2010). "Photoprotection in Plants : Optical Screening-based Mechanisms," in *Photoprotection in Plants:optical Screening-based Mechanisms*, ed. A. Solovchenko (Berlin: Springer).

Song, L., Ma, Q., Zou, Z., Sun, K., Yao, Y., Tao, J., et al. (2017). Molecular Link between Leaf Coloration and Gene Expression of Flavonoid and Carotenoid Biosynthesis in Camellia sinensis Cultivar 'Huangjinya'. *Front. Plant Sci.* 8:803. doi: 10.3389/fpls.2017.00803

Supriyadi, S., Nareswari, A. R., Fitriani, A., and Gunadi, R. (2021). Enhancement of Black Tea Aroma by Adding the β-Glucosidase Enzyme during Fermentation on Black Tea Processing. *Int. J. Food Sci.* 2021:5542109. doi: 10.1155/2021/5542109

Tian, Y., Wang, H., Sun, P., Fan, Y., Qiao, M., Zhang, L., et al. (2019). Response of leaf color and the expression of photoreceptor genes of Camellia sinensis cv. Huangjinya to different light quality conditions. *Sci. Horticult.* 251, 225–232. doi: 10.1016/j.scienta.2019.03.032

Tsai, C. C., Wu, Y. J., Sheue, C. R., Liao, P. C., Chen, Y. H., Li, S. J., et al. (2017). Molecular Basis Underlying Leaf Variegation of a Moth Orchid Mutant (Phalaenopsis aphrodite subsp. formosana). *Front Plant Sci.* 8:1333. doi: 10.3389/fpls.2017.01333

Wang, D., Kubota, K., Kobayashi, A., and Juan, I. M. (2001a). Analysis of glycosidically bound aroma precursors in tea leaves. 3. Change in the glycoside content of tea leaves during the oolong tea manufacturing process. *J. Agric. Food Chem.* 49, 5391–5396. doi: 10.1021/jf010235+

Wang, D., Kurasawa, E., Yamaguchi, Y., Kubota, K., and Kobayashi, A. (2001b). Analysis of glycosidically bound aroma precursors in tea leaves. 2. Changes in glycoside contents and glycosidase activities in tea leaves during the black tea manufacturing process. *J. Agric. Food Chem.* 49, 1900–1903. doi: 10.1021/jf001077+

Wang, J., Zhang, N., Zhao, M., Jing, T., Jin, J., Wu, B., et al. (2020). Carotenoid Cleavage Dioxygenase 4 Catalyzes the Formation of Carotenoid-Derived Volatile beta-Ionone during Tea (Camellia sinensis) Withering. *J. Agric. Food Chem.* 68, 1684–1690.

Wang, K. R., Li, N. N., Du, Y. Y., and Liang, Y. (2013). Effect of sunlight shielding on leaf structure and amino acids concentration of light sensitive albino tea plant. *Afr. J. Biotechnol.* 12, 5535–5539.

Wu, H., Huang, W., Chen, Z., Chen, Z., Shi, J., Kong, Q., et al. (2019). GC-MS-based metabolomic study reveals dynamic changes of chemical compositions during black tea processing. *Food Res. Int.* 120, 330–338. doi: 10.1016/j.foodres.2019.02.039

Wu, H., Shi, N., An, X., Liu, C., Fu, H., Cao, L., et al. (2018). Candidate Genes for Yellow Leaf Color in Common Wheat (Triticum aestivum L.) and Major Related Metabolic Pathways according to Transcriptome Profiling. *Int. J. Mol. Sci.* 19:1594. doi: 10.3390/ijms19061594

Yan, N., Du, Y., Liu, X., Chu, M., Shi, J., Zhang, H., et al. (2019). A comparative UHPLC-QqQ-MS-based metabolomics approach for evaluating Chinese and North American wild rice. *Food Chem.* 275, 618–627. doi: 10.1016/j.foodchem.2018.09.153

Yang, Z. (2013). Recent studies of the volatile compounds in tea. *Food Res. Int.* 2, 585–599. doi: 10.1016/j.foodres.2013.02.011

Zeng, L., Watanabe, N., and Yang, Z. (2019). Understanding the biosyntheses and stress response mechanisms of aroma compounds in tea (Camellia sinensis) to safely and effectively improve tea aroma. *Crit. Rev. Food Sci. Nutr.* 59, 2321–2334. doi: 10.1080/10408398.2018.1506907

Zhang, L., Ma, G., Kato, M., Yamawaki, K., Takagi, T., Kiriiwa, Y., et al. (2012). Regulation of carotenoid accumulation and the expression of carotenoid metabolic genes in citrus juice sacs in vitro. *J. Exp. Bot.* 63, 871–886. doi: 10.1093/jxb/err318

Zhang, S., Xuan, H., Zhang, L., Fu, S., Wang, Y., Yang, H., et al. (2017). TBC2health: a database of experimentally validated health-beneficial effects of tea bioactive compounds. *Brief Bioinform.* 18, 830–836. doi: 10.1093/bib/bbw055

Zhou, C., Mei, X., Rothenberg, D. O., Yang, Z., Zhang, W., Wan, S., et al. (2020). Metabolome and Transcriptome Analysis Reveals Putative Genes Involved in Anthocyanin Accumulation and Coloration in White and Pink Tea (Camellia sinensis) Flower. *Molecules* 25:190. doi: 10.3390/molecules25010190

Zhou, W., Niu, Y., Ding, X., Zhao, S., Li, Y., Fan, G., et al. (2020). Analysis of carotenoid content and diversity in apricots (Prunus armeniaca L.) grown in China. *Food Chem.* 330:127223. doi: 10.1016/j.foodchem.2020.127223

Zhu, G., Yang, F., Shi, S., Li, D., Wang, Z., Liu, H., et al. (2015). Transcriptome Characterization of Cymbidium sinense 'Dharma' Using 454 Pyrosequencing and Its Application in the Identification of Genes Associated with Leaf Color Variation. *PLoS One* 10:e0128592. doi: 10.1371/journal.pone.0128592

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.

frontiers in Plant Science

ORIGINAL RESEARCH
published: 15 December 2021
doi: 10.3389/fpls.2021.761862

Comparative Transcriptome Analysis Identifies Key Regulatory Genes Involved in Anthocyanin Metabolism During Flower Development in *Lycoris radiata*

*Ning Wang[1,2], Xiaochun Shu[1,2], Fengjiao Zhang[1,2], Weibing Zhuang[1,2], Tao Wang[1,2] and Zhong Wang[1,2]**

[1] Institute of Botany, Jiangsu Province and Chinese Academy of Sciences, Nanjing, China, [2] Jiangsu Key Laboratory for the Research and Utilization of Plant Resources, Institute of Botany, Jiangsu Province and Chinese Academy of Sciences, Nanjing, China

OPEN ACCESS

Edited by:
Xiumin Fu,
South China Botanical Garden, Chinese Academy of Sciences (CAS), China

Reviewed by:
Mingyue Fu,
Sichuan University, China
Md Abdur Rahim,
Sher-e-Bangla Agricultural University, Bangladesh

***Correspondence:**
Zhong Wang
wangzhong19@163.com

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

Received: 20 August 2021
Accepted: 17 November 2021
Published: 15 December 2021

Citation:
Wang N, Shu X, Zhang F, Zhuang W, Wang T and Wang Z (2021) Comparative Transcriptome Analysis Identifies Key Regulatory Genes Involved in Anthocyanin Metabolism During Flower Development in Lycoris radiata. Front. Plant Sci. 12:761862. doi: 10.3389/fpls.2021.761862

Lycoris is used as a garden flower due to the colorful and its special flowers. Floral coloration of *Lycoris* is a vital trait that is mainly regulated via the anthocyanin biosynthetic pathway. In this study, we performed a comparative transcriptome analysis of *Lycoris radiata* petals at four different flower development stages. A total of 38,798 differentially expressed genes (DEGs) were identified by RNA sequencing, and the correlation between the expression level of the DEGs and the anthocyanin content was explored. The identified DEGs are significantly categorized into 'flavonoid biosynthesis,' 'phenylpropanoid biosynthesis,' 'Tropane, piperidine and pyridine alkaloid biosynthesis,' 'terpenoid backbone biosynthesis' and 'plant hormone signal transduction' by Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analysis. The candidate genes involved in anthocyanin accumulation in *L. radiata* petals during flower development stages were also identified, which included 56 structural genes (especially *LrDFR1* and *LrFLS*) as well as 27 key transcription factor DEGs (such as *C3H*, *GATA*, *MYB*, and *NAC*). In addition, a key structural gene namely *LrDFR1* of anthocyanin biosynthesis pathway was identified as a hub gene in anthocyanin metabolism network. During flower development stages, the expression level of *LrDFR1* was positively correlated with the anthocyanin content. Subcellular localization revealed that LrDFR1 is majorly localized in the nucleus, cytoplasm and cell membrane. Overexpression of *LrDFR1* increased the anthocyanin accumulation in tobacco leaves and *Lycoris* petals, suggesting that *LrDFR1* acts as a positively regulator of anthocyanin biosynthesis. Our results provide new insights for elucidating the function of anthocyanins in *L. radiata* petal coloring during flower development.

Keywords: *Lycoris radiata*, transcriptome, anthocyanin, structural genes, phytohormone, transcription factors, dihydroflavonol 4-reductase

INTRODUCTION

Plant pigments, such as anthocyanins, carotenoids and chlorophylls, play important roles in affecting the appearance of flower, fruit and seed coloring (Tanaka et al., 2008; Rebecca et al., 2010; Rosas-Saavedra and Stange, 2016; Cui et al., 2021). As an important group of plant pigments, anthocyanins are water soluble and belong to the family of flavonoids. So far, more than 500 different anthocyanins have been isolated from plants (Francavilla and Joye, 2020). They are highly involved in determining flower, seed, fruit and vegetative tissue colors, ranging from pink through scarlet, purple, and blue (Tanaka et al., 2008; Khoo et al., 2017). There are six species of anthocyanins (namely cyanidin, delphinidin, peonidin, malvidin, pelargonidin, and petunidin) in colorful plants (Tanaka et al., 2008; Castaneda-Ovañdo et al., 2009), of which cyanidin is responsible for red-purple coloration, delphinidin contributes to purple or blue-red, and pelargonidin contributes to red and orange (Khoo et al., 2017). Besides, anthocyanins also play various vital functions in plant biological functions, including disease protection, resisting environmental stresses, and promoting pollination (Lev-Yadun and Gould, 2009; Zhang et al., 2014).

Anthocyanins are synthesized in cytosol, and stored in the vacuole. Studies on several plant species, including *Arabidopsis* (Baudry et al., 2006; Gonzalez et al., 2008; Qi et al., 2011; Xie et al., 2016), agricultural crops (Yang et al., 2019; Dong et al., 2020), fruits (Rahim et al., 2014; Zhou et al., 2015; Jiang et al., 2019; Li C. et al., 2020), vegetable and ornamental plants (Suzuki et al., 2016; Xu et al., 2017; Jin et al., 2018, 2019; Zhu et al., 2019) have revealed that biosynthesis of anthocyanins are controlled by structural and regulatory genes that take part in formation as well as regulation of specific enzymes. The key enzymes including phenylalanine ammonia lyase (PAL), cinnamic acid 4-hydroxylase (C4H), 4-coumarate-CoA ligase (4CL), chalcone synthase (CHS), chalcone isomerase (CHI), flavonone 3-hydroxylase (F3H), dihydroflavonol 4-reductase (DFR), flavonoid 3′-monooxygenase (F3′H), anthocyanin synthase (ANS), as well as UDP-glucose-flavonoid 3-*O*-glucosyltrasnferase (UFGT) are important in anthocyanin biosynthesis (Koes et al., 2005; Li et al., 2018). Among them, DFR catalyzes the conversion of dihydroflavonols to leucoanthocyanidins, which is one of the final stages of anthocyanin biosynthesis (Shimada et al., 2005; Luo et al., 2016). *DRF* gene is responsible for plant pigmentation (Lou et al., 2014), and its mutation has been associated with the loss of anthocyanins as well as proanthocyanidins (Liu H. et al., 2019; Jiang et al., 2020; Lim et al., 2020; Feng et al., 2021). Besides, enhancement or activation of *DFR* gene expression is vital in MYB transcription factor (TF)-based anthocyanin engineering. For example, regulatory roles of MYB TFs in anthocyanin biosynthesis such as Production of Anthocyanin Pigmentation 1 (*PAP1*, a *MYB75* TF), *PeMYB2/11/12*, *PsMYB114L*, *FtMYBF18*, *EsMYB90,* and *FhMYB5* depend on *DFR* expression (Hsu et al., 2015; Li et al., 2019; Zhang et al., 2019; Dong et al., 2020; Qi et al., 2020). *StMYB44* represses anthocyanin accumulation in leaves of tobacco by directly suppressing the activity of the *DFR* promoter (Liu Y. et al., 2019).

Moreover, some other TFs such as the MYB-bHLH-WD (MBW) complex, B-box, bZIP, MYC, NAC, WRKY, bHLH, MADS-box, and WD could also coordinate anthocyanin biosynthesis initiation by binding to the promoter regions of structural genes (Xu et al., 2015; Zhou et al., 2015; An et al., 2017; Lloyd et al., 2017; Lu et al., 2018; Fang et al., 2019; Jiang et al., 2019). For example, *Arabidopsis* bHLH TFs (GL3, TT8, and EGL3) and WD40 repeat protein TTG1 regulate anthocyanin biosynthetic gene expressions (Gonzalez et al., 2008; Gerats and Strommer, 2009; Saito et al., 2013). Similarly, anthocyanin biosynthesis in petunia petal cells is controlled by the MBW complex, comprising subgroups of MYB TF (PhAN2 or PhAN4) and bHLH TF (PhAN1 or PhJAF13), as well as the WD40 regulator PhAN11 (Quattrocchio et al., 2006). Strawberry FaMADS1a played a negative role in anthocyanin accumulation via repressing expression of *FaPAL6*, *FaCHS*, *FaDFR,* and *FaANS* (Lu et al., 2018). Furthermore, apple B-box zinc finger protein MdBBX20 promotes anthocyanin accumulation in response to ultraviolet-B radiation and low temperature (Fang et al., 2019). DhMYB2 was found to interact with DhbHLH1, thereby regulating anthocyanin secretion in *Dendrobium* hybrid petals (Li et al., 2017). Therefore, the regulatory mechanisms of TFs on plant color are diverse. The formation of plant flower color is affected by both structural genes and TFs.

The *Lycoris* species belongs to Amaryllidaceae family, and is a perennial bulb plant native to Northeast Asia, including China, South Korea, and Japan. It consists of about 20 species, of which 15 species and one variety are distributed in China (Zhang et al., 2020). Among them, *Lycoris radiata* is considered ornamentally and medicinally valuable, as the colorful and special flowers have been used for decoration and the bulbs are notable to produce alkaloids with various biological activities (Park et al., 2019, 2021). Anthocyanins are abundant in *Lycoris* flowers and also contribute to their color formation (He et al., 2011; Chun et al., 2013; Yue et al., 2019; Park et al., 2021). For example, four critical anthocyanins, namely cyanidin 3-sophoroside, cyanidin 3-xylosylglucoside, cyanidin 3-sambubioside, and pelargonidin 3-xylosylglucoside in *L. longituba* tepals of different colors have been well identified (He et al., 2011). In *L. radiata* flowers, three anthocyanins (cyanidin 3-diglucoside, cyanidin 3-sambubioside, and cyanidin 3-glucoside) were identified (Chun et al., 2013), and their presence during four flower development stages was confirmed more recently (Park et al., 2021). However, the molecular mechanisms of anthocyanins regulating color formation of *Lycoris* flower remain unclear. Thus, identifying the key genes related to color formation in *Lycoris* flower would provide a more sufficient genetic resource for manipulation of the related pathways to develop new cultivars with specific flower colors.

In recent years, transcriptome sequencing (RNA-seq) was used as a rapid technique to uncover DEGs, biosynthesis pathways, and TFs related to specific traits in plants (He et al., 2020; Li C. et al., 2020). In this study, we reported the changing profile of anthocyanins and gene expression dynamics in *L. radiata* petals at four developmental stages by integrated analyses of the physiology and transcriptome. We further identified modules with co-expressed genes and candidate hub genes

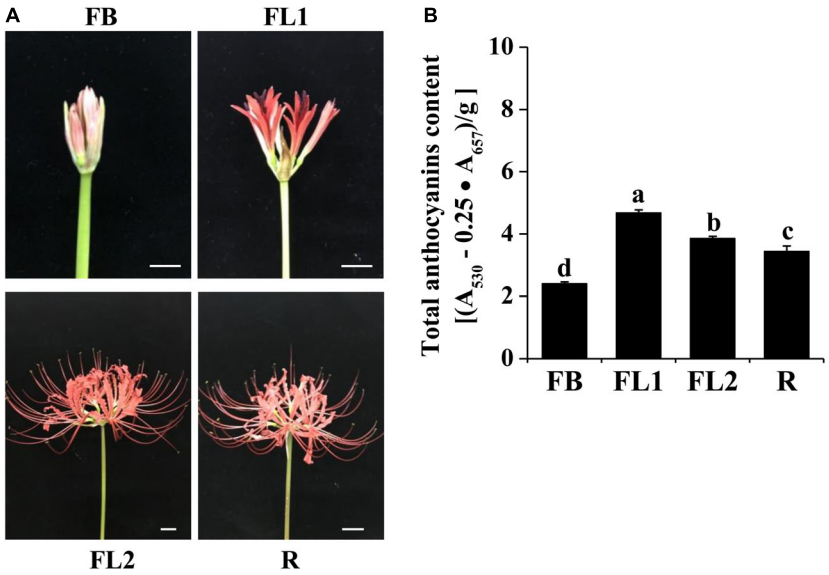


FIGURE 1 | Phenotypes and anthocyanins content in petals of *L. radiata* at different development stages. **(A)** Petals of *L. radiata* at four flower development stages. FB, floral bud stage; FL1, partially opening flower stage; FL2, fully opened flower stage, and R, senescent flower stage. Bars: 1 cm. **(B)** Anthocyanin levels in *L. radiata* petals at four flower development stages. Bars with different letters are significantly different at $p < 0.05$ according to Duncan's multiple range test.

for anthocyanin accumulation, and revealed *LrDFR1* acts as a positive regulator involved in anthocyanin biosynthesis. Our results may serve as a reference for understanding the regulation of key genes and transcription processes in color formation in the flowers of this esthetically important *Lycoris*.

MATERIALS AND METHODS

Plant Materials

Lycoris radiata (L'Her.) Herb. plants were grown in Experimental Plantation of Institute of Botany, Jiangsu Province and Chinese Academy of Sciences, Nanjing, China. According to the studies reported previously (Yue et al., 2019; Park et al., 2021), three biological replicates of *L. radiata* flowers were sampled at four development stages, which were FB (floral bud stage), FL1 (partially opening flower stage), FL2 (fully opened flower stage) and R (senescent flower stage), as shown in **Figure 1A**. Each biological replicate was taken from petals of five flowers and pooled together. For gene expression analysis, different *L. radiata* tissues, including scape, stamen, pistil, flower stalk, and petal samples were obtained during flowering time, while leaf, root, as well as bulb samples from the same plants were obtained during the vigorous vegetative growth stage. The fresh samples were harvested and instantly frozen in liquid nitrogen, then kept at –80°C until use.

Measurement of Total Anthocyanins

Extraction and determination of anthocyanins of *L. radiata* flowers was performed following the protocol of Mehrtens et al. (2005) with minor modifications. Briefly, approximately 0.1 g fresh petals were ground in 1 mL of acidic methanol (0.1 mol L^{-1} HCl) and then incubated overnight in the dark at 4°C with gentle shaking. After centrifugation for 10 min at 12,000 rpm, the supernatant was diluted four times with acidic methanol and the absorbance was measured at 530 and 657 nm using a UV-1600 spectrophotometer (SHIMADZU, Kyoto, Japan). The concentration of anthocyanins was calculated using the following formula: $Q_{Anthocyanins} = (A_{530} - 0.25 \times A_{657}) \times FW^{-1}$, where $Q_{Anthocyanins}$ is the amount of anthocyanins, A_{530} and A_{657} is the absorption at the indicated wavelengths and FW represents the weight of the fresh sample [g].

Construction of the cDNA Library, Sequencing, and Transcriptome Assembly

Total RNA was extracted with the mirVana miRNA isolation kit (Thermo Fisher Scientific, Waltham, MA, United States) following the manufacturer's protocol. The quality and quantity of the RNA were examined by the Agilent 2100 Bioanalyzer (Agilent Technologies, Santa Clara, CA, United States). Samples with RNA Integrity Number (RIN) ≥ 7 were subjected to cDNA library construction using the TruSeq Stranded mRNA LTSample Prep Kit (Illumina, San Diego, CA, United States). Sequencing of the cDNA libraries was done on the Illumina sequencing platform (Illumina HiSeq^{TM} 2500) by Shanghai OE Biotech. Co. Ltd. (Shanghai, China). Reads were cleaned by removing adapters, as well as low-quality and ambiguous regions, then subjected to *de novo* assembly using the Trinity software (Grabherr et al., 2011).

Functional Annotation

Alignment of the assembled unigenes was done against public databases including National Center for Biotechnology Information (NCBI) non-redundant protein (Nr) and nucleotide

(Nt) database, the Swiss-Prot protein database, Gene Ontology (GO) database, Protein Family (Pfam) database, Kyoto Encyclopedia of Genes and Genomes (KEGG) database, Eukaryotic Ortholog Groups (KOG) database, and eggNOG (evolutionary gene genealogy: Non-supervised Orthologous Groups) database.

Identification of Differentially Expressed Genes

The expression level of unigenes was calculated using fragments per kilobase per million fragments mapped (FPKM) method (Mortazavi et al., 2008). Identification of DEGs among samples at four development stage was done using the DESeq2 package implemented in R software, with cutoff values of $|\log2 \text{ (fold change)}| > 1$ and p-value < 0.05 algorithms (Young et al., 2010). To visualize the differential expression profiles, we generated a heatmap for the Trimmed Mean of M-values (TMM) normalized against FPKM via the pheatmap package in R.

Transcription Factors Analysis

To predict TFs involved in color formation of *L. radiata,* we utilized the getorf database (mini-size 150) to find the open reading frame (ORF) (Rice et al., 2000) and then used the HMM search database (version 3.0) to align the ORFs to the TF protein domain. The aligned sequences were described according to the TF families available on the PlantTF database version 3.0 (Zhang et al., 2011). Moreover, the Pearson's correlation coefficient (PCC) between these differentially expressed TFs, structure genes and total anthocyanin content of samples was calculated. The TFs with $|\text{PCC}| > 0.8$ were selected for subsequent analysis. The TF expression data, which included expression levels for MYB, bHLH, WD40, and the DEGs identified in the flavonoid biosynthetic pathway, was screened using blastx software, with an *e*-value of 1e-10. The target gene sequence was aligned to the protein sequence of the reference species contained in the string database, and the protein interaction relationship of the reference species was used to construct an interaction network. Network visualization for the interaction network related to DFR and DEGs was performed using Cytoscape version 3.6.1.

Gene Cloning and Construction of Expression Vectors

Cloning of *LrDFR1* was based on putative ORFs of unigenes from the RNA-seq database. Primers (**Supplementary Table 1**) were synthesized for ORF sequence amplification using Tks Gflex^{TM} DNA Polymerase (Takara, Dalian, China) from *L. radiata* petal cDNA. Reaction conditions were: 5 min of 95°C, 35 cycles for 30 s at 94°C, 30 s at 60°C, 1 min at 72°C, with extension at 72°C for 10 min. PCR products were cloned into pMD19-T simple vectors (Takara, Dalian, China). Afterward, those T-vectors were transferred into DH5α competent cells (Takara, Dalian, China) for amplification. The overexpression vectors of *LrDFR1* were established by linking their ORFs into a linear plant transformation vector, pBinGFP4, using the One Step Cloning Kit (Vazyme, Nanjing, China). Then the *35S:LrDFR1* recombinant vectors were transformed into *Agrobacterium tumefaciens* EHA105 competent cells.

Subcellular Localization and Proanthocyanidin Staining

The pBinGFP4 vector with *LrDFR1-GFP* was transformed into the *Agrobacterium tumefaciens* strain EHA105, and transferred into *Nicotiana benthamiana* epidermal cells (Sheludko et al., 2007). Cultivation of the transformed *N. benthamiana* leaves was done for 2–6 days. For co-localization with membrane-localized marker, *35S: PIP2;1*-mCherry construction was used (Huang et al., 2019). Assessment of transformed *N. benthamiana* epidermal cells was observed with confocal laser scanning microscopy (Zeiss LSM780 META, Jena, Germany). For staining of the nuclei, 10 mg/mL 4′6-diamidino-2-phenylindole (DAPI) was infiltrated into *N. benthamiana* leaves 6 h before observation.

Staining of proanthocyanidin was conducted as described by An et al. (2015). Briefly, light-treated *N. benthamiana* leaves were decolorized in a solution of ethanol: glacial acetic acid (3:1). A dimethylaminocinnamaldehyde (DMACA) reagent staining solution (Sigma-Aldrich, St. Louis, MO, United States) was then added for staining.

Agrobacterium-Mediated Transient Transformation System of *Lycoris* Petals

The *A. tumefaciens* harboring *35S:LrDFR1-GFP* construct and the control pBinGFP4 vector were prepared for injecting into *Lycoris* petals, respectively. The recombinant *Agrobacterium* strains were cultured in YEB broth containing 50 μg mL^{-1} kanamycin and incubated at 28 °C. Then, the collected recombinant *Agrobacterium* strains were resuspended to OD_{600} of 0.6 in a buffer with 10 mM 2-(4-Morpholino) ethanesulfonic acid, 10 mM $MgCl_2$, and 120 μM acetosyringone. Transformed *Lycoris* petals were stored for 48 h in the dark after which they were transferred to a phytotron at a constant photon flux density of 100 μmol m^{-2} s^{-1}. With 5 days cultivation, *Lycoris* petals were obtained for anthocyanin level assessment and RNA extraction.

Validation RNA-Seq by Quantitative Real-Time PCR

For validating gene expression using qRT-PCR, 32 unigenes associated with anthocyanin biosynthesis and phytohormone metabolism were randomly selected (**Supplementary Table 2**). Total RNA isolation was conducted by using the RNAprep Pure Plant Kit (Tiangen, Beijing, China). First-strand cDNA was synthesized with TransScript One-Step gDNA Removal and cDNA Synthesis SuperMix kit (Takara, Dalian, China), and the extracted RNA was used as template according to manufacturer's instructions. A list of gene-specific primers is provided in **Supplementary Table 1**. The quantified expression levels of the tested genes were normalized against the house keeping genes *TIP41-like protein* (*TIP41*) according to previous study on *L. aurea* (Ma et al., 2016). qRT-PCR assays were conducted by the SYBR Premix Ex Taq^{TM} II kit (Tli RNaseH Plus) (Takara, Dalian, China) in a Bio-Rad iQ5 Gradient RT-PCR system. Reaction conditions were: 30 s of denaturation at 95°C and 40

amplification cycles (5 s at 95°C, 30 s at 60°C). Calculation of relative target gene expression levels was done using the $2^{-\Delta\Delta Ct}$ method (Livak and Schmittgen, 2001). Experiments were conducted using three independent biological and three technical replicates.

Statistical Analysis

Statistical analyses were done by SPSS version 10.0 software (IBM Corporation, Armonk, NY, United States). The significant difference among sets of data was determined by one-way analysis of variance (ANOVA) with Duncan's multiple range test ($p < 0.05$) or a significant *t*-test ($^{**}p < 0.01$, $^{*}p < 0.05$). All the results are presented as the mean ± standard deviation (SD).

RESULTS

Anthocyanin Levels in *Lycoris radiata* Petal During Flower Development Stages

During the red flower development of *L. radiata*, petals underwent a rapid color change from slight red to brilliant red (**Figure 1A**). At the flower bud (FB) stage, a slight red color was observed, then the color intensity was significantly increased with rapid elongation of petals in FL1. Subsequently, the intensity of *L. radiata* decreased at FL2 and R stages (**Figure 1A**). We thus investigated the changes of anthocyanin contents in *L. radiata* at four different petal development stages. Notably, anthocyanin content at FL1 stage was significantly higher than that of FB, FL2 and R stages (**Figure 1B**), suggesting that changes in anthocyanin levels could be the main reason for red color formation of *L. radiata*.

Transcriptome Sequencing and *de novo* Assembly

To further study the molecular mechanism of *L. radiata* petal coloring during flower development, twelve libraries were established using samples at four flower development (FB, FL1, FL2, and R) stages (three biological replicates for samples at each development stage), and a total of 644.93 million raw reads as well as 96.73 Gb raw bases were obtained. After eliminating the adaptor, poor-quality sequences, and ambiguous reads, 634.09 million clean reads and 89.86 Gb clean bases were retrieved from 12 samples (**Supplementary Table 3**). The quality score above 30 (Q30) of each library was 93.75–94.91%, and GC percentages ranged from 44.99–46.51% (**Supplementary Table 3**). By using Trinity software, the *de novo* assembly of 12 petal transcriptomes totally generated 87,584 unigenes with an average length of 942 bp (**Supplementary Table 4**). Sequence length distribution showed that 27,073 (30.91%) unigenes had a mean length ≥ 1000 bp (**Supplementary Figure 1** and **Supplementary Table 4**). The N50 was determined to be 1,334 bp, which indicated that the quality of sequence assembly was good.

FPKM values were used to estimate the transcription levels of unigenes. More than 50.0% of unigenes had FPKM values above 1 (**Supplementary Figure 2**). In addition, the use of relative unigene expression obtained from FPKM for principal component analysis (PCA) showed 52.10% variability among the samples (**Supplementary Figure 3**). Moreover, heatmap coefficient matrix analysis of the samples based on the FPKM values showed that most biological replicates (except FB3 sample) exhibited similar expression patterns, indicating relatively high reliability of our sequencing data (**Supplementary Figure 4**).

Functional Annotations and Unigene Classifications

All of the unigenes were annotated by BLAST search against the public databases. The results revealed that 40,974 (46.78%), 29,476 (33.65%), 37,487 (42.8%), and 22,318 (25.48%) unigenes were annotated to the Nr, Swiss-Prot, eggNOG, and Pfam databases, respectively. Taken the entire public databases together, a total of 41,534 (47.42%) unigenes could be successfully annotated (**Supplementary Table 5**). To elucidate their main biological functions, GO, KOG, and KEGG pathway assessments were also performed (**Supplementary Figures 5–7**). Consequently, 27,296 (31.17%) unigenes were assigned into three main categories including "biological process" (BP), "cellular component" (CC), and "molecular function (MF)," which could be further distributed under 50 GO terms (**Supplementary Figure 5**). In addition. 15,122 (17.27%) unigenes were associated with 126 KEGG pathways, and category 'Metabolism' (6187 unigenes) was the most abundant (**Supplementary Figure 6**). Moreover, the KOG analysis showed that 23,858 (27.24%) annotated unigenes were assigned into 25 classes (**Supplementary Figure 7**).

Identification of Differentially Expressed Genes in *Lycoris radiata* Petal During Flower Development Stages

To identify the key DEGs involved in *L. radiata* petal color transitions, six pair-wise comparison groups (FL1 vs. R, FB vs. R, FL2 vs. R, FL2 vs. FL1, FL1 vs. FB, and FL2 vs. FB) were conducted (**Figure 2**). A total of 38,798 DEGs were identified among all samples based on a $|\log_2 \text{fold change}| > 1$ at $p < 0.05$. Among these comparison groups, the largest abundance of DEGs (23,202) was found between FB and R libraries, of which 10,958 and 12,244 genes were down-regulated and up-regulated, respectively (**Figure 2A**). Conversely, the smallest abundance of DEGs (9,057) was recorded between FL2 and FL1 libraries, with 5,033 and 4,024 of them down-regulated and up-regulated, respectively (**Figure 2A**). Furthermore, the overlap DEGs among the six comparison groups were screened. The results indicated that 38 genes were differentially expressed among all the comparisons, which indicated that these DEGs might have key functions in the color expression of different petals (**Figure 2B** and **Supplementary Table 6**).

Functional Annotation of Differentially Expressed Genes

To elaborate the functions of DEGs and identify genes involved in regulating anthocyanin accumulation in *L. radiata*, all the DEGs were firstly subjected to GO analyses, and 14,555 of the 38,798 DEGs were assigned to GO annotations (**Supplementary Table 7**

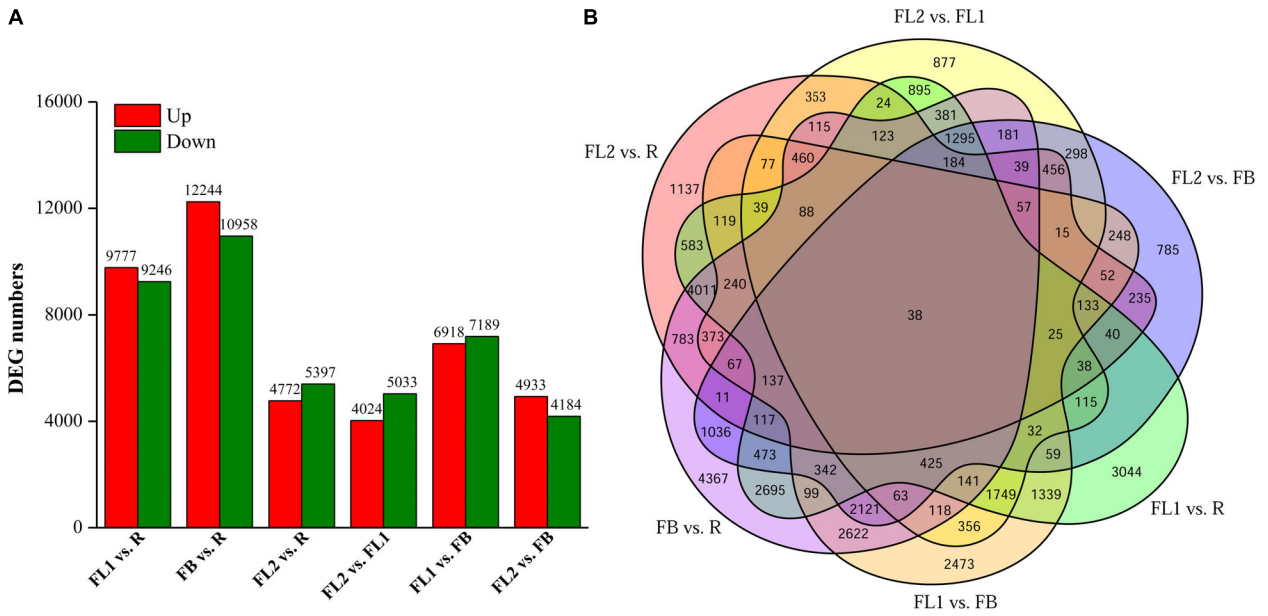


FIGURE 2 | Statistics of differentially expressed genes (DEGs) between two different samples at flower development stages. **(A)** Numbers of DEGs in various pair-wise comparisons. **(B)** Venn diagram for the numbers of DEGs as shown by pair-wise comparisons. FB, floral bud stage; FL1, partially opening flower stage; FL2, fully opened flower stage, and R, senescent flower stage.

and **Figure 3A**). In the biological process category, most of the DEGs were mapped to 'cellular process' (9,476, 20.23%), 'metabolic process' (8,073, 17.23%), and 'response to stimulus' (4,181, 8.92%) terms. In the cellular component category, more than 63.08% of DEGs were enriched in 'cell,' 'cell part' and 'organelle' terms, but for molecular function, nearly 86.13% of DEGs were mapped to 'catalytic activity' and 'binding' terms (**Figure 3A**). For the KEGG annotation results, 7,631 DEGs among all samples were also mapped to 126 KEGG pathways (**Supplementary Table 7**). Comparisons across the samples at four petal development stages revealed significant enrichment of DEGs in 'flavonoid biosynthesis,' 'phenylpropanoid biosynthesis,' 'Tropane, piperidine and pyridine alkaloid biosynthesis,' 'terpenoid backbone biosynthesis' as well as 'plant hormone signal transduction' pathways (**Figure 3B** and **Supplementary Figure 8**). For example, the significantly enriched KEGG pathway term 'Tropane, piperidine and pyridine alkaloid biosynthesis' was shared in all the comparisons. The 'flavonoid biosynthesis' pathway was enriched in FL1 vs. R, FB vs. R, FL2 vs. R, FL2 vs. FL1, and FL2 vs. FB, but not in FL1 vs. FB. In addition, the 'plant hormone signal transduction' pathway was enriched in FB vs. R, FL2 vs. FB, and FL2 vs. FL1 (**Supplementary Figure 8**).

Identification of Key Differentially Expressed Genes Responsible for the Anthocyanin Biosynthesis Pathway

To elucidate the molecular basis underlying difference in anthocyanin biosynthesis among the four flower development stages in *L. radiata*, DEGs involved in the anthocyanin synthesis pathway were identified. The results revealed that 56 DEGs were enriched in the anthocyanin synthesis pathway, including *PAL*, *C4H*, *4CL*, *CHS*, *CHI*, *F3H*, *F3′H*, *DFR*, *ANS*, *UFGT*, *FLS*, *ANR*, and *LAR* (**Figure 4A**). Moreover, the Pearson's correlation coefficient between the expression level of these DEGs and the total anthocyanins content was further calculated (**Figure 4B**). The results showed that 23 DEGs negatively regulated anthocyanin synthesis, whereas 33 DEGs positively regulated the anthocyanin synthesis. Among them, the expression level of two DEGs, namely *LrDFR1* (DN43960) and *LrDFR2* (DN42380) indicated a significant positive correlation with the total anthocyanins content in petals during the flower development stages, while *LrFLS* (DN37334) indicated a significant negative correlation with the total anthocyanins content ($|PCC| > 0.8$, **Figure 4B** and **Table 1**), suggesting that these three DEGs may have an essential role in anthocyanin accumulation.

Identification of Transcription Factors Related to Anthocyanin Biosynthesis in the Petals of *Lycoris radiata*

Transcription factors were subsequently predicted to whether modulate anthocyanin accumulation and biosynthesis in *L. radiata* petals during flowering development stages. In this study, a total of 1,631 TFs were identified by searching the TF database. The classified results indicated that most of these TFs belonged to the MYB, C2C2, AP2/ERF, C2H2, and bHLH family (**Supplementary Figure 9**). Furthermore, the differentially expressed TFs (721) were characterized by analyzing their FPKM values (**Supplementary Figure 9**). Importantly, co-expression modules of these 721 TF DEGs were analyzed with

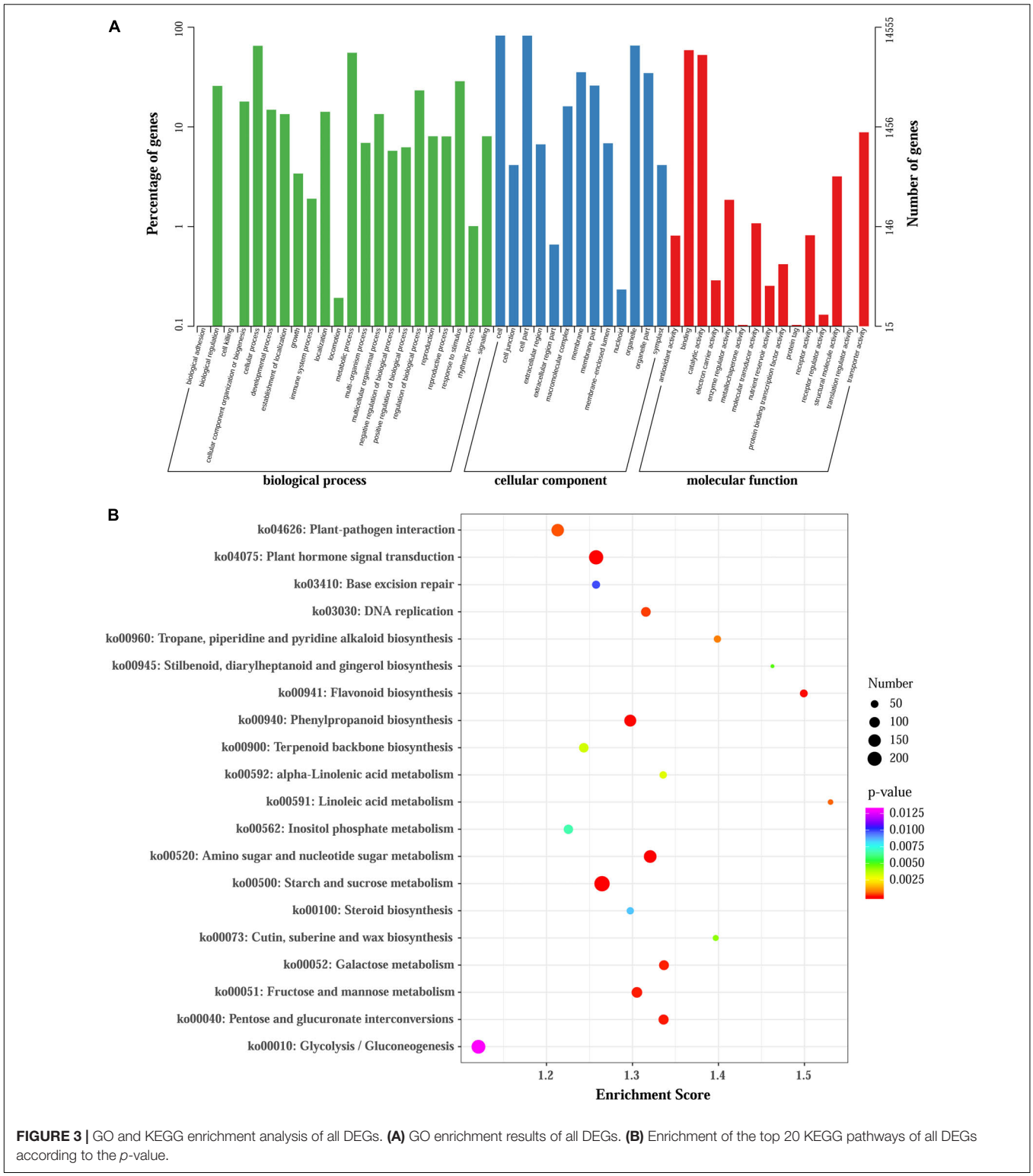


FIGURE 3 | GO and KEGG enrichment analysis of all DEGs. **(A)** GO enrichment results of all DEGs. **(B)** Enrichment of the top 20 KEGG pathways of all DEGs according to the *p*-value.

Short Time-series Expression Miner (STEM) software. In all, six clusters of 272 TF DEGs were detected (**Figure 5A** and **Supplementary Table 8**).

On the other hand, by calculating the PCC between the expression level of 721 TF DEGs and the total anthocyanins content, 27 TFs genes (|PCC| > 0.8) involved in the accumulation of anthocyanins were identified, including 10 positive regulators and 17 negative regulators (**Table 1**). These 10 positive regulators, including *MYB* (1), *AP2/ERF* (1), *bHLH* (1), *bZIP* (1), *NAC* (1), *NF-X1* (1), and *Trihelixs* (4) genes,

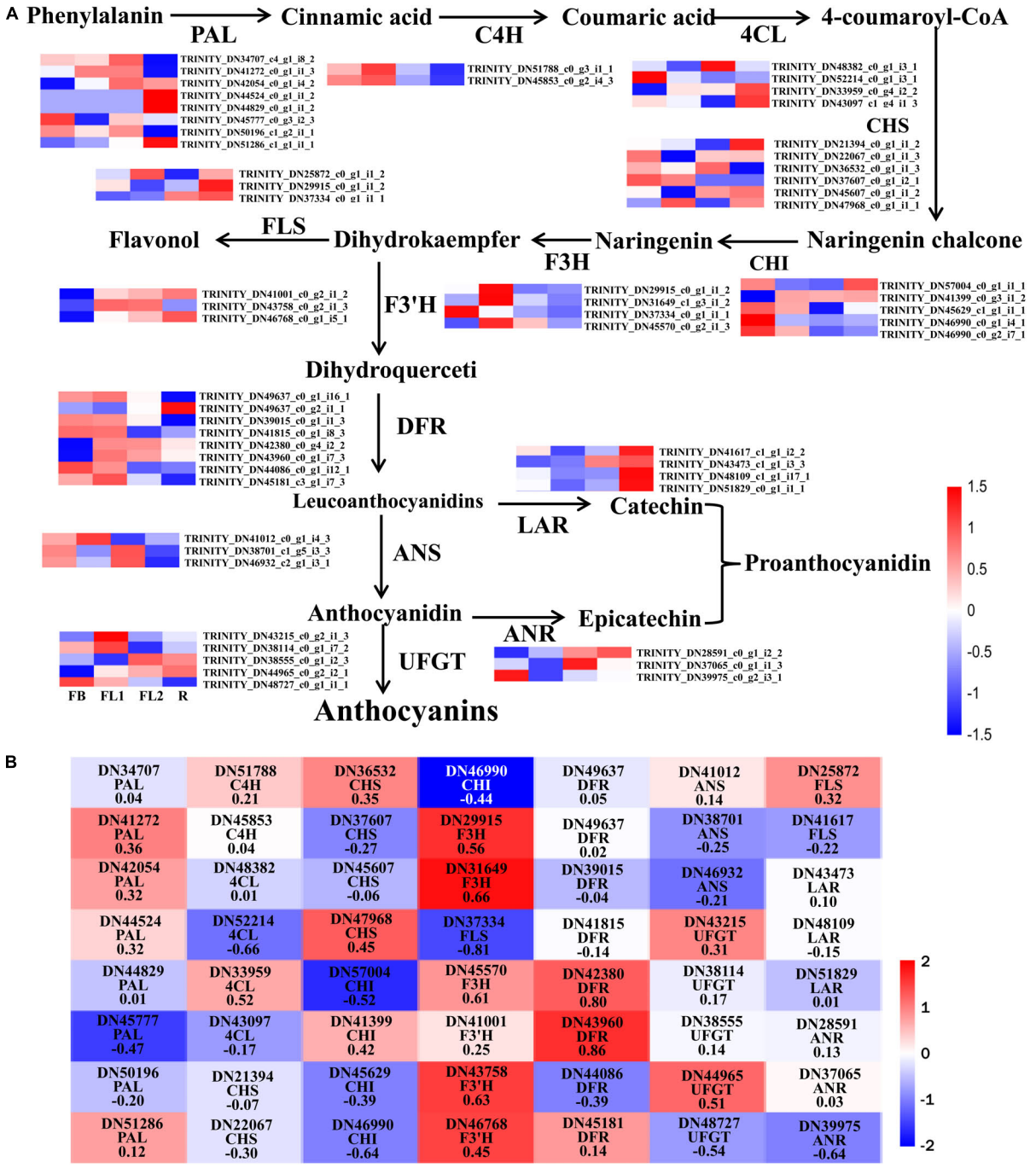


FIGURE 4 | Analysis of DEGs involved in anthocyanin biosynthesis pathway in *L. radiata*. **(A)** Anthocyanin biosynthesis pathway and the $\log_2$ transformed FPKM values of DEGs associated with structural enzyme genes were used to draw the heatmap. The enzymes include 4-coumarateCoA ligase (4CL), phenylalanine ammonia lyase (PAL), chalcone synthase (CHS), flavone 3-hydroxylase (F3H), chalcone isomerase (CHI), flavonoid 3′-hydroxylase (F3′H), dihydroflavonol reductase (DFR), flavonol synthase (FLS), UDP-flavonoid glucosyl transferase (UFGT), anthocyanidin reductase (ANR), and leucoanthocyanidin reductase (LAR). FB, floral bud stage; FL1, partially opening flower stage; FL2, fully opened flower stage, and R, senescent flower stage. Color gradients comprise red, white, and blue, representing genes that were upregulated, not regulated, as well as downregulated, respectively. **(B)** The heatmap analysis of all DEGs in anthocyanin biosynthesis pathway according to the FPKM value.

likely act to improve anthocyanin synthesis during *L. radiata* petal development stages. However, the 17 negative regulators, including *Alfin*-like (1), *AP2/ERF* (1), *GATA* (1), *GRFs* (2), *bHLHs* (2), *MYBs* (2), *C2H2s* (2), *C3Hs* (4), *MADS* (1), and *NAC* (1), might act as repressors in *L. radiata* anthocyanin biosynthesis (**Table 1**). Notably, 11 of 17 negative TF regulators (subclass 4,

TABLE 1 | The candidate TFs and the key structural gene involved in anthocyanin accumulation.

Gene family	Gene ID	Annotation	Correlation with total anthocyanin	*p*-value
FLS	TRINITY_DN37334_c0_g1_i1_1	Flavonol synthase/flavanone 3-hydroxylase-like	−0.8058	0.0015
DFR	TRINITY_DN42380_c0_g4_i2_2	Dihydroflavonol 4-reductase LrDFR2	0.8045	0.0016
	TRINITY_DN43960_c0_g1_i7_3	Dihydroflavonol 4-reductase (LrDFR1)	0.8655	0.0002
Alfin-like	TRINITY_DN45802_c0_g1_i2_1	PHD finger protein ALFIN-LIKE 6-like	−0.9302	1.16E-05
AP2/ERF	TRINITY_DN42881_c0_g1_i1_3	AP2 domain-containing transcription factor 2	0.8584	0.0003
	TRINITY_DN13573_c0_g1_i1_1	AP2 domain-containing transcription factor 2	−0.8304	0.0008
bHLH	TRINITY_DN36174_c0_g1_i1_1	Transcription factor bHLH30-like	−0.8631	0.0002
	TRINITY_DN41224_c0_g1_i1_1	Transcription factor bHLH57-like	−0.8494	0.0004
	TRINITY_DN48856_c0_g1_i6_1	Transcription factor bHLH48-like	0.8305	0.0008
bZIP	TRINITY_DN27549_c0_g1_i3_2	bZIP transcription factor 11-like	0.8317	0.0007
GATA	TRINITY_DN30983_c0_g2_i1_2	GATA transcription factor 3-like isoform X2	−0.8175	0.0011
GRF	TRINITY_DN37111_c0_g1_i4_2	Growth-regulating factor 4-like	−0.8349	0.0007
	TRINITY_DN48073_c0_g1_i1_1	Growth-regulating factor 7-like	−0.8084	0.0014
C2H2	TRINITY_DN44933_c0_g1_i2_1	Histone deacetylase HDT2-like	−0.8566	0.0003
	TRINITY_DN48967_c0_g1_i23_1	Histone deacetylase HDT2	−0.8434	0.0005
C3H	TRINITY_DN36142_c1_g2_i3_3	Zinc finger CCCH domain-containing protein 8	−0.8517	0.0004
	TRINITY_DN42291_c0_g1_i1_1	Zinc finger CCCH domain-containing protein 59	−0.8752	0.0001
	TRINITY_DN60484_c0_g1_i1_1	Zinc finger CCCH domain-containing protein 44	−0.8130	0.0013
	TRINITY_DN48680_c0_g1_i6_1	Zinc finger CCCH domain-containing protein 8	−0.8535	0.0004
MADS	TRINITY_DN50153_c1_g1_i10_1	Transcription factor, MADS-box	−0.8239	0.0009
MYB	TRINITY_DN33872_c0_g1_i1_3	MYB transcription factor	0.8101	0.0013
	TRINITY_DN45447_c0_g1_i7_1	MYB transcription factor	−0.8298	0.0008
	TRINITY_DN51496_c0_g1_i9_1	MYB transcription factor	−0.8245	0.0009
NAC	TRINITY_DN39353_c0_g2_i1_1	NAC domain-containing protein 43	−0.8351	0.0007
	TRINITY_DN39797_c0_g2_i6_3	NAC domain-containing protein 17-like	0.8215	0.0010
NF-X1	TRINITY_DN33064_c0_g1_i1_2	NF-X1-type zinc finger protein NFXL2	0.8369	0.0006
Trihelix	TRINITY_DN39907_c1_g1_i5_1	Trihelix transcription factor GTL1	0.8138	0.0012
	TRINITY_DN41155_c0_g2_i2_2	Trihelix transcription factor GTL1-like	0.8208	0.0010
	TRINITY_DN43481_c0_g2_i1_1	Trihelix transcription factor ASIL2-like	0.8571	0.0003
	TRINITY_DN45677_c0_g1_i7_3	Trihelix transcription factor GTL1	0.8042	0.0016

5, and 6) and one positive regulator (subclass 2) were enriched in TF co-expression modules (**Supplementary Table 8**).

Previous studies have reported that bHLH, MYB and WD40 TFs regulate anthocyanin biosynthesis thereby activating or repressing transcription of anthocyanin structural genes. We then performed unigenes regarding to MYB, bHLH and WD40, as well as 56 DEGs involved in anthocyanin biosynthesis (**Figure 5**) to analyze their interaction network and hope to identify the hub TF genes that could affect anthocyanin biosynthesis pathway. The results showed that four *DFRs*, four *MYBs*, two *WD40s*, two *4CLs*, one *F3′H*, one *UFGT*, one *CHS*, one *ANS*, one *FLS*, and one *CHI* were selected as hub genes based on their connection position in the network modules, expression pattern and functional annotation (**Supplementary Table 9a** and **Supplementary Figure 10**). Furthermore, those genes (shown in **Supplementary Table 9a**) and 27 key TF genes (**Table 1**) were selected to build the interaction network for further analysis. Among them, *LrDFR1* (DN43960) and *LrFLS* (DN37334) could be regarded as key hub genes for participating anthocyanin biosynthesis. Two *MYBs* (DN45447 and DN33872), two *NACs* (DN39353 and DN39797), one *C3H* (DN42291), and one *GATA* (DN30983) TF genes were identified as hub genes in regulating anthocyanin biosynthesis (**Figure 5B** and **Supplementary Table 9b**). The above results indicate that these eight genes may play essential roles in anthocyanin synthesis in *L. radiata* during petal development.

Validation of RNA-seq Data by qRT-PCR

To validate the accuracy and transcription profiles revealed by the RNA-seq data, 32 unigenes were selected for qRT-PCR assays. The relative expression levels of these 32 genes were normalized to the expression of *LrTIP41*, and compared with the RNA-Seq data, as shown in **Figure 6A**. Further linear regression analysis revealed that the expression levels of these genes were well correlated with the RNA-Seq results (**Figure 6B**, $R^2 > 0.76$), indicating that the RNA-seq data were credible and accurate.

LrDFR1 Is Involved in Anthocyanin Biosynthesis in *Lycoris radiata*

In this study, we cloned *LrDFR1* gene (DN43960) from *L. radiata*. The full-length cDNA of *LrDFR1* is 1113 bp in length and it encodes a 370 amino acid protein with a molecular weight of 41.67 kDa (**Supplementary Table 10**).

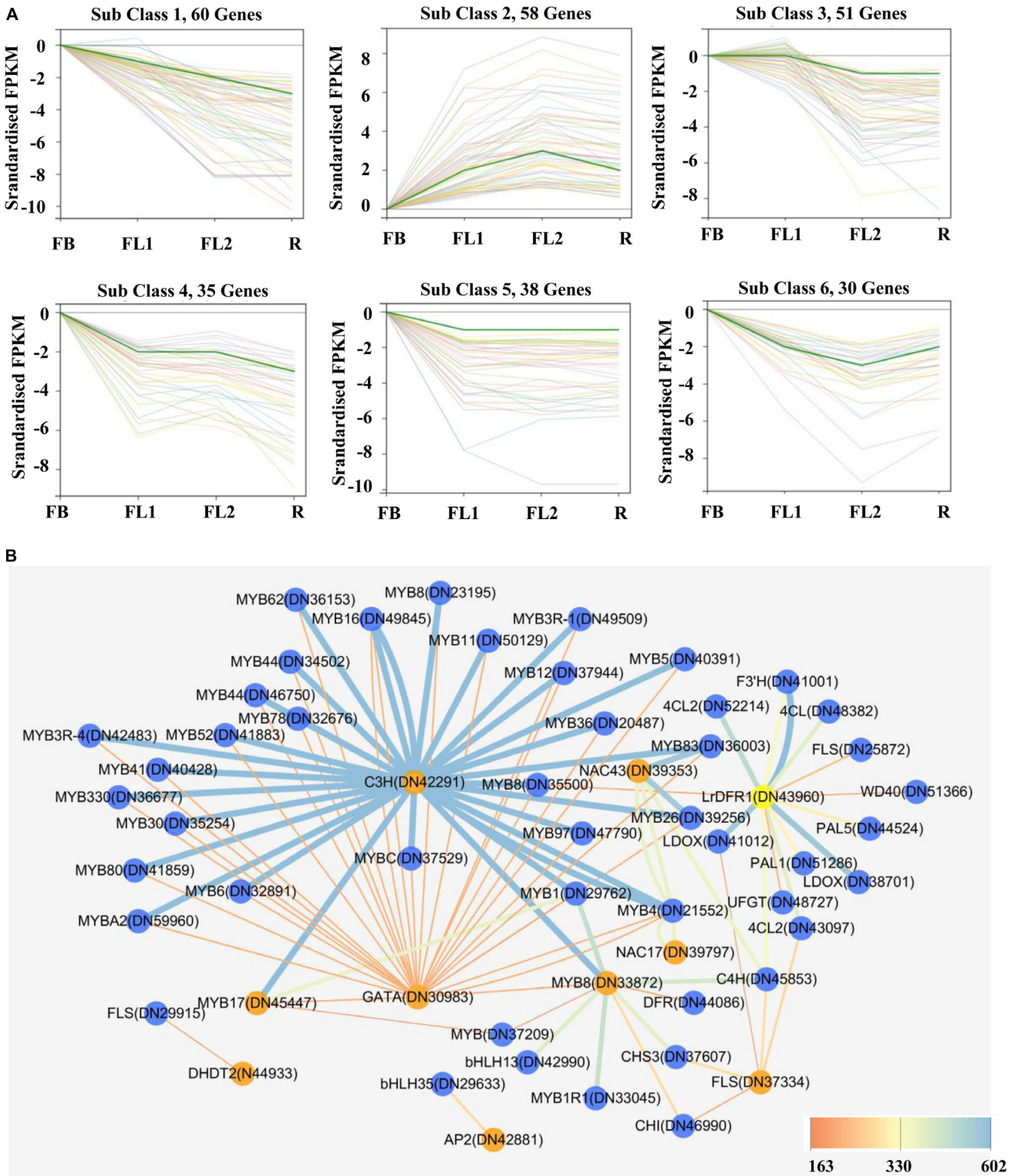


FIGURE 5 | Gene expression profiles of identified transcription factor and protein-protein interaction network of key structural enzymes and TFs involved in anthocyanin biosynthesis in *L. radiata* flowers. **(A)** *K*-means clusters of DEGs based on standardized (log_2 transformed) FPKM of *L. radiata* petals at flower development stage (FB, FL1, FL2, and R). Number of genes that were clustered in every subclass are shown above each Figure. **(B)** Protein–protein interaction network constituted by protein sequences of differentially expressed transcription factors and structural genes involved in anthocyanin synthesis of *L. radiata* petals. Genes that have the higher weight are depicted in 'yellow and orange,' the 'blue edges' correspond to co-expressed strong links and the 'yellow edges' correspond to co-expressed weak links.

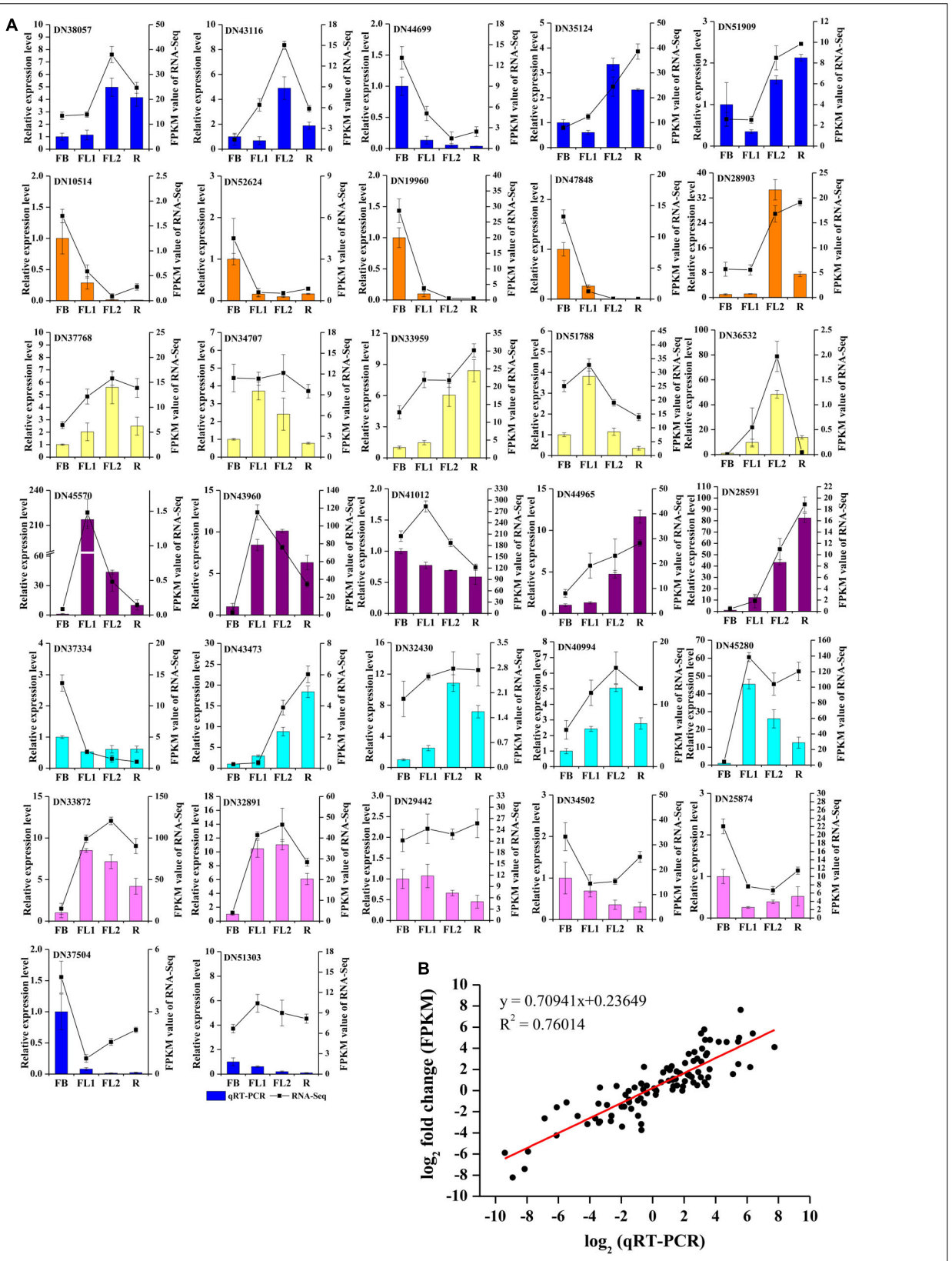


FIGURE 6 | qRT-PCR validation of gene expression level in the transcriptome. **(A)** qRT-PCR validation of gene expression level in the transcriptome. Thirty two unigenes were selected for qRT-PCR validation. **(B)** Correlation analysis of the results between qRT-PCR and RNA-Seq.

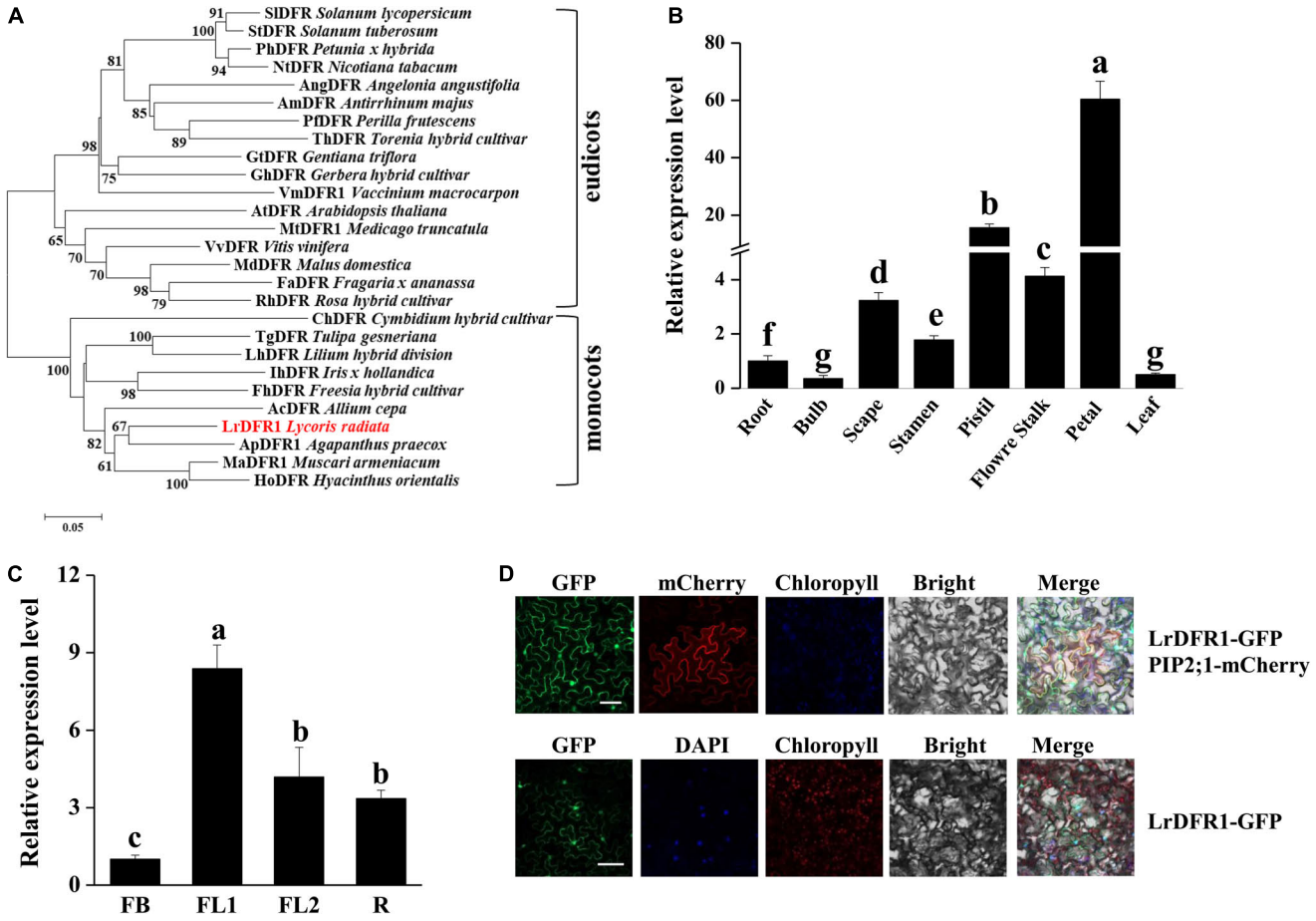


FIGURE 7 | Phylogenetic tree analysis, transcription expression profiles, and subcellular localization of LrDFR1. **(A)** Phylogenetic assessment of LrDFR1 with other DFR proteins from different plants. Phylogenetic tree generation was achieved using the maximum likelihood method in MEGA 7.0 software. Numbers at every interior branch show bootstrap values of 1000 replicates. The bar shows a 0.05 genetic distance. Plant species as well as GenBank accession numbers of DFR proteins used in phylogenetic analyses are: *Solanum tuberosum* StDFR (AF449422), *Solanum lycopersicum* SlDFR (CAA79154.1), *Petunia hybrida* PhDFR (AF233639), *Angelonia angustifolia* AngDFR (KJ817183), *Nicotiana tabacum* NtDFR (NP_001312559.1), *Antirrhinum majus* AmDFR (X15536), *Perilla frutescens* PfDFR (AB002817), *Gentiana triflora* GtDFR (D85185), *Torenia hybrid* ThDFR (AB012924), *Gerbera hybrid* GhDFR (Z17221), *Vaccinium macrocarpon* VmDFR1 (AF483835), *Arabidopsis thaliana* AtDFR (AB033294), *Medicago truncatula* MtDFR1 (AY389346), *Vitis vinifera* VvDFR (Y11749), *Malus domestica* MdDFR (AAO39816), *Rosa hybrid* RhDFR (D85102), *Cymbidium hybrid* ChDFR (AF017451), *Fragaria ananassa* FaDFR (AF029685), *Tulipa gesneriana* TgDFR (BAH98155.1), *Lilium hybrid* LhDFR (AB058641), *Iris hollandica* IhDFR (BAF93856.1), *Allium cepa* AcDFR (AY221250.2), *Agapanthus praecox* ApDFR (AB099529.1), *Muscari aucheri* MaDFR (MH636605), *Freesia hybrid* FhDFR (KU132389), and *Hyacinthus orientalis* HoDFR (AFP58815.1). **(B)** Expression profiles of *LrDFR1* in various tissues of *L. radiata*. Expressions of *LrDFR1* were assessed by qRT-PCR, and normalized to *LrTIP41*. Expressions of *LrDFR1* in root tissues were defined as 1.0. Data are shown as mean ± SD. Bars with different letters are significantly different at $p < 0.05$ according to Duncan's multiple range test. **(C)** Expression profiles of *LrDFR1* during the FB stage, FL1 stage, FL2 stage and R stage of *L. radiata*. Expression levels of *LrDFR1* were assessed by qRT-PCR, and normalized to *LrTIP41*. Expression levels of *LrDFR1* in FB stage were defined as 1.0. Data are shown as mean ± SD. Bars with different letters are significantly different at $p < 0.05$ according to Duncan's multiple range test. **(D)** Subcellular localization of LrDFR1 in *N. benthamiana* epidermal cells. Scale bars = 20 μm. The nuclei are indicated by DAPI staining.

The deduced amino acid sequence of LrDFR1 revealed a high similarity with DFR proteins from *Agapanthus praecox* (75.33%), *Muscari armeniacum* (74.74%), and *Hyacinthus orientalis* (72.72%) (**Figure 7A**). Multiple amino acid sequence alignments showed the highly preserved NADPH-binding motif (VTGAAGFIGSWLIMRLLERGY) (Gang, 2005) and the substrate-binding domain (T128–K154) (Johnson et al., 2001) in the LrDFR1 sequence (**Supplementary Figure 11**). qRT-PCR was then performed to assess whether expression patterns of *LrDFR1* in different tissues and flower development stages were coincided with anthocyanin accumulation in *L. radiata*. *LrDFR1* was found to be expressed in all tissues, with the highest expression levels in petals (**Figure 7B**). Moreover, expression levels of *LrDFR1* were significantly increased from stage FB to stage R, peaking at stage FL1 (**Figure 7C**). These findings imply tissue-specific expression levels for *LrDFR1*, which is associated with anthocyanin accumulation in *L. radiata* petals.

Moreover, we transiently expressed *LrDFR1* in tobacco epidermal cells to assess subcellular localization of *LrDFR1*. As shown in **Figure 7D**, the fluorescent signal of LrDFR1-GFP was localized into the nucleus, cytoplasm and cell membrane, while GFP was evenly distributed in the cell (**Figure 7D**

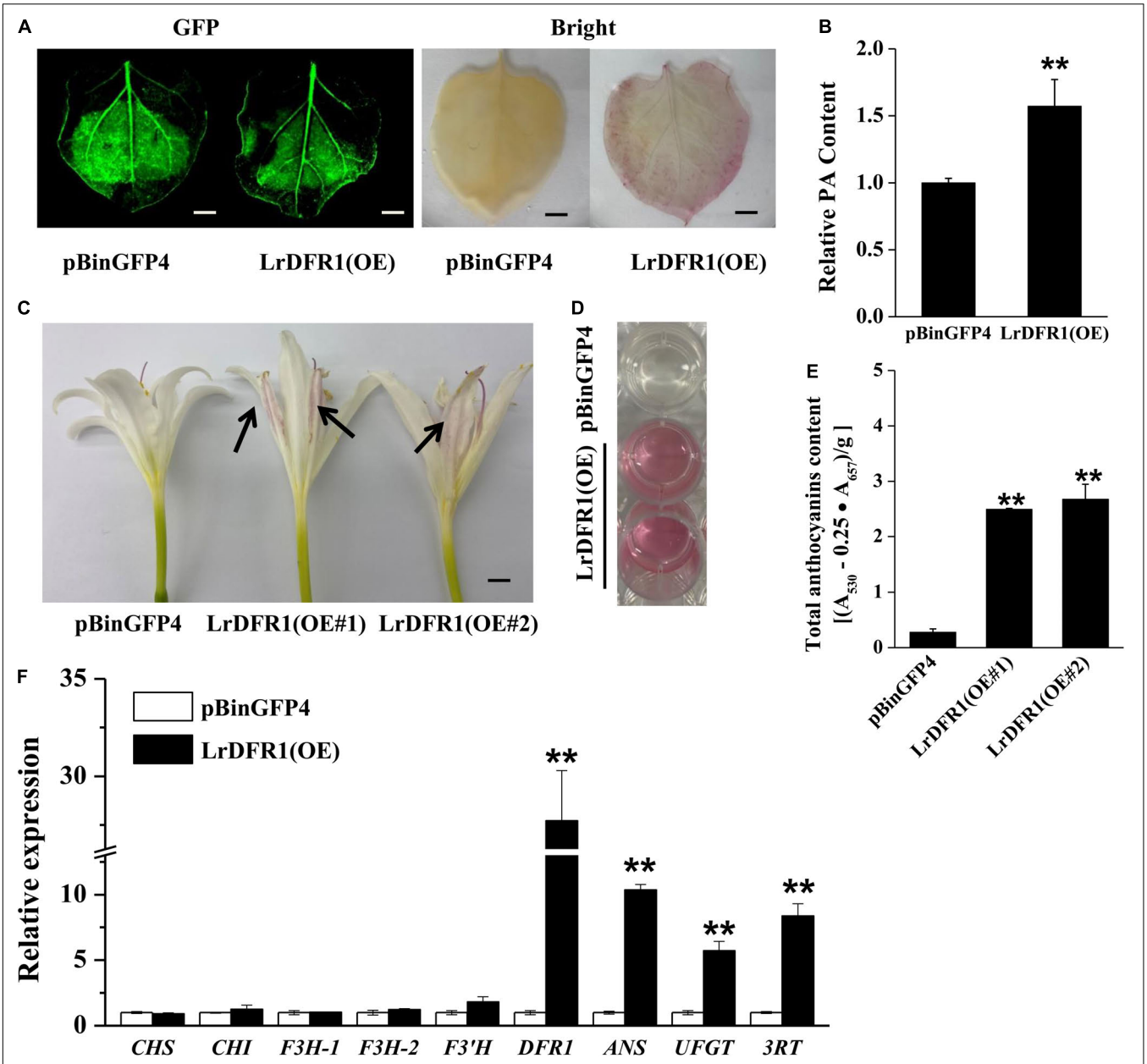


FIGURE 8 | Overexpression of *LrDFR1* promotes anthocyanin and proanthocyanidin biosynthesis. **(A)** Proanthocyanidin staining and **(B)** relative proanthocyanidin (PA) levels in transiently transformed tobacco leaves (pBinGFP4: empty vector controls; LrDFR1-OE: *LrDFR1*-overexpressing leaves). Tobacco leaves were kept in a phytotron at 24°C under constant lighting for 5 days. DMACA was used to stain proanthocyanidin. Every experiment was performed using 8–10 leaves for each genotype. Experiments were conducted in triplicates, and a representative image is shown. Proanthocyanidin levels of empty vector controls were set as the reference to 1. Asterisks represent significant differences between control and *LrDFR1*-overexpressing leaves (**$p < 0.01$). Bars = 1 cm. **(C)** Phenotypes of anthocyanin accumulation. Arrow indicates the transfected petals. **(D,E)** Relative anthocyanin levels in transiently transformed *Lycoris* petals (pBinGFP4: empty vector controls; LrDFR1-OE: *LrDFR1*-overexpressing petals). *Lycoris* petals were kept in a phytotron at 24°C with a constant light for 5 days. Every experiment was performed using 8–10 petals per genotype. Data are shown as mean ± SD. **$p < 0.01$. Bar = 0.5 cm. **(F)** Relative expression levels of endogenous anthocyanin biosynthetic genes in pBinGFP4 (empty vector controls) as well as *LrDFR1*-overexpressing petals. Expression patterns of early biosynthetic genes (*CHS*, *F3H*, *CHI*, and *F3′H*) as well as late biosynthetic genes (*DFR*, *UFGT, ANS*, and 3RT) in petals were investigated. Asterisks represent significant differences between control and *LrDFR1*-overexpressing petals (**$p < 0.01$).

and **Supplementary Figure 12**). To determine the roles of *LrDFR1* in regulating anthocyanin as well as proanthocyanidin biosynthesis in *L. radiata*, an *LrDFR1*-overexpressing plasmid was transfected into *Lycoris* petals and tobacco epidermal cells (**Figure 8A**). Overexpression of *LrDFR1* in tobacco and *Lycoris* petals markedly enhanced proanthocyanidin and anthocyanin accumulation (**Figures 8B–E**). To assess the effects of *LrDFR1* on endogenous *Lycoris* petals genes that are involved in anthocyanin

synthesis, the expression levels of *CHS*, *CHI*, *F3H*, *F3′H*, *DFR*, *ANS*, *UFGT,* and *3RT* were determined (**Figure 8F**). Among them, the expressions of *LrDFR1*, *ANS*, *UFGT* and *3RT* were significantly higher in *LrDFR1*-overexpressing plants than in control plants (**Figure 8F**). These results suggest that *LrDFR1* may play important roles in anthocyanins biosynthesis of *Lycoris* petals.

DISCUSSION

Changes of Anthocyanin Contents in the *Lycoris radiata* Petals During Flower Development Stages

Flowering plants exhibit a wide variation in their flora, foliage, and fruit colors, as a result of genetic factors and variations in environments. Flavonoids/anthocyanins, betalains and carotenoids are the major metabolites for coloration in plant reproductive organs (Griesbach, 2005; Tanaka et al., 2008). Most of the red, purple, and blue-colored flowers (such as red rose, lavender, and blue chicory) as well as fruits (such as berries, currants, and grapes) contained high anthocyanins content (Khoo et al., 2017). The genus *Lycoris* is used as a garden flower due to the colorful and special flowers, and the flower colors of *Lycoris* are diverse. For example, the flower color of *L. radiata* and *L. rosea* was red, that of *L. aurea* and *L. chinensis* was yellow. *L. sprengeri* and *L. haywardii* showed red and blue color, while *L. longituba* displays an exceptionally wide diversity of flower colors from purple, red, orange, to yellow (He et al., 2011). Similar to the flowers of other species, the petals of *Lycoris* are rich in anthocyanins, and their color formation are largely related to anthocyanins (He et al., 2011; Chun et al., 2013; Yue et al., 2019; Park et al., 2021). In this study, we determined the content of anthocyanins in the petals of *L. radiata*, and the results showed that the color intensity of the *L. radiata* petals was changed with the different anthocyanin contents. The anthocyanins increased then decreased during the flower development stages (**Figure 1**), which are similar to the results recently reported by Park et al. (2021).

Key Structural Genes Responsible for Anthocyanin Synthesis in *Lycoris radiata* Petals During Flower Development Stages

To data, transcriptome sequencing is highly employed for predicting novel genes, gene function, and genome evolution for plant breeding and horticulture research (Rameneni et al., 2020). For example, transcriptome analysis has revealed the role of anthocyanin in flower color formation in several horticultural crops, such as *Camellia sinensis* (Zhou et al., 2020), "Tiny Padhye" (*Lilium* spp.) (Xu et al., 2017), lilies (*Lilium* spp.) (Suzuki et al., 2016), *Magnolia sprengeri* (Shi et al., 2014), *Paeonia lactiflora* (Zhao et al., 2014), *Paeonia delavayi* (Shi et al., 2015), and *Silene littorea* (Casimiro-Soriguer et al., 2016). For better understanding of petals color formation during flower development stages in *L. radiata*, a comparative transcriptomics analysis was carried out. The results showed that approximately 70.27 GB of high-quality data, and 87,584 unigenes were obtained. Further analyses, based on NR, Swiss-Prot, KEGG, KOG, GO, Pfam, and eggNOG databases, predicted 38,798 DEGs associated with a specific or general function (**Supplementary Tables 4, 5**).

The variations in floral coloration emanates from different processes, such as pathways competition, expression levels of structural genes involved in pigment formation, and mutations of structural or regulatory genes (Grotewold, 2006; Cui et al., 2021). In plants, phenylpropanoids represent a vital group of physiologically active secondary metabolites derived from phenylalanine, and anthocyanins, flavonols, isoflaconoids and flavonols have a similar metabolism pathway during their biosynthesis (Ferrer et al., 2008). KEGG pathway analysis showed that the 'phenylpropanoid biosynthesis,' 'flavonoid biosynthesis,' as well as 'flavone and flavonol biosynthesis' pathways were enriched between each two transcriptomes of *L. radiata* petals during flower development stages (**Figure 3B** and **Supplementary Figure 8**). Given that anthocyanin biosynthesis pathway is well known to modulate color formation in plants, we mainly focused on them as the candidate pathways to elucidate their involvement in petal/flower color formation in *L. radiata*. Subsequently, we identified the main functional genes participated in the anthocyanin biosynthetic pathway, and found that most of structural genes such as *F3′H*, *UFGT*, *DFR,* and *FLS* were elevated in *L. radiata* petals at FL1 and FL2 stages (**Figure 4A**). Therefore, these genes might have contributed to the increasing anthocyanin content in petals from the FB stage to the FL1 and FL2 stage, as evidenced in **Figure 1**. For example, three *F3′H* genes (*DN41001*, *DN43758,* and *DN46768*) were highly expressed in petals at FL1 and FL2 stages (**Figure 4A**). Another prominent gene, *UFGT* (*DN44965*), which glycolyzes anthocyanidin into anthocyanin (Xie et al., 2003), was also highly expressed in petals at FL1, FL2 and R stages, as compared to that of the samples at FB stage (**Figure 4A**). All of these genes were positively correlated with the biosynthesis of anthocyanins (Niu et al., 2010). Notably, two *DFR* genes (DN42380, DN43960) and one *FLS* (DN37334) (**Table 1**) were found to be highly associated with the total anthocyanins content ($|PCC| > 0.8$), suggesting they may have an essential function in the phenotypic expression of petal color (**Figure 4B**). In anthocyanin biosynthesis, DFR catalyze the reduction of dihydroquercetin to leucoanthocyanidins, and the level of *DFR* expression have been associated with flower color changes (Nakatsuka et al., 2003; Zhao et al., 2012). qRT-PCR also indicated that the hub gene *LrDFR1* were mostly expressed the most in the FL1 samples (**Figure 7C**). Our results suggest that these enzymes may be the most important enzymes to catalyze anthocyanin biosynthesis in *L. radiata* petals.

Transcriptional Regulation of Color Formation in *Lycoris radiata* Petals

Transcription factors play critical functions in flavonoid biosynthesis, by regulating expression of structural genes. For example, the class of TFs identified were previously implicated in regulation of petal color formation in roses

(Li D. et al., 2020). Particularly, MYB-bHLH-WD40 complexes have been implicated in multi-level regulation of flavonoid biosynthesis (Gallego et al., 2018), whereas the R2R3-MYB family was shown to play a vital role in regulation of spatiotemporal expressions of genes involved in anthocyanin biosynthetic in plants (Gonzalez et al., 2008; Zhao and Tao, 2015). Besides, MYB-domain TFs are important mediators of anthocyanin accumulation and participate in colorations of various organs in horticultural as well as ornamental plants (Tang et al., 2017; Liu Y. et al., 2019; Xiang et al., 2019; Zhai et al., 2019; He et al., 2020; Sun et al., 2020; Wang et al., 2020; Zhong et al., 2020).

In this study, the most abundant TFs including *AP2*/*ERF*, *bHLH*, *bZIP*, *C2C2*, *HSF*, *MYB*, *NAC*, *TIFY*, and *WRKY* families were predicted (**Supplementary Table 8**). In addition, we employed a *K*-means clustering, as proposed earlier by Handhayani and Hiryanto (2015), which permitted the clustering of 272 TF unique genes among the samples (FB, FL1, FL2, and R) into six sub-clusters with some members in Cluster 2 associated with genes from the *MYB* and *bHLH* TFs (**Figure 5A**). Based on the expression level of TFs obtained from the transcriptome data, 27 TFs (**Table 1**) were found to highly associate with the total anthocyanin content (|PCC| $>$ 0.8), and these TFs may have an essential function in the phenotypic expression of *L. radiata* petal color. Interestingly, among these TFs, three *MYBs* showed two different expression patterns. The expression level of two *MYBs* (DN45447, DN51496) was highest in FB, followed by FL1, FL2, and R, which was contrary to the total anthocyanin content trend. Conversely, the expression of *LrMYB1* (DN33872) exhibited a similar trend to the total anthocyanin content in the *L. radiata* petals (**Supplementary Figure 13**), indicating that *MYBs* (DN45447 and DN51496) negatively regulated anthocyanin accumulation, whereas *LrMYB1* (DN33872) was identified as one of the eight hub genes may positively regulate anthocyanin accumulation in *L. radiata* (**Table 1** and **Figure 5B**).

Subsequently, two negatively correlated *bHLHs* (DN36174 and DN41224) and one positively correlated *LrbHLH1* (DN48856) were identified (**Table 1**). In plants, MYB often forms protein complexes with bHLH and WD40 to participate in anthocyanin biosynthesis rather than regulate anthocyanin biosynthesis directly (Feng et al., 2020). In apple, *MdMYB1*, *MdMYB9*, *MdMYB10,* and *MdMYBA* act as positive modulators of anthocyanin biosynthesis, by activating the expressions of *MdDFR* and *MdUF3GT* (Takos et al., 2006; Ban et al., 2007; Espley et al., 2007; An et al., 2015). On the contrary, downregulation of *MdMYB1* inhibits anthocyanin accumulation mediated by ethylene, abscisic acid (ABA), wounding, drought, and different light intensities (An et al., 2018, 2019, 2020a,b). Notably, our results also revealed a significant upregulation of *LrMYB1* (DN33872) and *LrbHLH1* (DN48856), of which the expression was positively correlated with *LrDFR1, LrCHS*, *LrCHI*, *F3′H, LrUFGT* and *LrANS* genes during petal development stages (**Supplementary Figure 13**). This is similar to that of *LhMYB12*-Lat, which has previously been associated with activation of accumulation of anthocyanin in lily petals (Yamagishi et al., 2014). In our co-expression networks, the module that was positively correlated with anthocyanin contents and modules negatively correlated with anthocyanin content were identified. Overall, whether these *MYB* TFs interact with *bHLH* TFs to regulate anthocyanin biosynthesis in *L. radiata* remains to be further investigated.

The *LrDFR1* Drives Anthocyanin Accumulation in *Lycoris radiata* Petals

In the anthocyanin biosynthesis pathway, *DFR* catalyzes dihydroflavonol conversion to leucoanthocyanidins (Zhang et al., 2014). *DFR* belongs to the superfamily of short chain dehydrogenase reductase (SDR), which has a highly preserved NADPH-binding domain "VTGAAGFIGSWLIMRLLERGY" as well as a substrate-binding domain in plants (Martens et al., 2002; Haselmair-Gosch et al., 2018). In this study, based on the expression level of the anthocyanin structure genes obtained from the transcriptome data, *LrDFR1* and *LrDFR2* (**Table 1**) were found to highly associate with the total anthocyanin content (PCC $>$ 0.8), suggesting *DFR* may have an essential function in the phenotypic expression of *L. radiata* petal color. *LrDFR1* was then identified as one of the hub genes (**Figure 5B**) and important to positively regulate anthocyanin production in *L. radiata* petals. Multiple amino acid alignments showed that LrDFR1 contains the NADPH-binding domains and substrate-binding domains. Phylogenic tree analysis revealed a high similarity between LrDFR1 and other characterized DFRs, implying that LrDFR1 belongs to the monocot DFR family and exhibits catalytic characteristics.

The *DFR* genes of *Iris* and *Gentiana* have been reported to be associated with the absence of brick-red flowers (Noda et al., 2017). Moreover, heterologous *MaDFR* expressions in *N. tabacum* has been associated with enhanced anthocyanin accumulation, which leads to darker flower colors, suggesting that *MaDFR* is involved in flower color development (Liu H. et al., 2019). After the introduction of maize (*Zea mays*) *DFR* into white-flowered petunia varieties, transgenic plant flowers accumulate non-native pelargonidin, which results in novel brick red-flower varieties (Meyer et al., 1987). In this study, the expression patterns of *LrDFR1* was first temporally and spatially tested in various tissues and petal development stages of *L. radiata*. It showed that the expression levels of *LrDFR1* were correlated with total anthocyanin accumulation. These findings imply that *LrDFR1* is associated with petal color development in *L. radiata* (**Figures 1**, **7C,D**). The spatial and temporal expression characteristics of *LrDFR1* gene were found similarly in several other species (Liu H. et al., 2019; Lim et al., 2020). In order to investigate the functional divergence of *LrDFR1* gene in the flavonoid biosynthesis, we performed transient expression analyses using *Lycoris* petals and tobacco leaves. Overexpressed *LrDFR1* was associated with significantly elevated anthocyanin content and proanthocyanidin content in *Lycoris* petals and tobacco leaves. Interestingly, overexpression of *LrDFR1* also enhanced the expression of downstream genes (*LrANS* and *LrUFGT*) involved in anthocyanins biosynthesis in transgenic *Lycoris* petals (**Figures 8B–F**). In addition, for plant breeders, a single *DFR* gene maybe ideal for determining flower colors. *DFR* is vital for pigmentation, when compared to other

anthocyanin biosynthetic genes, which only regulate plant flower color hue. Thus, whether *LrDFR1* has a high preference for dihydromyricetin, and is accountable for the limited flower colors in *L. radiata* needs further study.

CONCLUSION

In this study, we provided a dynamic transcriptome profile of *L. radiata* petals during flower development stages. Overall, 56 structural genes and 27 key TF DEGs were identified as key genes responsible for *L. radiata* petal coloration. In the protein-protein interaction network analysis, *LrDFR1* was identified as a hub gene in the anthocyanin biosynthesis pathway, and was highly associated with anthocyanin accumulation. Overexpression of *LrDFR1* in *Lycoris* petals and tobacco leaves induced anthocyanin accumulation. In addition, the structural genes and co-expressed TFs reported in this study would serve as useful genetic resources for further functional characterization and molecular breeding programs in *L. radiata*. Taken together, our results elucidate on the molecular basis of petal development in *L. radiata*.

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: https://ngdc.cncb.ac.cn/, CRA004779.

AUTHOR CONTRIBUTIONS

NW and ZW designed the research and wrote the manuscript. NW performed most of the experiments and data analysis. XS and FZ collected the experimental materials. TW assisted with the data analysis. WZ provided helpful comments on the manuscript. ZW provided guidance on the whole study and contributed with valuable discussions. All authors read and approved the final manuscript.

FUNDING

This research was funded by the Jiangsu Key Laboratory for the Research and Utilization of Plant Resources (Grant No. JSPKLB202020). This work was jointly funded by Projects of Independently Development of Jiangsu Provincial Department of Science and Technology (BM2018021-5); Jiangsu Agricultural Science and Technology Innovation Fund [Grant Nos. CX(20)3171 and CX(19)3033]; the National Natural Science Foundation of China under Grant (Grant No. 31801900); and the Natural Science Foundation of Jiangsu Province (Grant No. BK20180310).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2021.761862/full#supplementary-material

REFERENCES

An, J. P., Liu, Y. J., Zhang, X. W., Bi, S. Q., Wang, X. F., You, C. X., et al. (2020a). Dynamic regulation of anthocyanin biosynthesis at different light intensities by the BT2-TCP46-MYB1 module in apple. *J. Exp. Bot.* 71, 3094–3109. doi: 10.1093/jxb/eraa056

An, J. P., Zhang, X. W., Bi, S. Q., You, C. X., Wang, X. F., and Hao, Y. J. (2020b). The ERF transcription factor MdERF38 promotes drought stress-induced anthocyanin biosynthesis in apple. *Plant J.* 101, 573–589. doi: 10.1111/tpj.14555

An, J. P., Qu, F. J., Yao, J. F., Wang, X. N., You, C. X., Wang, X. F., et al. (2017). The bZIP transcription factor MdHY5 regulates anthocyanin accumulation and nitrate assimilation in apple. *Hortic. Res.* 4:17023. doi: 10.1038/hortres.2017.56

An, J. P., Wang, X. F., Li, Y. Y., Song, L. Q., Zhao, L. L., You, C. X., et al. (2018). EIN3-LIKE1, MYB1, and ethylene response factor 3 act in a regulatory loop that synergistically modulates ethylene biosynthesis and anthocyanin accumulation. *Plant Physiol.* 178, 808–823.

An, J. P., Wang, X. F., Zhang, X. W., Bi, S. Q., You, C. X., and Hao, Y. J. (2019). MdBBX22 regulates UV-B-induced anthocyanin biosynthesis through regulating the function of MdHY5 and is targeted by MdBT2 for 26S proteasome-mediated degradation. *Plant Biotechnol. J.* 17, 2231–2233. doi: 10.1104/pp.18.00068

An, X. H., Tian, Y., Chen, K. Q., Liu, X. J., Liu, D. D., Xie, X. B., et al. (2015). MdMYB9 and MdMYB11 are involved in the regulation of the JA-induced biosynthesis of anthocyanin and proanthocyanidin in apples. *Plant Cell Physiol.* 56, 650–662. doi: 10.1093/pcp/pcu205

Ban, Y., Honda, C., Hatsuyama, Y., Igarashi, M., Bessho, H., and Moriguchi, T. (2007). Isolation and functional analysis of a MYB transcription factor gene that is a key regulator for the development of red coloration in apple skin. *Plant Cell Physiol.* 48, 958–970. doi: 10.1093/pcp/pcm066

Baudry, A., Caboche, M., and Lepiniec, L. (2006). TT8 controls its own expression in a feedback regulation involving TTG1 and homologous MYB and bHLH factors, allowing a strong and cell-specific accumulation of flavonoids in *Arabidopsis thaliana*. *Plant J.* 46, 768–779. doi: 10.1111/j.1365-313X.2006.02733.x

Casimiro-Soriguer, I., Narbona, E., Buide, M. L., Del Valle, J. C., and Whittall, J. B. (2016). Transcriptome and biochemical analysis of a flower color polymorphism in *Silene littorea* (Caryophyllaceae). *Front. Plant Sci.* 7:204. doi: 10.3389/fpls.2016.00204

Castaneda-Ovañdo, A., Pacheco-Hernández, M. D., Páez-Hernández, M. E., Rodríguez, J. A., and Galán-Vidal, C. A. (2009). Chemical studies of anthocyanins: a review. *Food Chem.* 113, 859–871. doi: 10.1016/j.foodchem.2008.09.001

Chun, J. H., Jang, I. H., Arasu, M. V., Aldhabi, N. A., Duraipandiyan, V., Lee, D. H., et al. (2013). Isolation and identification of alkaloids and anthocyanins from flower and bulb of *Lycoris radiata* using HPLC and LC-ESI-MS. *J. Agric. Chem. Environ.* 1, 22–2610.

Cui, X., Deng, J., Huang, C., Tang, X., Li, X., Li, X., et al. (2021). Transcriptomic analysis of the anthocyanin biosynthetic pathway reveals the molecular mechanism associated with purple color formation in Dendrobium Nestor. *Life* 11:113. doi: 10.3390/life11020113

Dong, Q., Zhao, H., Huang, Y., Chen, Y., Wan, M., Zeng, Z., et al. (2020). FtMYB18 acts as a negative regulator of anthocyanin/proanthocyanidin biosynthesis in Tartary buckwheat. *Plant Mol. Biol.* 104, 309–325. doi: 10.1007/s11103-020-01044-5

Espley, R. V., Hellens, R. P., Putterill, J., Stevenson, D. E., Kutty-Amma, S., and Allan, A. C. (2007). Red colouration in apple fruit is due to the activity of the MYB transcription factor, MdMYB10. *Plant J.* 49, 414–427. doi: 10.1111/j.1365-313X.2006.02964.x

Fang, H. C., Dong, Y. H., Yue, X. X., Hu, J. F., Jiang, S. H., Xu, H. F., et al. (2019). The B-box zinc finger protein MdBBX20 integrates anthocyanin accumulation

in response to ultraviolet radiation and low temperature. *Plant Cell Environ.* 42, 2090–2104. doi: 10.1111/pce.13552

Feng, C., Ding, D., Feng, C., and Kang, M. (2020). The identification of an R2R3-MYB transcription factor involved in regulating anthocyanin biosynthesis in *Primulina swinglei* flowers. *Gene* 752:144788. doi: 10.1016/j.gene.2020.144788

Feng, X., Zhang, Y., Wang, H., Tian, Z., Xin, S., and Zhu, P. (2021). The dihydroflavonol 4-reductase BoDFR1 drives anthocyanin accumulation in pink-leaved ornamental kale. *Theor. Appl. Genet.* 134, 159–169. doi: 10.1007/s00122-020-03688-9

Ferrer, J.-L., Austin, M. B., Stewart, C. Jr., and Noel, J. P. (2008). Structure and function of enzymes involved in the biosynthesis of phenylpropanoids. *Plant Physiol. Biochem.* 46, 356–370. doi: 10.1016/j.plaphy.2007.12.009

Francavilla, A., and Joye, I. J. (2020). Anthocyanins in whole grain cereals and their potential effect on health. *Nutrients* 12:2922. doi: 10.3390/nu12102922

Gallego, A. M., Rojas, L. F., Parra, O., Rodriguez, H. A., Mazo Rivas, J. C., Urrea, A. I., et al. (2018). Transcriptomic analyses of cacao cell suspensions in light and dark provide target genes for controlled flavonoid production. *Sci. Rep.* 8:13575. doi: 10.1038/s41598-018-31965-7

Gang, D. R. (2005). Evolution of flavors and scents. *Annu. Rev. Plant Biol.* 56, 301–325. doi: 10.1146/annurev.arplant.56.032604.144128

Gerats, T., and Strommer, J. (eds) (2009). "Petunia: evolutionary, developmental and physiological genetics," in *Petunia*, (New York, NY: Springer). doi: 10.1007/978-0-387-84796-2

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, 644–652. doi: 10.1038/nbt.1883

Griesbach, R. J. (2005). "Biochemistry and genetics of flower color," in *Plant Breeding Reviews*, ed. J. Janick (Hoboken, NJ: Wiley-Blackwell), 89–114. doi: 10.1002/9780470650301.ch4

Grotewold, E. (2006). The genetics and biochemistry of floral pigments. *Annu. Rev. Plant Biol.* 57, 761–780. doi: 10.1146/annurev.arplant.57.032905.105248

Handhayani, T., and Hiryanto, L. (2015). Intelligent Kernel K-means for clustering gene expression. *Procedia Comput. Sci.* 59, 171–177. doi: 10.1016/j.procs.2015.07.544

Haselmair-Gosch, C., Miosic, S., Nitarska, D., Roth, B. L., Walliser, B., Paltram, R., et al. (2018). Great cause-small effect: undeclared genetically engineered orange petunias harbor an inefficient dihydroflavonol 4-reductase. *Front. Plant Sci.* 9:149. doi: 10.3389/fpls.2018.00149

He, C., Liu, X., Teixeira da Silva, J. A., Liu, N., Zhang, M., and Duan, J. (2020). Transcriptome sequencing and metabolite profiling analyses provide comprehensive insight into molecular mechanisms of flower development in *Dendrobium officinale* (Orchidaceae). *Plant Mol. Biol.* 104, 529–548. doi: 10.1007/s11103-020-01058-z

He, Q. L., Cui, S. J., Gu, J. L., Zhang, H., Wang, M. X., Zhou, Y., et al. (2011). Analysis of floral transcription factors from *Lycoris longituba*. *Genomics* 96, 119–127. doi: 10.1016/j.ygeno.2010.04.002

Hsu, C. C., Chen, Y. Y., Tsai, W. C., Chen, W. H., and Chen, H. H. (2015). Three R2R3-MYB transcription factors regulate distinct floral pigmentation patterning in *Phalaenopsis* spp. *Plant Physiol.* 168, 175–191. doi: 10.1104/pp.114.254599

Huang, Y., Xuan, H., Yang, C., Guo, N., Wang, H., Zhao, J., et al. (2019). GmHsp90A2 is involved in soybean heat stress as a positive regulator. *Plant Sci.* 295, 26–33. doi: 10.1016/j.plantsci.2019.04.016

Jiang, G., Li, Z., Song, Y., Zhu, H., Lin, S., Huang, R., et al. (2019). LcNAC13 Physically interacts with LcR1MYB1 to coregulate anthocyanin biosynthesis-related genes during Litchi fruit ripening. *Biomolecules* 9:135. doi: 10.3390/biom9040135

Jiang, L., Fan, Z., Tong, R., Zhou, X., Li, J., and Yin, H. (2020). Functional diversification of the dihydroflavonol 4-reductase from *Camellia nitidissima* Chi. in the control of polyphenol biosynthesis. *Genes* 11:1341. doi: 10.3390/genes11111341

Jin, S. W., Rahim, M. A., Afrin, K. S., Park, J. I., Kang, J. G., and Nou, I. S. (2018). Transcriptome profiling of two contrasting ornamental cabbage (*Brassica oleracea* var. acephala) lines provides insights into purple and white inner leaf pigmentation. *BMC Genomics* 19:797. doi: 10.1186/s12864-018-5199-3

Jin, S. W., Rahim, M. A., Jung, H. J., Afrin, K. S., Kim, H. T., Park, J. I., et al. (2019). Abscisic acid and ethylene biosynthesis-related genes are associated with anthocyanin accumulation in purple ornamental cabbage (*Brassica oleracea* var. acephala). *Genome* 62, 513–526. doi: 10.1139/gen-2019-0038

Johnson, E. T., Ryu, S., Yi, H., Shin, B., Cheong, H., and Choi, G. (2001). Alteration of a single amino acid changes the substrate specificity of dihydroflavonol 4-reductase. *Plant J.* 25, 325–333. doi: 10.1046/j.1365-313x.2001.00962.x

Khoo, H. E., Azlan, A., Tang, S. T., and Lim, S. M. (2017). Anthocyanidins and anthocyanins: colored pigments as food, pharmaceutical ingredients, and the potential health benefits. *Food Nutr. Res.* 61:1361779. doi: 10.1080/16546628.2017.1361779

Koes, R., Verweij, W., and Quattrocchio, F. (2005). Flavonoids: a colorful model for the regulation and evolution of biochemical pathways. *Trends Plant Sci.* 10, 236–242. doi: 10.1016/j.tplants.2005.03.002

Lev-Yadun, S., and Gould, K. S. (2009). "Role of Anthocyanins in Plant Defence," in *Anthocyanins*, ed. C. Winefield, K. Davies and K. Gould K (Cham: Springer), 22–28.

Li, C., Qiu, J., Ding, L., Huang, M., Huang, S., Yang, G., et al. (2017). Anthocyanin biosynthesis regulation of *DhMYB2* and *DhbHLH1* in *Dendrobium* hybrids petals. *Plant Physiol. Biochem.* 112, 335–345. doi: 10.1016/j.plaphy.2017.01.019

Li, C., Wu, J., Hu, K. D., Wei, S. W., Sun, H. Y., Hu, L. Y., et al. (2020). PyWRKY26 and PybHLH3 cotargeted the PyMYB114 promoter to regulate anthocyanin biosynthesis and transport in red-skinned pears. *Hortic. Res.* 7:37. doi: 10.1038/s41438-020-0254-z

Li, D., Liu, X., Shu, L., Zhang, H., Zhang, S., Song, Y., et al. (2020). Global analysis of the AP2/ERF gene family in rose (*Rosa chinensis*) genome unveils the role of RcERF099 in Botrytis resistance. *BMC Plant Biol.* 20:553. doi: 10.1186/s12870-020-02740-6

Li, Y., Shan, X., Zhou, L., Gao, R., Yang, S., Wang, S., et al. (2019). The R2R3-MYB factor FhMYB5 from *Freesia hybrida* contributes to the regulation of anthocyanin and proanthocyanidin biosynthesis. *Front. Plant Sci.* 9:1935. doi: 10.3389/fpls.2018.01935

Li, Z., Zhao, M., Jin, J., Zhao, L., and Xu, Z. (2018). Anthocyanins and their biosynthetic genes in three novel-colored *Rosa rugosa* cultivars and their parents. *Plant Physiol. Biochem.* 129, 421–428. doi: 10.1016/j.plaphy.2018.06.028

Lim, S. H., Park, B., Kim, D. H., Park, S., Yang, J. H., Jung, J. A., et al. (2020). Cloning and functional characterization of dihydroflavonol 4-reductase gene involved in anthocyanin biosynthesis of chrysanthemum. *Int. J. Mol. Sci.* 21:7960. doi: 10.3390/ijms21217960

Liu, H., Lou, Q., Ma, J., Su, B., Gao, Z., and Liu, Y. (2019). Cloning and functional characterization of dihydroflavonol 4-Reductase gene involved in anthocyanidin biosynthesis of grape hyacinth. *Int. J. Mol. Sci.* 20:4743. doi: 10.3390/ijms20194743

Liu, Y., Lin-Wang, K., Espley, R. V., Wang, L., Li, Y., Liu, Z., et al. (2019). StMYB44 negatively regulates anthocyanin biosynthesis at high temperatures in tuber flesh of potato. *J. Exp. Bot.* 70, 3809–3824. doi: 10.1093/jxb/erz194

Livak, K. J., and Schmittgen, T. D. (2001). Analysis of relative gene expression data using real-time quantitative PCR and the 2- $\Delta\Delta$CT method. *Methods* 25, 402–408. doi: 10.1006/meth.2001.1262

Lloyd, A., Brockman, A., Aguirre, L., Campbell, A., Bean, A., Cantero, A., et al. (2017). Advances in the MYB–bHLH–WD repeat (MBW) pigment regulatory model: addition of a WRKY factor and co-option of an anthocyanin MYB for betalain regulation. *Plant Cell Physiol.* 58, 1431–1441. doi: 10.1093/pcp/pcx075

Lou, Q., Liu, Y., Qi, Y., Jiao, S., Tian, F., Jiang, L., et al. (2014). Transcriptome sequencing and metabolite analysis reveals the role of delphinidin metabolism in flower colour in grape hyacinth. *J. Exp. Bot.* 65, 3157–3164. doi: 10.1093/jxb/eru168

Lu, W., Chen, J., Ren, X., Yuan, J., Han, X., Mao, L., et al. (2018). One novel strawberry MADS-box transcription factor FaMADS1a acts as a negative regulator in fruit ripening. *Sci. Hortic.* 227, 124–131.

Luo, P., Ning, G. G., Wang, Z., Shen, Y. X., Jin, H. A., Li, P. H., et al. (2016). Disequilibrium of flavonol synthase and dihydroflavonol-4-reductase expression associated tightly to White vs. Red color flower formation in plants. *Front. Plant Sci* 6:1257. doi: 10.3389/fpls.2015.01257

Ma, R., Xu, S., Zhao, Y. C., Xia, B., and Wang, R. (2016). Selection and validation of appropriate reference genes for quantitative real-time PCR analysis of gene expression in *Lycoris aurea*. *Front. Plant Sci.* 7:536. doi: 10.3389/fpls.2016.00536

Martens, S., Teeri, T., and Forkmann, G. (2002). Heterologous expression of dihydroflavonol 4-reductases from various plants. *FEBS Lett.* 531, 453–458. doi: 10.1016/s0014-5793(02)03583-4

Mehrtens, F., Kranz, H., Bednarek, P., and Weisshaar, B. (2005). The Arabidopsis transcription factor MYB12 is a flavonol-specific regulator of phenylpropanoid biosynthesis. *Plant Physiol.* 138, 1083–1096. doi: 10.1104/pp.104.058032

Meyer, P., Heidmann, I., Forkmann, G., and Saedler, H. (1987). A new petunia flower colour generated by transformation of a mutant with a maize gene. *Nature* 330, 677–678. doi: 10.1038/330677a0

Mortazavi, A., Williams, B. A., Mccue, K., Schaeffer, L., and Wold, B. (2008). Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nat. Methods* 5, 621–628. doi: 10.1038/nmeth.1226

Nakatsuka, A., Izumi, Y., and Yamagishi, M. (2003). Spatial and temporal expression of chalcone synthase and dihydroflavonol 4-reductase genes in the Asiatic hybrid lily. *Plant Sci.* 165, 759–767

Niu, S. S., Xu, C. J., Zhang, W. S., Zhang, B., Li, X., Lin-Wang, K., et al. (2010). Coordinated regulation of anthocyanin biosynthesis in chinese bayberry (*Myrica rubra*) fruit by an R2R3 MYB transcription factor. *Planta* 231, 887–899. doi: 10.1007/s00425-009-1095-z

Noda, N., Yoshioka, S., Kishimoto, S., Nakayama, M., Douzono, M., Tanaka, Y., et al. (2017). Generation of blue chrysanthemums by anthocyanin B-ring hydroxylation and glucosylation and its coloration mechanism. *Sci. Adv.* 3:e1602785. doi: 10.1126/sciadv.1602785

Park, C. H., Yeo, H. J., Kim, Y. J., Nguyen, B. V., Park, Y. E., Sathasivam, R., et al. (2021). Profiles of secondary metabolites (phenolic acids, carotenoids, anthocyanins, and galantamine) and primary metabolites (carbohydrates, amino acids, and organic acids) during flower development in *Lycoris radiata*. *Biomolecules* 11:248. doi: 10.3390/biom11020248

Park, C. H., Yeo, H. J., Park, Y. E., Baek, S. A., Kim, J. K., and Park, S. U. (2019). Transcriptome analysis and metabolic profiling of *Lycoris radiata*. *Biology* 8:63. doi: 10.3390/biology8030063

Qi, T., Song, S., Ren, Q., Wu, D., Huang, H., Chen, Y., et al. (2011). The Jasmonate-ZIM-domain proteins interact with the WD-Repeat/bHLH/MYB complexes to regulate jasmonate-mediated anthocyanin accumulation and trichome initiation in *Arabidopsis thaliana*. *Plant Cell* 23, 1795–1814. doi: 10.1105/tpc.111.083261

Qi, Y., Gu, C., Wang, X., Gao, S., Li, C., Zhao, C., et al. (2020). Identification of the Eutrema salsugineum EsMYB90 gene important for anthocyanin biosynthesis. *BMC Plant Biol.* 20:186. doi: 10.1186/s12870-020-02391-7

Quattrocchio, F., Verweij, W., Kroon, A., Spelt, C., Mol, J., and Koes, R. (2006). PH4 of petunia is an R2R3 MYB protein that activates vacuolar acidification through interactions with basic-helix-loop-helix transcription factors of the anthocyanin pathway. *Plant Cell* 18, 1274–1291. doi: 10.1105/tpc.105.034041

Rahim, M. A., Busatto, N., and Trainotti, L. (2014). Regulation of anthocyanin biosynthesis in peach fruits. *Planta* 240, 913–929. doi: 10.1007/s00425-014-2078-2

Rameneni, J. J., Choi, S. R., Chhapekar, S. S., Man-Sun, K., Singh, S., Yi, S. Y., et al. (2020). Red Chinese cabbage transcriptome analysis reveals structural genes and multiple transcription factors regulating reddish purple color. *Int. J. Mol. Sci.* 21:2901. doi: 10.3390/ijms21082901

Rebecca, O. P. S., Boyce, A. N., and Somasundram, C. (2010). Pigment identification and antioxidant properties of red dragon fruit (Hylocereus polyrhizus). *Afr. J. Biotechnol.* 9, 1450–1454. doi: 10.5897/AJB09.1603

Rice, P., Longden, L., and Bleasby, A. (2000). EMBOSS: the European molecular biology open software suite. *Trends Genet.* 16, 276–277. doi: 10.1016/S0168-9525(00)02024-2

Rosas-Saavedra, C., and Stange, C. (2016). Biosynthesis of carotenoids in plants: enzymes and color. *Subcell. Biochem.* 79, 35–69. doi: 10.1007/978-3-319-39126-7_2

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

Sheludko, Y. V., Sindarovska, Y. R., Gerasymenko, I. M., Bannikova, M. A., and Kuchuk, N. V. (2007). Comparison of several *Nicotiana* species as hosts for high-scale *Agrobacterium*-mediated transient expression. *Biotechnol. Bioeng.* 96, 608–614. doi: 10.1002/bit.21075

Shi, Q., Zhou, L., Wang, Y., Li, K., Zheng, B., and Miao, K. (2015). Transcriptomic analysis of *Paeonia delavayi* wild population flowers to identify differentially expressed genes involved in purple-red and yellow petal pigmentation. *PLoS One* 10:e0135038. doi: 10.1371/journal.pone.0135038

Shi, S. G., Yang, M., Zhang, M., Wang, P., Kang, Y. X., and Liu, J. J. (2014). Genome-wide transcriptome analysis of genes involved in flavonoid biosynthesis between red and white strains of *Magnolia sprengeri* pamp. *BMC Genomics* 15:706. doi: 10.1186/1471-2164-15-706

Shimada, N., Sasaki, R., Sato, S., Kaneko, T., Tabata, S., Aoki, T., et al. (2005). A comprehensive analysis of six dihydroflavonol 4-reductases encoded by a gene cluster of the Lotus japonicus genome. *J. Exp. Bot.* 56, 2573–2585. doi: 10.1093/jxb/eri251

Sun, C., Deng, L., Du, M., Zhao, J., Chen, Q., Huang, T., et al. (2020). A transcriptional network promotes anthocyanin biosynthesis in tomato flesh. *Mol. Plant* 13, 42–58. doi: 10.1016/j.molp.2019.10.010

Suzuki, K., Suzuki, T., Nakatsuka, T., Dohra, H., Yamagishi, M., Matsuyama, K., et al. (2016). RNA-seq-based evaluation of bicolor tepal pigmentation in *Asiatic hybrid* lilies (*Lilium* spp.). *BMC Genomics* 17:611. doi: 10.1186/s12864-016-2995-5

Takos, A. M., Jaffé, F. W., Jacob, S. R., Bogs, J., Robinson, S. P., and Walker, A. R. (2006). Light-induced expression of a MYB gene regulates anthocyanin biosynthesis in red apples. *Plant Physiol.* 142, 1216–1232. doi: 10.1104/pp.106.088104

Tanaka, Y., Sasaki, N., and Ohmiya, A. (2008). Biosynthesis of plant pigments: anthocyanins, betalains and carotenoids. *Plant J.* 54, 733–749. doi: 10.1111/j.1365-313X.2008.03447.x

Tang, Q., Tian, M., An, G., Zhang, W., Chen, J., and Yan, C. (2017). Rapid identification of the purple stem Ps gene of Chinese kale (*Brassica oleracea* var. alboglabra) in a segregation distortion population by bulked segregant analysis and RNA sequencing. *Mol. Breed.* 37:153. doi: 10.1007/s11032-017-0752-3

Wang, Q., Wang, Y., Sun, H., Sun, L., and Zhang, L. (2020). Transposon-induced methylation of the RsMYB1 promoter disturbs the anthocyanin accumulation in red-fleshed radish. *J. Exp. Bot.* 71, 2537–2550. doi: 10.1093/jxb/eraa010

Xiang, L., Liu, X., Li, H., Yin, X., Grierson, D., Li, F., et al. (2019). CmMYB#7, an R3 MYB transcription factor, acts as a negative regulator of anthocyanin biosynthesis in chrysanthemum. *J. Exp. Bot.* 70, 3111–3123.

Xie, D. Y., Sharma, S. B., Paiva, N. L., Ferreira, D., and Dixon, R. A. (2003). Role of anthocyanidin reductase, encoded by BANYULS in plant flavonoid biosynthesis. *Science* 299, 396–399. doi: 10.1126/science.1078540

Xie, Y., Tan, H., Ma, Z., and Huang, J. (2016). DELLA proteins promote anthocyanin biosynthesis via sequestering MYBL2 and JAZ suppressors of the MYB/bHLH/WD40 complex in *Arabidopsis thaliana*. *Mol. Plant* 9, 711–721. doi: 10.1016/j.molp.2016.01.014

Xu, L., Yang, P., Feng, Y., Xu, H., Cao, Y., Tang, Y., et al. (2017). Spatiotemporal transcriptome analysis provides insights into bicolor tepal development in Lilium "Tiny Padhye". *Front. Plant Sci.* 8:398. doi: 10.3389/fpls.2017.00398

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

Yamagishi, M., Toda, S., and Tasaki, K. (2014). The novel allele of the LhMYB12 gene is involved in splatter-type spot formation on the flower tepals of Asiatic hybrid lilies (*Lilium* spp.). *New Phytol.* 201, 1009–1020. doi: 10.1111/nph.12572

Yang, X., Xia, X., Zhang, Z., Nong, B., Zeng, Y., Wu, Y., et al. (2019). Identification of anthocyanin biosynthesis genes in rice pericarp using PCAMP. *Plant Biotechnol. J.* 17, 1700–1702. doi: 10.1111/pbi.13133

Young, M. D., Wakefield, M. J., Smyth, G. K., and Oshlack, A. (2010). Gene ontology analysis for RNA-seq: accounting for selection bias. *Genome Biol.* 11:R14. doi: 10.1186/gb-2010-11-2-r14

Yue, Y., Liu, J., Shi, T., Chen, M., Li, Y., Du, J., et al. (2019). Integrating transcriptomic and GC-MS metabolomic analysis to characterize color and aroma formation during tepal development in *Lycoris longituba*. *Plants* 8:53. doi: 10.3390/plants8030053

Zhai, R., Wang, Z., Yang, C., Lin-Wang, K., Espley, R., Liu, J., et al. (2019). PbGA2ox8 induces vascular-related anthocyanin accumulation and contributes to red stripe formation on pear fruit. *Hortic. Res.* 6:137. doi: 10.1038/s41438-019-0220-9

Zhang, F., Wang, T., Shu, X., Wang, N., Zhuang, W., and Wang, Z. (2020). Complete chloroplast genomes and comparative analyses of *L. chinensis*, *L. anhuiensis*, and *L. aurea* (Amaryllidaceae). *Int. J. Mol. Sci.* 21:5729. doi: 10.3390/ijms21165729

Zhang, H., Jin, J., Tang, L., Zhao, Y., Gu, X., Gao, G., et al. (2011). PlantTFDB 2.0: update and improvement of the comprehensive plant transcription factor database. *Nucleic Acids Res.* 9, D1114–D1117. doi: 10.1093/nar/gkq1141

Zhang, X., Xu, Z., Yu, X., Zhao, L., Zhao, M., Han, X., et al. (2019). Identification of two novel R2R3-MYB transcription factors, PsMYB114L and PsMYB12L, related to anthocyanin biosynthesis in *Paeonia suffruticosa*. *Int. J. Mol. Sci.* 20:1055. doi: 10.3390/ijms20051055

Zhang, Y., Butelli, E., and Martin, C. (2014). Engineering anthocyanin biosynthesis in plants. *Curr. Opin. Plant Biol.* 19, 81–90. doi: 10.1016/j.pbi.2014.05.011

Zhao, D., Jiang, Y., Ning, C., Meng, J., Lin, S., Ding, W., et al. (2014). Transcriptome sequencing of a chimaera reveals coordinated expression of anthocyanin biosynthetic genes mediating yellow formation in herbaceous peony (*Paeonia lactiflora* Pall). *BMC Genomics* 15:689. doi: 10.1186/1471-2164-15-689

Zhao, D., and Tao, J. (2015). Recent advances on the development and regulation of flower color in ornamental plants. *Front. Plant Sci.* 6:261. doi: 10.3389/fpls.2015.00261

Zhao, D., Tao, J., Han, C., and Ge, J. (2012). Flower color diversity revealed by differential expression of flavonoid biosynthetic genes and flavonoid accumulation in herbaceous peony (*Paeonia lactiflora* Pall.). *Mol. Biol. Rep.* 39, 11263–11275. doi: 10.1007/s11033-012-2036-7

Zhong, C., Tang, Y., Pang, B., Li, X., Yang, Y., Deng, J., et al. (2020). The R2R3-MYB transcription factor GhMYB1a regulates flavonol and anthocyanin accumulation in *Gerbera hybrida*. *Hortic. Res.* 7:78. doi: 10.1038/s41438-020-0296-2

Zhou, C., Mei, X., Rothenberg, D. O'., Yang, Z., Zhang, W., Wan, S., et al. (2020). Metabolome and transcriptome analysis reveals putative genes involved in anthocyanin accumulation and coloration in white and pink tea (*Camellia sinensis*) Flower. *Molecules* 25:190. doi: 10.3390/molecules25010190

Zhou, H., Lin-Wang, K., Wang, H., Gu, C., Dare, A. P., Espley, R. V., et al. (2015). Molecular genetics of blood-fleshed peach reveals activation of anthocyanin biosynthesis by NAC transcription factors. *Plant J.* 82, 105–121. doi: 10.1111/tpj.12792

Zhu, H. H., Yang, J. X., Xiao, C. H., Mao, T. Y., Zhang, J., and Zhang, H. Y. (2019). Differences in flavonoid pathway metabolites and transcripts affect yellow petal colouration in the aquatic plant *Nelumbo nucifera*. *BMC Plant Biol.* 19:277. doi: 10.1186/s12870-019-1886-8

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.

frontiers in Plant Science

ORIGINAL RESEARCH
published: 11 February 2022
doi: 10.3389/fpls.2022.829089

Identifying *Bixa orellana* L. New Carotenoid Cleavage Dioxygenases 1 and 4 Potentially Involved in Bixin Biosynthesis

*Rosa Us-Camas[1], Margarita Aguilar-Espinosa[1], Jacobo Rodríguez-Campos[2], Alba Adriana Vallejo-Cardona[3], Víctor Manuel Carballo-Uicab[4], Hugo Serrano-Posada[4] and Renata Rivera-Madrid[1]**

[1] Unidad de Bioquímica y Biología Molecular de Plantas, Centro de Investigación Científica de Yucatán A.C., Mérida, Mexico, [2] Unidad de Servicios Analíticos y Metrológicos, Centro de Investigación y Asistencia en Tecnología y Diseño del Estado de Jalisco, Guadalajara, Mexico, [3] Unidad de Biotecnología Médica y Farmacéutica, CONACYT, Centro de Investigación y Asistencia en Tecnología y Diseño del Estado de Jalisco, Guadalajara, Mexico, [4] CONACYT, Laboratorio de Biología Sintética, Estructural y Molecular, Laboratorio de Agrobiotecnología, Colima, Mexico

OPEN ACCESS

Edited by:
Li Li,
Cornell University, United States

Reviewed by:
José-Antonio Daròs,
Polytechnic University of Valencia, Spain
Tianhu Sun,
Cornell University, United States

***Correspondence:**
Renata Rivera-Madrid
renata@cicy.mx

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

Received: *04 December 2021*
Accepted: *19 January 2022*
Published: *11 February 2022*

Citation:
Us-Camas R, Aguilar-Espinosa M, Rodríguez-Campos J, Vallejo-Cardona AA, Carballo-Uicab VM, Serrano-Posada H and Rivera-Madrid R (2022) Identifying Bixa orellana L. New Carotenoid Cleavage Dioxygenases 1 and 4 Potentially Involved in Bixin Biosynthesis. Front. Plant Sci. 13:829089. doi: 10.3389/fpls.2022.829089

Carotene cleavage dioxygenases (CCDs) are a large family of Fe^{2+} dependent enzymes responsible for the production of a wide variety of apocarotenoids, such as bixin. Among the natural apocarotenoids, bixin is second in economic importance. It has a red-orange color and is produced mainly in the seeds of *B. orellana*. The biosynthesis of bixin aldehyde from the oxidative cleavage of lycopene at 5,6/5′,6′ bonds by a CCD is considered the first step of bixin biosynthesis. Eight *BoCCD* (*BoCCD1-1, BoCCD1-3, BoCCD1-4, CCD4-1, BoCCD4-2, BoCCD4-3* and *BoCCD4-4*) genes potentially involved in the first step of *B. orellana* bixin biosynthesis have been identified. However, the cleavage activity upon lycopene to produce bixin aldehyde has only been demonstrated for BoCCD1-1 and BoCCD4-3. Using *in vivo* (*Escherichia coli*) and *in vitro* approaches, we determined that the other identified BoCCDs enzymes (BoCCD1-3, BoCCD1-4, BoCCD4-1, BoCCD4-2, and BoCCD4-4) also participate in the biosynthesis of bixin aldehyde from lycopene. The LC-ESI-QTOF-MS/MS analysis showed a peak corresponding to bixin aldehyde (*m/z* 349.1) in pACCRT-EIB *E. coli* cells that express the BoCCD1 and BoCCD4 proteins, which was confirmed by *in vitro* enzymatic assay. Interestingly, in the *in vivo* assay of BoCCD1-4, BoCCD4-1, BoCCD4-2, and BoCCD4-4, bixin aldehyde was oxidized to norbixin (*m/z* 380.2), the second product of the bixin biosynthesis pathway. *In silico* analysis also showed that BoCCD1 and BoCCD4 proteins encode functional dioxygenases that can use lycopene as substrate. The production of bixin aldehyde and norbixin was corroborated based on their ion fragmentation pattern, as well as by Fourier transform infrared (FTIR) spectroscopy. This work made it possible to clarify at the same time the first and second steps of the bixin biosynthesis pathway that had not been evaluated for a long time.

Keywords: *BoCCD1*, *BoCCD4*, bixin aldehyde, norbixin, bixin biosynthesis, *Bixa orellana* L., apocarotenoids, carotenoids

INTRODUCTION

Carotenoids comprise a broad group of mostly colored molecules on which the oxidative cleavage of carotenoid cleavage oxygenases [CCOs — also know as carotenoid cleavage dioxygenases (CCDs)] — act to produce numerous derivatives known as apocarotenoids. Some of these apocarotenoids have color, such as bixin, a red-orange monomethyl ester of the C_{24} dicarboxylic acid norbixin, that is accumulated mainly in the seeds of *Bixa orellana* (achiote or annatto-tree) (Rivera-Madrid et al., 2016). Achiote is a perennial tree species native to the Amazonia and tropical regions (Akshatha et al., 2011; Moreira et al., 2015). Its economic importance lies in the high content of bixin in its seeds (about 80% of total carotenoids). Bixin is a red-orange apocarotenoid widely used in the food, pharmaceutical, cosmetic, and textile industries (Tirimanna, 1981; Giuliano et al., 2003; Rivera-Madrid et al., 2006, 2016). The demand for bixin has increased in recent years due to its potential use as a natural dye and its health and environmental benefits over synthetic dyes (Evans, 2000; Devi et al., 2013; Shahid ul et al., 2016). The apocarotenoid bixin is produced by the oxidative cleavage at the double bonds of the carotenoid polyene chain by the action of the carotenoid cleavage dioxygenases (CCDs), a large family of non-iron (II) dependent enzymes present in all taxa (Auldridge et al., 2006; Hou et al., 2016; Ramamoorthy et al., 2020). In plants, CCDs are divided into two functionally different groups: the 9-*cis* epoxycarotenoid dioxygenases (NCEDs) and the CCDs (Walter et al., 2010). NCEDs participate in the biosynthesis of ABA through cleavage at the 11,12 double bonds of 9-*cis*-violaxanthin and 9-*cis*-neoxanthin to produce xanthoxin (C_{15}), the ABA precursor (Tan et al., 2003; Us-Camas et al., 2020). CCDs have recently been classified into six subfamilies: CCD1, CCD2, CCD4, CCD7, CCD8, and ZAS (zaxinone synthase), which vary in their substrate specificity and cleavage sites (Auldridge et al., 2006; Frusciante et al., 2014; Hou et al., 2016; Fiorilli et al., 2019; Rivera-Madrid and Ramamoorthy, 2020). *In vitro* analyzes indicate that CCD1 enzymes can cleave different cyclic and linear carotenoids at the 5,6/5′,6′, 7,6/7′,6′, and 9,10/9′,10′ double bonds, generating a great diversity of apocarotenoids as products (Schwartz et al., 2001; Mathieu et al., 2005; Rubio-Moraga et al., 2008; Vogel et al., 2008; Ilg et al., 2009; Lashbrooke et al., 2013). In planta, CCD1 cleaves C_{27} apocarotenoid at 9,10/9′,10′ positions to produce volatile C_{13} (β-ionone) and C_{14} (dialdehydes) apocarotenoids involved in flavor, aroma, and root colonization (Floss et al., 2008; Floss and Walter, 2009). On the other hand, CCD4 enzymes are capable of cleaving the 5,6/5′,6′, 7,6/7′,6′, and 9,10/9′,10′ positions of different carotenoids (Bouvier et al., 2003; Rubio-Moraga et al., 2008; Lashbrooke et al., 2013; Ma et al., 2013; Rodrigo et al., 2013). CCD4 enzymes are mainly involved in the regulation of volatile apocarotenoids biosynthesis (Rubio-Moraga et al., 2008; Brandi et al., 2011; Lashbrooke et al., 2013), as well as in the regulation of pigmentation of flower petals, fruits, and seeds (Ohmiya et al., 2006; Campbell et al., 2010; Brandi et al., 2011; Gonzalez-Jorge et al., 2013; Rodrigo et al., 2013; Ko et al., 2018). CCD4 enzymes generally have a chloroplast transit peptide in their sequence. In contrast, CCD1 enzymes are located in the cytosol (Rubio-Moraga et al., 2008). Some CCD4 enzymes are associated with the plastoglobules of chloroplasts, where they perform carotenoid cleavage (Ytterberg et al., 2006; Rottet et al., 2016).

The investigation carried out by Bouvier et al. (2003) demonstrated that bixin biosynthesis begins with the oxidative cleavage of C_{40} lycopene at the 5,6/5′,6′ double bonds by the CCD4 enzyme lycopene dioxygenase (BoLCD), to produce bixin aldehyde. Subsequently, the aldehyde groups are oxidized to carboxyls by bixin aldehyde dehydrogenase (BoALDH) producing norbixin, and finally, bixin is produced when one of the carboxyl group of norbixin is methylated by norbixin methyltransferase (BoMTH) (Jako et al., 2002; Bouvier et al., 2003; **Figure 1**). The sequential activity of these three enzymes in lycopene-producing *Escherichia coli* cells produced bixin (Bouvier et al., 2003). However, the *BoLCD* (CCD4 type) sequence, as well as the *BoALDH* and *BoMTH* genes involved in bixin biosynthesis (Bouvier et al., 2003) were not found in the leaves nor in mature and immature seed transcriptomes of the local Yucatecan accession "Peruana Roja" of *B. orellana* (Cárdenas-Conejo et al., 2015). Instead, other members of the *BoCCDs*, *BoALDHs* and, *SABATHs* (methyltransferases) gene families potentially involved in the bixin biosynthesis were identified. Derived from achiote transcriptomes analysis, Cárdenas-Conejo et al. (2015) identified eight candidate *BoCCDs* genes involved in bixin biosynthesis, distributed in hypothetical cytosolic and plastidic pathways, four of them belonging to the *BoCCD1* subfamily: *BoCCD1-1*, *BoCCD1-2*, *BoCCD1-3*, and *BoCCD1-4*; and four of them to the *BoCCD4* subfamily: *BoCCD4-1*, *BoCCD4-2*, *BoCCD4-3* and, *BoCCD4-4* (Cárdenas-Conejo et al., 2015). Recently, Carballo-Uicab et al. (2019) demonstrated that the recombinant proteins BoCCD1-1 and BoCCD4-3 exhibited cleavage activity at the 5,6/5′,6′ positions of lycopene producing bixin aldehyde, indicating their involvement in the first step of bixin biosynthesis (Carballo-Uicab et al., 2019). Other *BoCCD1* and *BoCCD4* genes have been isolated from different *B. orellana* cultivars, but their role in bixin biosynthesis has not been unveiled (Rodríguez-Ávila et al., 2011b; Soares et al., 2011; Sankari et al., 2016). In general, the expression levels of all *BoCCD* genes identified in the transcriptomes increased in the immature stages of the seed (S3 and S4) when bixin reached its maximum levels and decreased in the mature stages of the seed (S5) when bixin levels dropped dramatically in the two accessions that produce contrasting levels of bixin (N4P and P13W) (Carballo-Uicab et al., 2019). Therefore, we hypothesize spatio-temporal participation in bixin biosynthesis of all the BoCCD1 and BoCCD4 enzymes identified in the transcriptomes of *B. orellana*. Furthermore, it has been suggested that the BoCCD1 and BoCCD4 enzymes are distributed in two hypothetical pathways, a cytosolic and a plastidic, respectively, where they can act individually or in coordination to synthesize bixin, reinforcing our hypothesis (Cárdenas-Conejo et al., 2015). For this reason, in the present study, we used different approaches to determine if other BoCCD1 and BoCCD4 enzymes, in addition to BoCCD1-1, and BoCCD4-3, have 5/6, 5′/6′ cleavage activity on lycopene and if they are involved in the first step of bixin biosynthesis.

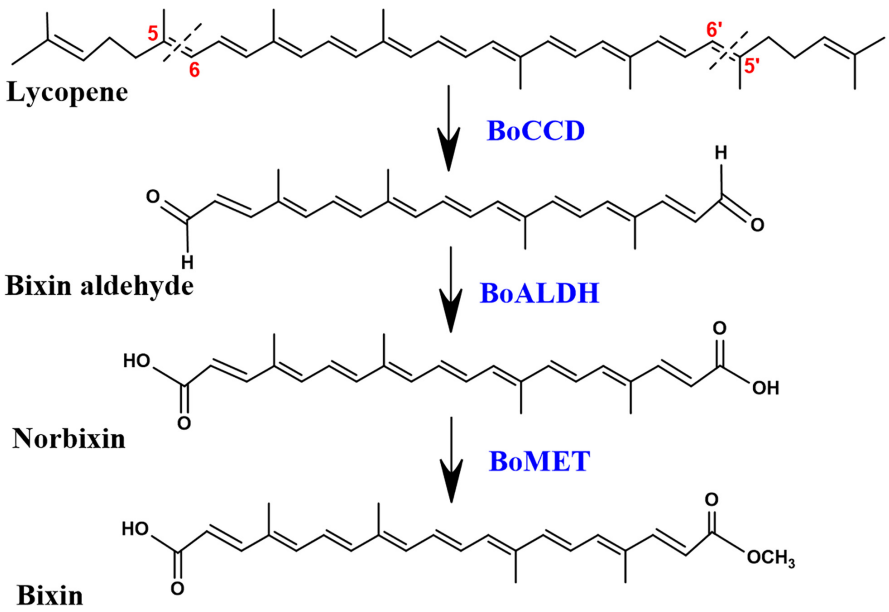


FIGURE 1 | Bixin biosynthesis pathway. All compounds are in the *trans* configuration. BoCCD, carotene cleavage dioxygenase; BoALDH, bixin aldehyde dehydrogenase; BoMET, norbixin methyltransferase.

We isolated, cloned, and expressed seven of the eight *BoCCD* genes reported in *B. orellana*; three of them belong to the *BoCCD1* subfamily: *BoCCD1-1*, *BoCCD1-3,* and *BoCCD1-4*; and four belonging to the *BoCCD4* subfamily: *BoCCD4-1*, *BoCCD4-2*, *BoCCD4-3,* and *BoCCD4-4*. Sequence alignment and phylogenetic analysis indicated that all isolated BoCCD1 and BoCCD4 proteins encode functional carotenoid cleavage dioxygenases. The production of bixin aldehyde (*m/z* 349) from lycopene was confirmed by Liquid Chromatography Electrospray Ionization Quadrupole Time-of-Flight Mass Spectrometry (LC-ESI-QTOF-MS/MS) analyzes in the *in vivo* (*E. coli*) assay. *In vitro* enzymatic assays indicated that in addition to BoCCD1-1 and BoCCD4-3; BoCCD1-3, BoCCD1-4, BoCCD4-1, BoCCD4-2, and BoCCD4-4 enzymes have cleavage activity at the 5,6/5′,6′ double bonds of lycopene. Three-D structural modeling and docking analyzes also indicated that all the isolated BoCCD1 and BoCCD4 proteins can use lycopene as a substrate. Interestingly, we found that in the *in vivo* assay of BoCCD1-4, BoCCD4-1, BoCCD4-2, and BoCCD4-4, bixin aldehyde (*m/z* 349.1) was oxidized to norbixin (*m/z* 380.2), the second product of the bixin biosynthesis pathway. Our results suggest that the second product of bixin biosynthesis pathway, norbixin, can be generated by an oxidation process in *E. coli* cells, but further studies are required. The production of bixin aldehyde and norbixin were corroborated based on their ion fragmentation patterns and by Fourier transform infrared (FTIR) spectroscopy analysis. This work allowed the identification of new BoCCD enzymes potentially involved in the first step of bixin biosynthesis in *B. orellana* and at the same time clarified the first and second step of bixin biosynthesis pathway that had not been evaluated for a long time.

MATERIALS AND METHODS

Plant Material

The study used *B. orellana* N4P accession, which has distinctive morphological characteristics: pink flowers, dehiscent green fruits with red spines, and seeds with high bixin contents (Trujillo-Hdz et al., 2016). Young leaves from the N4P accession were collected and immediately immersed in liquid nitrogen then stored at −80°C until subsequent RNA extraction.

Total RNA Extraction and cDNA Synthesis

Total RNA was extracted from the leaves of *B. orellana* using the PureLink RNA Mini Kit (Cat. No. 12183018A, Invitrogen) as described by Rodríguez-Ávila et al. (2009). First-strand cDNA was synthesized from 1 μg of total RNA using the SuperScriptTM III reverse transcriptase (Cat. No. 18080-093) according to the manufacturer's instructions.

Isolation and Cloning of *BoCCD1* and *BoCCD4* Genes

The coding region of *BoCCD1-1*, *BoCCD1-3*, *BoCCD1-4*, *BoCCD4-1*, *BoCCD4-2*, *BoCCD4-3*, and *BoCCD4-4* was isolated from leaf cDNA of the N4P accession of *B. orellana* using the primers listed in **Supplementary Table 1**. The primers were designed using the sequences of the *BoCCDs* found in the transcriptome and reported in the GeneBank (Cárdenas-Conejo et al., 2015). The PCR reactions were performed using the Phusion High-Fidelity DNA Polymerase (Cat. No. F534S, Thermo Fisher Scientific). The reaction contained 0.5 μl of cDNA at ∼1000 ng/μl, 4 μl of 5X Phusion Green HF Buffer,

0.1 μl of $MgCl_2$ at 50 mM, 0.4 μl of a mix of dNTPs at 10 mM, 0.5 μl of specific forward and reverse primers at 10 Mm, 0.1 μl of Phusion High-Fidelity DNA Polymerase at 2 U/μl, and sterile water to a final volume of 20 μl. PCR reactions were carried out with the following program: initial denaturation at 98°C for 30 s, denaturation at 98°C for 15 s, alignment for 30 s at the temperature indicated in **Supplementary Table 1**, extensión at 72°C for 30 s for 35 cycles, and a final extension at 72°C for 10 min. The PCR products were resolved on SYBR Green I-stained agarose gels (Cat. No. S7563, Thermo Fisher Scientific), and the fragments with the expected sizes (**Supplementary Figure 1**) were purified using the QIAquick gel extraction kit (Cat. No. 27104, QIAGEN), following the supplier's specifications. The individually purified ORFs were subjected to a 3′A addition reaction and subsequently cloned into the vector pCR8/GW/TOPO (Cat. No. K250020, Thermo Fisher Scientific). DH5α cells were transformed, and the positive colonies were selected in semisolid Luria–Bertani (LB) media supplemented with spectinomycin (100 μg/ml). Finally, the ORFs were recombined in-frame to the N-terminal HIS-tag of the pDEST17 expression vector (Cat. No. 11803-012, Thermo Fisher Scientific), using the Gateway Technology (Cat. No. 11791020, Thermo Fisher Scientific). The transformed DH5α cells were selected using ampicillin (100 μg/ml). The correct cloning of the ORFs was confirmed by enzymatic digestion and sequencing.

Bioinformatic Analysis

The pDEST17-*BoCCDs* plasmids were sequenced (Macrogene, Seoul, South Korea), and the identity of the sequences was confirmed by comparing them with the NCBI sequence database, using the Basic Local Alignment Search Tools (BLASTn and BLASTp)[1]. The identity and similarity scores of the sequenced coding regions of all BoCCDs with their homologs reported in the GeneBank from the transcriptome were estimated using the Clustal W program with the default settings[2]. Amino acid sequences were obtained using the translate tool of Expasy server[3]. Multiple alignments of BoCCD1 and BoCCD4 sequences were performed using the Clustal W algorithm and the DNAMAN program. The potential subcellular location of the BoCCDs were predicted using the Ipsort[4], ProtComp 9.0[5], and PREDATOR[6] programs.

Phylogenetic Analysis

The phylogenetic tree was constructed using the amino acid sequences of all BoCCD1 and BoCCD4 proteins, and other functionally characterized CCD1 and CCD4 protein sequences from different plants species. Most of the sequences were retrieved from the GeneBank and some from Sol Genomics Network databases[7]. The phylogenetic tree was constructed using the Maximum likelihood method based on Jones-Taylor-Thornton (JTT) substitution model and Gamma distributed with Invariant sites (G + I), using 1000 bootstrap through MEGA X software (Kumar et al., 2018). The gaps and missing data were partially eliminated. The amino acid sequences were aligned using CLUSTAL W with the default parameter on MEGAX. The substitution model was estimated using the best-fit substitution model (ML) function included in MEGA X.

3D Modeling and Docking Analysis

The 3D structure of the BoCCD1 and BoCCD4 proteins were modeled with the online software PHYRE2[8] using a profile-profile alignment algorithm against the fold library of PHYRE2. Sequence alignment was performed using the online version of CLUSTALW. For docking analysis, Fe^{+2} coordinates in each CCD model complex were selected to define the center of a 25 Å cubic grid box for the search for the geometry of the lycopene docking. Lycopene ligand (zinc8214943) was selected from the Zinc database[9]. The initial geometry was optimized by minimizing energy with the Yasara software[10] and eliminating hydrogens with the Phenix program. As the final model of each complex, the one with the lowest protein-ligand affinity free energy was selected. Structures and docking calculations were obtained using AutoDock Vina. Finally, the analysis of the structures and the visualization were carried out with UCSF Chimera 1.11, and Discovery studio.

SDS-PAGE and Protein Blot Analysis

Five milliliters of the cultures grown at 37°C for 20 h at 175 rpm and induced with IPTG at a final concentration of 0.5 mM were pelleted and solubilized with 8 M urea for 10 min at 96°C. The amount of total protein was equalized with urea 8 M based on culture optical density OD_{600}. Sixteen μl of the solubilized sample were mixed with 4 μl of 5X loading buffer and boiled for 6 min at 96°C. Samples were loaded on 10% SDS/PAGE gels and run in a Mini-PROTEAN Tetra Cell chamber. Protein bands were stained with comasie blue or transferred to a nitrocellulose membrane. Protein blot analysis was performed following the procedure described by Nic-Can et al. (2013), using the anti-His tag (6E6) monoclonal antibody and the HRP conjugate as a secondary antibody. The signal was detected using the commercial ECL chemiluminescence solution.

In vivo Expression of BoCCD1 and CCD4 Proteins

BL21 (DE3) competent *E. coli* cells were cotransformed with the plasmid pACCRT-EIB that produces lycopene and pDEST17-*BoCCDs* plasmids. The negative controls were BL21 (DE3) cells containing the empty pDEST17 vector. Double recombinant cells were selected in semi-solid LB medium with ampicillin (100 μg/ml) and chloramphenicol (34 μg/ml). Two positive colonies were placed in 6 ml of liquid LB medium with the same antibiotics at the same concentrations and were incubated

[1]https://www.ncbi.nlm.nih.gov/

[2]https://www.ebi.ac.uk/Tools/msa/clustalo/

[3]https://web.expasy.org/translate/

[4]http://ipsort.hgc.jp/

[5]http://linux1.softberry.com/all.htm

[6]https://urgi.versailles.inra.fr/predotar/

[7]https://solgenomics.net/

[8]http://www.sbg.bio.ic.ac.uk/phyre2/html/page.cgi?id=index

[9]http://zinc.docking.org/

[10]http://www.yasara.org/minimizationserver.htm

overnight at 37°C at 200 rpm. Two milliliters of the overnight cultures were transferred into 100 ml of LB liquid medium with the same antibiotics at the same concentrations until they reached an OD_{600} of 0.4-0.5. Protein expression was induced with IPTG at a final concentration of 0.5 mM. Cultures were incubated in dark conditions at 37°C for 20 h at 175 rpm. Cells were pelleted in two tubes of 50 ml at 6000 g (Universal 32 R; Hettich zentrifugen, Tuttlingen, Germany) for 20 min at 4°C. For pigment extraction, the pellets were resuspended in 1.7 ml of cold 2:1 chloroform/methanol, centrifuged at 14000 rpm for 20 minutes at 4°C, and the organic phase was placed in a new tube. The organic phase was filtered (Durapore PVDP, 13 mm diameter, 0.22 μm of size, Millipore) and subsequently dried under a stream of nitrogen. All steps were performed on ice and under dark conditions. Each sample was stored at −80°C until it was analyzed by mass spectrometry to identify the compounds.

In vitro Functional Assay of BoCCD1 and BoCCD4 Proteins

For *in vitro* analysis the *PDEST17-BoCCDs* constructs were transformed into BL21 (DE3) cells. One milliliter of the overnight cultures was transferred into 50 ml of LB liquid medium with half-strength of the antibiotic ampicillin (50 μg/ml), incubated at 37°C at 200 rpm until reaching an OD_{600} of 0.4-0.5. Protein expression was induced with IPTG at a final concentration of 0.5 mM, at 25°C for 20 h at 175 rpm. The crude extracts and the *in vitro* reactions were prepared as described by Vogel et al. (2008) with slight modifications. Fifty milliliters of the growth pellet were resuspended in 4.5 ml of cold phosphate-buffered saline (PBS) with 1 mM of phenylmethylsulfonyl fluoride. Samples were sonicated four times for 30 s at máximum power on ice, and then 0.5 ml of 20% Triton X-100 in PBS buffer was added. The cell slurry was shaken for 20 min on ice and then centrifuged at 12,000 × g for 20 min at 4°C. The soluble fractions were filtered through a 0.22 μm filter, and then glycerol was added to a final concentration of 20%; the lysate was aliquoted and stored at −20°C. To obtain an aqueous solution for the enzymatic reactions, lycopene standard was prepared by adding 1% of Tween 40 in chloroform. The solution was vigorously vortexed, and 60 ppm aliquots were removed for each reaction. The aliquots were dried under a stream of nitrogen. The *in vitro* reaction was carried out in a final volume of 1 ml containing 50 mM NaH_2PO_4 (pH 7.2), 300 mM NaCl, 5 μM $FeSO_4$, 5 Mm ascorbic acid, and 200 μl of the total protein extract. The reactions were incubated for 20 h at 30°C and stopped by adding chloroform/methanol (3:1). Compounds were extracted as previously described.

LC-ESI-QTOF-MS/MS

The mass spectra of the different compounds were obtained on a QTOF Xevo G2-S (Waters Corporation, Manchester, UK) with an electrospray source in positive ion mode. Bacterial pellets were resuspended in 100 μl of dichloromethane. Fifty microliters were diluted in 950 μl of isopropanol: acetonitrile (1:1.375). The samples were injected into the mass spectrometer by direct infusion, and the sample infusion flow rate was 15 μL/min. The conditions of the ionization were optimized as follows: the source temperature and desolvation temperatures were 100°C and 400°C, respectively. The desolvation and cone gas flow were 800 and 50 L/h, respectively. The cone and capillary voltages were 30 and 4,000 V, respectively. For the MS/MS acquisition, the collision energies were 6 and 15 eV. The acquisition mass range was 50 to 1500 *m/z* with a scan time of 1 s and resolution mode.

FTIR Analysis

The functional groups of the oxidative cleavage products of the *in vitro* and *in vivo* assay of BoCCD1 and BoCCD4 protein expression were determined by Fourier transform infrared (FTIR) absorption spectra (range 3600 to 500 cm^{-1}) using a TENSOR II FTIR Spectrometer (Bruker Corporation) with default settings. For the analyzes, 20 μl of the dissolved bacterial extracts were mixed with 20 μl of Milli-Q water. Standard lycopene was used as a control. Two microliters were used for each analysis.

RESULTS

Sequence Analysis of BoCCD1 and BoCCD4 Proteins Reveals They Preserve the Active Site Structure of the Carotene Cleavage Dioxygenases

The full-length cDNA of *BoCCD1-1*, *BoCCD1-3*, *BoCCD1-4*, *BoCCD4-1*, *BoCCD4-2*, *BoCCD4-3*, and *BoCCD4-4* genes were amplified from *B. orellana* mRNA using specific primers (**Supplementary Table 1**). Fragments with the expected sizes were cloned and sequenced (**Supplementary Figure 1**). The cloned genes were sequenced, and the identities of all isolated *BoCCD*s genes were confirmed by comparing them with their homologous sequences obtained from the transcriptome and reported in the Genebank (Cárdenas-Conejo et al., 2015). Sequence analysis of the members of the *BoCCD1* family showed that *BoCCD1-1* contains an open frame of 1,629 nucleotides (nt) and encodes a predicted protein of 542 amino acids (aa); *BoCCD1-3* contains 1,644 nt and encodes a protein of 547 aa, and *BoCCD1-4* contains 1,515 nt and encodes a protein of 504 aa. The analysis of the members of the *BoCCD4* family showed that *BoCCD4-1* contains 1,800 nt and encodes a predicted protein of 599 aa, *BoCCD4-2* contains 1,749 nt and encodes a protein of 582 aa, *BoCCD4-3* contains 1, 773 nt and encodes a protein of 590 aa and finally, *BoCCD4-4* contains 1, 863 nt and encodes a protein of 620 aa (**Supplementary Table 2**). The open reading frame of all the isolated *BoCCDs* showed changes at the nucleotide level in their sequence that generated changes at the amino acid level without altering the size of their open reading frame. Only BoCCD4-3 presented the insertion of a proline in position 72 (**Supplementary Figure 2**). Additionally, we found that the missing nucleotide in the *BoCCD4-3* sequence reported in the Genebank (accession number: KT359024.1) corresponds to an adenine, and the complete codon encodes a lysine (**Supplementary Figure 2**). Finally, the level of homology between the isolated genes and

those reported in the transcriptome were > 98%, and 97% at the nucleotide and amino acid levels, respectively (**Supplementary Table 2**), indicating that the isolated *BoCCD1* and *BoCCD4* genes are those previously reported.

To elucidate the phylogenetic relationship of the isolated BoCCD1 and BoCCD4 proteins, a phylogenetic analysis of the deduced BoCCD protein sequences with other functionally characterized CCD1 and CCD4 proteins from several plant species was performed. The protein sequence of *Synechocystis* sp. apocarotenoid cleavage oxygenase (ACO) was used as an external group (**Supplementary Figure 3**). The results show the formation of two major clades corresponding to the CCD1 and CCD4 subfamilies of enzymes. As expected, the results confirmed that the isolated BoCCD1-1, BoCCD1-3, and BoCCD1-4 belong to the CCD1 subfamily; and BoCD4-1, BoCCD4-2, BoCCD4-3, and BoCCD4-4 belong to the CCD4 subfamily (**Supplementary Figure 3**). **Supplementary Figure 3** also shows that the BoLCD (Accession No. CAD33263.1) published by Bouvier et al. (2003) is grouped with the CCD4 enzymes of *Crocus sativus* and is located outside the subgroup formed by the BoCCD4 enzymes, as previously reported (Cárdenas-Conejo et al., 2015). Interestingly, the BoCCD4 enzymes are closely related to citrus CCD4 enzymes, determinants of the reddish-orange color of the flavedo of citrus and mandarins fruits through the biosynthesis of the C_{30} apocarotenoid β-citraulin by cleaving zeaxanthin and β-cryptoxanthin at the 7/8 or 7′/8′ positions (Ma et al., 2013; Rodrigo et al., 2013; Zheng et al., 2019). Citrus CCD4b was also able to cleave lycopene at 7/8 and 5/6 positions producing the C_{30} apocarotenoid apo-8′-lycopenal and the C_{27} apocarotenoid apo-10′-lycopenal, respectively. CCD4b can also cleave apo-8′-lycopenal and apo-10′-lycopenal at 5/6 positions producing C_{22} and C_{19} dialdehydes, respectively (Rodrigo et al., 2013). BoCCD1 proteins formed two close subgroups; in the first subgroup, BoCCD1-1 showed to be more closely related to the BoCCD1 described by Rodríguez-Ávila et al. (2011b) and obtained from the same accession of *B. orellana*; in the second subgroup BoCCD1-3 and BoCCD1-4 were grouped together (**Supplementary Figure 3**; Cárdenas-Conejo et al., 2015; Carballo-Uicab et al., 2019).

To determine if the seven isolated BoCCDs conserved the active site residues necessary for dioxygenase activity, we performed an alignment using the predicted amino acid sequence of the isolated BoCCDs and the sequences of the CCD1 and CCD4 enzymes of *Arabidopsis thaliana* and *C. sativus* functionally characterized (**Supplementary Figure 4**). The alignment analysis revealed that the seven BoCCD contain four highly conserved histidine residues, as a Fe^{+}-ligating cofactor, characteristic of the active site, two conserved glutamic acid residues, and one semi-conserved aspartate residue involved in fixing the iron-linked histidine residues (Harrison and Bugg, 2014; Daruwalla and Kiser, 2020; **Supplementary Figure 4**). According to the consensus of protein subcellular localization prediction programs, all BoCCD4 proteins contain the characteristic chloroplast transit peptide in the N-terminal region. In contrast, all BoCCD1 proteins appear to be cytosolic (Cárdenas-Conejo et al., 2015; Carballo-Uicab et al., 2019; **Supplementary Table 3**). Overall, these results suggest that the isolated *BoCCD1-1*, *BoCCD1-3*, *BoCCD1-4*, *BoCCD4-1*, *BoCCD4-2*, *BoCCD4-3,* and *BoCCD4-4* genes encode functional enzymes that have conserved their role as carotene cleavage dioxygenases in carotenoid metabolism.

BoCCD1 and BoCCD4 Enzymes Are Involved in Bixin Aldehyde Biosynthesis

To determine the participation of the BoCCDs in the first step of bixin biosynthesis, the *in vivo* lycopene cleavage activity of BoCCD1-1, BoCCD1-3, BoCCD1-4, BoCCD4-1, BoCCD4-2, BoCCD4-3, and BoCCD4-4 was analyzed. The coding regions of the seven *BoCCDs* were recombined, producing a fused N-terminal HIS-tag protein in the pDEST17 expression vector, providing an IPTG-inducible expression in *E. coli*. The empty pDEST17 vector was used as a negative control. SDS/PAGE analysis showed that BoCCD1-1, BoCCD1-3, BoCCD1-4, BoCCD4-1, BoCCD4-2, BoCCD4-3, and BoCCD4-4 fusion proteins were expressed with the addition of IPTG, showing an apparent molecular mass of 60.97 kDa, 61.98 kDa, 57.42 kDa, 67.99 kDa, 65.88 kDa, 66.13 kDa, and 69.74 kDa, respectively (**Supplementary Figures 5A,B**). Protein blot analysis using anti-His antibody confirmed the expression of the seven BoCCDs (**Supplementary Figure 5C**).

The recombinant proteins were expressed in *E. coli* cells genetically engineered to produce all-*trans* lycopene by plasmid pACCRT-EIB (Misawa et al., 1995). The pACCRT-EIB *E. coli* cells display a red coloration given by the accumulation of lycopene (**Figure 2**). **Supplementary Figure 6** shows the mass spectra obtained by LC-ESI-QTOF-MS/MS of the lycopene standard (STD) and the lycopene produced by the pACCRT-EIB *E. coli* cells, respectively. In both samples, the molecular ion of lycopene in the MS/MS spectra is observed at *m/z* 536.3, along with its M + 1 ion at *m/z* 537.3 and M + 2 ion at *m/z* 538.3 (**Supplementary Figure 6A** and **Table 1**). The lycopene molecular ion *m/z* 536.3 at 15 eV of positive ion collision is dissociated in two main ion fragments corresponding to *m/z* 444.3 $[M - 92]^{+}$ and *m/z* 467.3 $[M - 69]^{+}$ (**Supplementary Figures 6B–D**; van Breemen, 2001; Rivera et al., 2014; Amorim et al., 2018). The mass spectrum of the lycopene standard and that of lycopene extracted from pACCRT-EIB *E. coli* cells showed a similar ion pattern that corresponds to lycopene (**Supplementary Figure 6** and **Table 1**).

The color of the pellet generated after the induction of the expression of BoCCD1 and BoCCD4 proteins in pACCRT-EIB *E. coli* cells showed that only BoCCD1-1, BoCCD1-3, and BoCCD4-3 generated an evident decoloration compared to the control (empty vector). Pellet decoloration was also observed, although to a lesser extent (pellets presented a pink tone), in lycopene-producing *E. coli* cells that expressed BoCCD1-4, BoCCD4-1, and BoCCD4-2 proteins. The pellet color of *E. coli* cells that expressed BoCCD4-4 was similar to the control; both presented a red/orange color (**Figure 2**).

Pellet decoloration intensity is an indication that lycopene is being cleaved by the activity of the expressed CCD proteins, generating new colorless compounds (Carballo-Uicab et al., 2019; Gómez-Gómez et al., 2020; Nawade et al., 2020). Mass

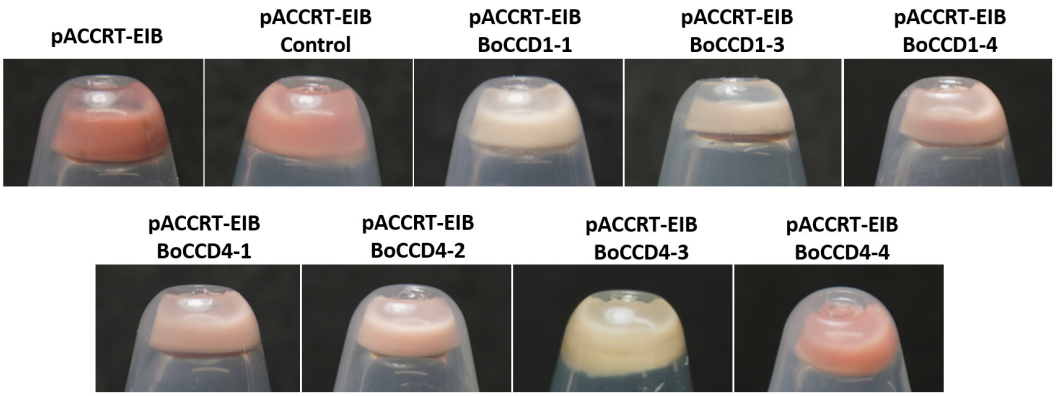


FIGURE 2 | Discoloration of pCCART-EIB *E. coli* cells that accumulate lycopene by the expression of BoCCD1-1, BoCCD1-3, BoCCD1-4, BoCCD4-1, BoCCD4-2, BoCCD4-3, and BoCCD4-4 proteins. The empty pDEST17 vector was used as control.

spectrometry analysis by direct injection of lycopene-producing *E. coli cell* extracts expressing the BoCCD1 and BoCCD4 proteins revealed a decrease in the lycopene peak (*m/z* 536.3) compared to the control, even in BoCCD4-4, which has a red orange color (**Supplementary Figure 7**). These results suggest a cleavage activity on lycopene, which may be related to the pellet color observed after the expression of the BoCCD proteins (**Figure 2**).

A previous study indicated that BoCCD1-1 and BoCCD4-3 enzymes perform lycopene cleavage at 5,6/5′,6′ bonds to produce bixin aldehyde (*m/z* 348), identified in its protonated form *m/z* 349 [M + 1] (Carballo-Uicab et al., 2019). To determine if the rest of the isolated BoCCD1 and BoCCD4 proteins are capable of producing bixin aldehyde, we performed a MS and MS/MS analysis directed at the *m/z* 349 peak. LC-ESI-QTOF-MS/MS analyzes showed that BoCCD1-1 and BoCCD4-3 produce a corresponding peak of *m/z* 349.1 (**Supplementary Figure 8A**), and revealed that BoCCD1-4, BoCCD4-1, and BoCCD4-2 produced the same peak of *m/z* 349.1 (**Figure 3A**). Unfortunately, there is still no commercial standard for bixin aldehyde. Therefore, to establish that the *m/z* 349.1 peak corresponds to bixin aldehyde we carried out its ion fragmentation. In all the bacterial extracts, the peak corresponding to *m/z* 349.1 was found, as well as in most the ion fragments of *m/z* 145, *m/z* 175 and *m/z* 265 generated by the ion fragmentation of bixin aldehyde according to the database of The Metabolomics Innovation Center (TMIC). We also found other ion fragments corresponding to bixin aldehyde according to the TMIC database (**Figure 3A**, **Table 1**, and **Supplementary Figures 8A, 9**). Ion fragments appear in very low intensities, even increasing the positive collision energy (> 20eV). It has been reported that diapocarotenoids or dyaldehydes produced from the oxidative cleavage of carotenoids have a low ionization efficiency in mass spectrometry (Mi et al., 2019). In the case of BoCCD1-3 and BoCCD4-4 we found a very small peak corresponding to *m/z* 350 $[M + 2]^+$ and *m/z* 349 $[M + 1]^+$, respectively, that was impossible to fragment (**Supplementary Figure 10**). In all the bacterial extracts the bixin aldehyde peak (*m/z* 349.1) was not prominent when injected directly (**Supplementary Figure 7**). It is also possible that the low intensities are due to the low amount of bixin aldehyde, or its high consumption efficiency (Rubio-Moraga et al., 2008; Mi et al., 2019). Results indicate that the BoCCD1-4, BoCCD4-1, and BoCCD4- 2 enzymes are capable of cleaving lycopene at 5,6/5′,6′ double bonds to produce bixin aldehyde, as BoCCD1-1 and BoCCD4-3 do.

In vitro Analysis Confirms Cleavage Activity of BoCCD1 and BoCCD4 Enzymes Upon Lycopene to Produce Bixin Aldehyde

To confirm the *in vivo* cleavage activity at 5,6, 5′/6′ double bonds of lycopene, we evaluated the *in vitro* enzymatic activity of BoCCD1 and BoCCD4 proteins and their specificity to lycopene using the crude protein extracts and lycopene standard as substrate. The results of mass spectrophotometry confirmed that BoCCD1-1 and BoCCD4-3, as well as BoCCD1-4, BoCCD4-1, and BoCCD4-2, could produce the peak of *m/z* 349.1 corresponding to bixin aldehyde when they were incubated with the lycopene standard (**Figure 3B** and **Supplementary Figure 8B**). Bixin aldehyde was also identified based on its ion fragments (**Figure 3B**, **Table 1**, and **Supplementary Figures 8B, 9**). Interestingly, the *m/z* 349.1 peak was also found in the *in vitro* assay of BoCCD1-3 and BoCCD4-4 (**Figure 3B**). Diapocarotenoids detection is more efficient in the *in vitro* enzymatic assays because of their low quantity and high reactivity (Ruch et al., 2005; Ilg et al., 2014; Mi et al., 2019; Wang et al., 2019). This is probably why bixin aldehyde produced by BoCCD1-3 and BoCCD4-4 was better detected by *in vitro* assays.

In *Escherichia coli* Cells Bixin Aldehyde Is Oxidized to Norbixin

During the direct injection of the extracts obtained from the *in vivo* assay, we observed a prominent peak of *m/z* 380.2. This was observed in the extracts obtained from the *in vivo* expression of the BoCCD1 and BoCCD4 proteins (**Supplementary Figure 7**). This led us to suggest that *in vivo*,

TABLE 1 | ESI-MS and MS/MS fragments of lycopene, bixin aldehyde, and norbixin reported and found in bacterial extracts that express BoCCD1 and BoCCD4 proteins.

Compound	MS (*m/z*)	MS/MS (*m/z*)
Lycopene	536 $[M]^+$	536 $[M]^+$
	537 $[M + 1]^+$	647 $[M - 69]^+$
	538 $[M + 2]^+$	444 $[M - 92]^+$
BoCCDs	536	536
	537	647
	538	444
Bixin aldehyde	349 $[M + 1]^+$	349 $[M + 1]^+$
		145
		174
		227
		265
		267
		319
BoCCDs	349 $[M + 1]^+$	349 $[M + 1]^+$
		145 $[M + 203]^+$
		175 $[M + 1 - 174]^+$
		227 $[M - 121]^+$
		228 $[M + 1 - 121]^+$
		264 $[M - 84]^+$
		265 $[M + 1 - 84]^+$
		266 $[M - 82]^+$
		319 $[M - 29]^+$
Norbixin	380 $[M]^+$	381 $[M + 1]^+$
		121
		197
		253
		267
		281
		309
BoCCDs	380 $[M]^+$	380 $[M]^+$
		120 $[M + 1 - 261]^+$
		196 $[M - 184]^+$
		252 $[M - 128]^+$
		266 $[M + 1 - 115]^+$
		280 $[M - 100]^+$
		308 $[M - 72]^+$

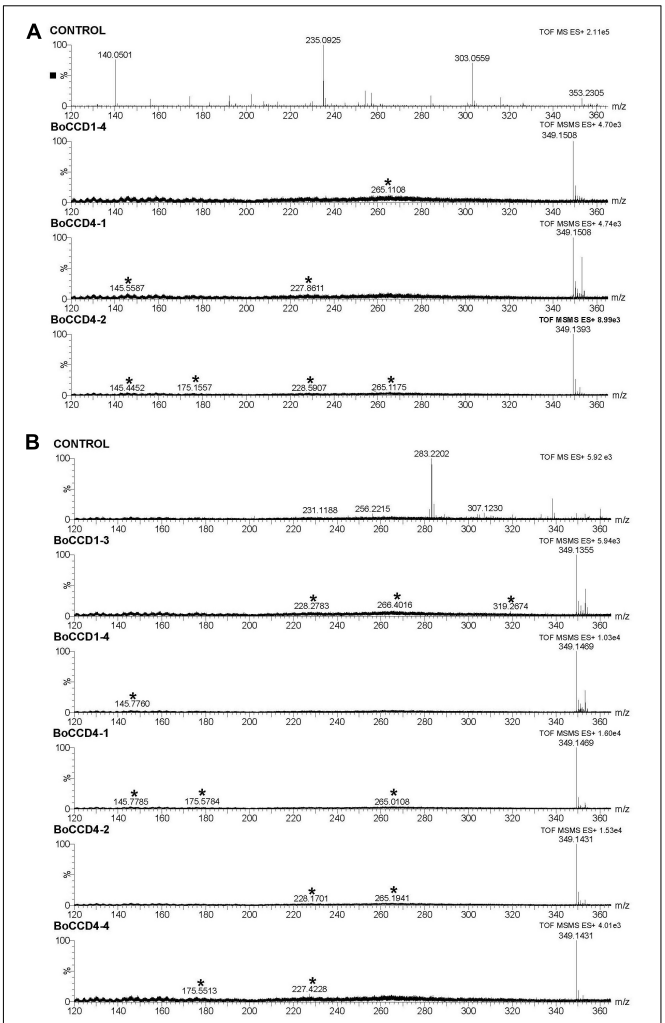


FIGURE 3 | Bixin aldehyde MS/MS spectra (*m/z* 349.1) of the extracts obtained from the *in vivo* and *in vitro* expression of the BoCCD1 and BoCCD4 proteins. **(A)** MS/MS spectra of the *in vivo* expression of BoCCD1-4, BoCCD4-1 and CCD4-2. **(B)** MS/MS spectra of the *in vitro* expression of BoCCD1-3, BoCCD1-4, BoCCD4-1, BoCCD4-2, and CCD4-4. The empty pDEST17 vector was used as a control.

the aldehyde groups of bixin aldehyde are oxidized to produce norbixin whose exact mass is 380.198. To test our hypothesis, we directed the search toward the peak of *m/z* 380.2, and we analyzed whether its ionic fragmentation corresponded to norbixin (**Table 1**). Mass spectrophotometry results showed that only in the bacterial extracts obtained from the expression of BoCCD1-4, BoCCD4-1, BoCCD4-2, and BoCCD4-4 proteins, the peak of *m/z* 380.2 presented the ion fragmentation pattern without a positive charge that corresponds to norbixin according to the TMIC database, indicating the previous formation of bixin aldehyde (**Figure 4**, **Table 1**, and **Supplementary Figure 11**). Furthermore, the intensity of *m/z* 380.2 decreased after its fragmentation (**Supplementary Figure 12**). Although the *m/z* 380.2 peak was found in the control, it did not fragment, suggesting that it could be another compound with the same mass. The results of this analysis indicate that *in vivo*, bixin aldehyde produced by BoCCD1-4, BoCCD4-1, BoCCD4-2, and BoCCD4-4 genes is oxidized to a more stable compound— norbixin— the second product of bixin biosynthesis pathway.

The *in vivo* formation of norbixin (which contains the carboxyl acid functional group) and *in vitro* formation of bixin aldehyde (which contains the aldehyde functional group) was corroborated by FTIR spectroscopy. The extracts of BoCCD1-4 and BoCCD4-1 were analyzed in a spectrum range from 500 to 3,600 cm^{-1} (**Figure 5**). The FTIR spectra within BoCCD1-4 and BoCCD4-1 in the *in vivo* and *in vitro* assays were very similar. In the initial time (t_0) of the *in vitro* enzymatic reaction, the vibration signal of the C = O group at 1727 cm^{-1} corresponding to the aldehyde group was observed. Such signal is observed 20 h

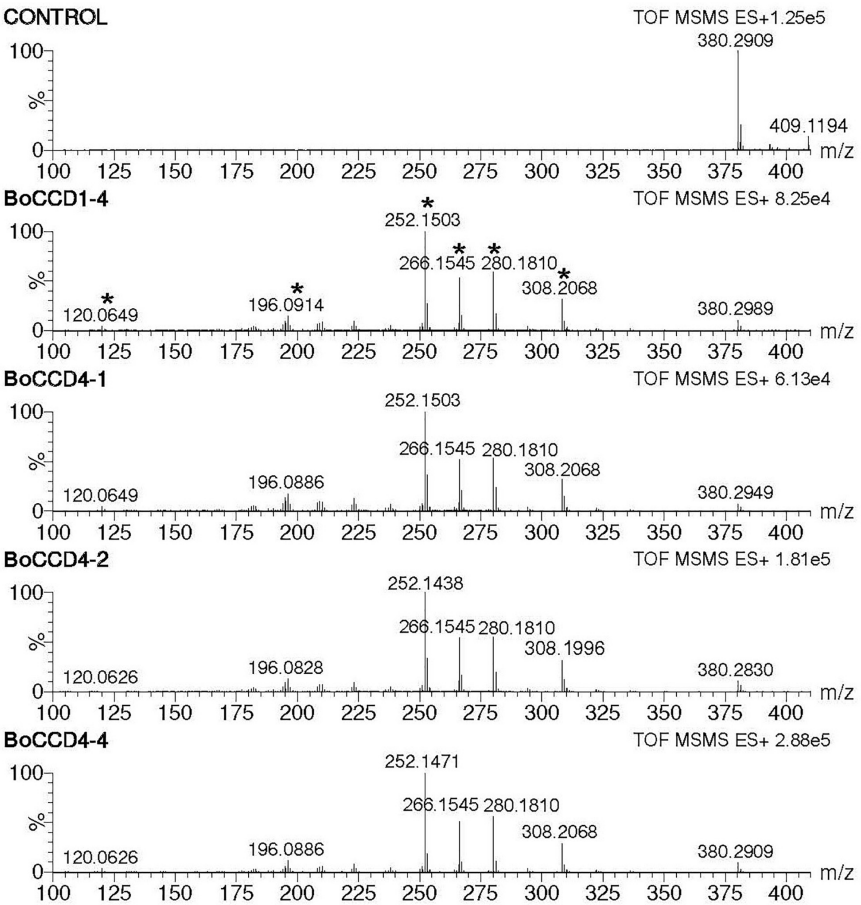


FIGURE 4 | Norbixin MS/MS spectra (*m/z* 380.2) of the extracts obtained from the *in vivo* expression of BoCCD1-4, BoCCD4-1, BoCCD4-2, and CCD4-4. The empty pDEST17 vector was used as a control.

later, at the end of the *in vitro* reaction (t_f). The vibration signal at 2250 cm^{-1} is also observed in both t_0 and t_f, attributed to the different products generated by the changing bonds of the CHO (*), CH (**) and CH3 (***) groups until bixin aldehyde is produced from lycopene. The C triple bond signal is also observed at 2100 to 2250 cm^{-1} (**Figures 5A,B,D,E**; Carballo-Uicab et al., 2019). On the other hand, in the *in vivo* assay of BoCCD1-4 and BoCCD4-4, the vibration signal of the C = O group at 1727 cm^{-1} corresponding to the aldehyde group, and of the CHO (*), CH (**) and CH3 (***) groups between 2100 cm^{-1} to 2250 cm^{-1} is almost imperceptible. Instead, the vibration signal of the C-O-H group at 935 cm^{-1}, which corresponds to the carboxyl acid group is observed. We also observed the vibration signals of the C-O at 1294 cm^{-1}, C = O at 1714 cm^{-1}, and the OH stretch between 3,400 to 2,400 cm^{-1}, characteristic of compounds with carboxyl acid groups (**Figures 5C,F**). The FTIR spectra of the *in vivo* and *in vitro* assays corroborate the mass spectrometry results. It also suggests that the formation of bixin aldehyde occurs by a rapid enzymatic reaction because the C = O vibration signal is observed since t_0 in the *in vitro* enzymatic reaction (**Figures 5A,D**). It is also possible that the *in vivo* formation of norbixin from bixin aldehyde occurs continuously, which could explain why the bixin aldehyde signal was almost imperceptible in these samples.

Structural 3D Modeling and Docking Analysis of BoCCD1 and BoCCD4 Proteins

To identify the structural 3D domains of the BoCCD1 and BoCCD4 proteins, structural modeling of sequence homology was performed. Sequence alignment software predicted VP14 — a 9-Z-epoxycarotenoid dioxygenase from maize (PBD accession: 3npeA) — as the best template to model the different BoCCDs proteins (Schwartz et al., 1997; Messing et al., 2010). In all cases, we found reasonable identity and confidence percentages between VP14 and the BoCCD1 and BoCCD4 proteins, indicating good coverage between the residues (**Supplementary Table 4**). The structural models obtained for BoCCD1-1, BoCCD1-3, BoCC1-4, BoCCD4-1, BoCCD4-2, BoCCD4-3, and BoCCD4-4 proteins using the VP14 protein as a template showed that all models obeyed the basic conformation of the CCD enzymes (**Figure 6** and **Supplementary Figure 13**). In general, the seven CCDs contain the two main functional

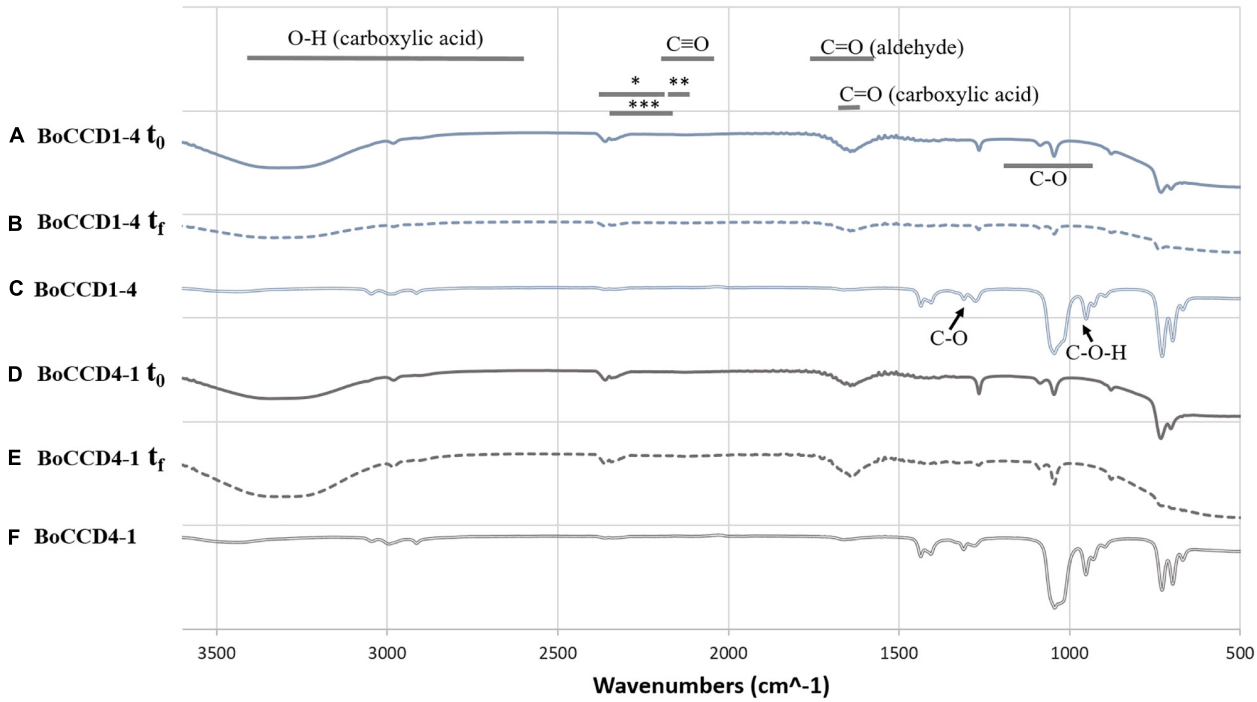


FIGURE 5 | FTIR spectrum of the *in vitro* and *in vivo* extracts of BoCCD1-4 and BoCCD4-1. FTIR spectrum of the initial time (t_0) of the *in vitro* enzymatic reaction of **(A)** BoCCD1-4 and **(D)** BoCCD4-1. FTIR spectrum of the end time (t_f) of the *in vitro* enzymatic reaction of **(B)** BoCCD1-4 and **(E)** BoCCD4-1. FTIR spectrum of the *in vivo* assay of **(C)** BoCCD1-4 and **(F)** BoCCD4-1.

domains exhibiting the seven characteristic blades that form the β-propeller domain, in addition to the three α-helices (α1, α2, and α3) that form the α helical domain located in the upper part of the β-propeller. The β-propeller domain is conserved in the CCD protein family, it forms a toroidal structure with a long tunnel in the central axis surrounded by the seven blades. The four histidine residues coordinated with Fe^{2+} necessary for dioxygenase activity are in the center of the structure (**Figure 6** and **Supplementary Figure 13**; Kloer and Schulz, 2006; Messing et al., 2010; Harrison and Bugg, 2014; Priya et al., 2017; Vega-Teijido et al., 2019; Daruwalla and Kiser, 2020). In VP14, the parallel helices α1 and α3 form a hydrophobic patch that is associated with the interaction of this protein with the thylakoid membranes (Tan et al., 2001; Messing et al., 2010; Daruwalla and Kiser, 2020).

To gain a better understanding of the molecular mechanism of BoCCD1 and BoCCD4, a docking analysis was directed toward the center of the BoCCD1 and BoCCD4 proteins where Fe^{2+} is located, using lycopene as a ligand. Results showed that all BoCCD1 and BoCCD4 proteins were able to form molecular interactions with lycopene. The spatial location of the iron (II) ion coordinated with the four conserved histidine residues were very similar in all BoCCDs. The only significant differences were found in the spatial orientation of lycopene, which probably influenced the differential residues interacting with it (**Figure 6** and **Supplementary Figure 13**). The lowest RMSD values for all BoCCDs indicate the accuracy of the docking geometries and suggest the high conservation of the binding site structure of all BoCCD1 and BoCCD4 proteins (**Supplementary Table 5**). Differences were found in their binding affinities estimated by ΔG° values and dissociation constants (Kd). The range of ΔG° values is between −5 and −7 in all BoCCDS, and it is reasonable assuming that the same ligand was used. The best affinity value is for CCD4-2, followed by BoCCD1-1, BoCCD4-1, BoCCD4-4, BoCCD1-3, BoCCD1-4, and finally BoCCD4-3 (**Supplementary Table 5**). The low Kd values also suggest that CCD4-2 and CCD1-1 may be playing a more relevant role in the formation of bixin aldehyde from lycopene. The results found in the *in silico* analysis indicate that all BoCCD1 and BoCCD4 proteins contain the structural and functional characteristics of carotenoid cleavage dioxygenases and confirm that lycopene can be used as a substrate to produce bixin aldehyde.

DISCUSSION

Bixin Aldehyde Biosynthesis Is Performed by Different Members of the BoCCD1 and BoCCD4 Subfamily of Enzymes

Bixin biosynthesis pathway begins with the biosynthesis of bixin aldehyde through the symmetric cleavage of lycopene at 5, 6/5′,6′ double bonds by a lycopene dioxygenase (BoLCD) (Jako et al., 2002; Bouvier et al., 2003; Carballo-Uicab et al., 2019). The transcriptomes of leaf and mature/immature seed of a local

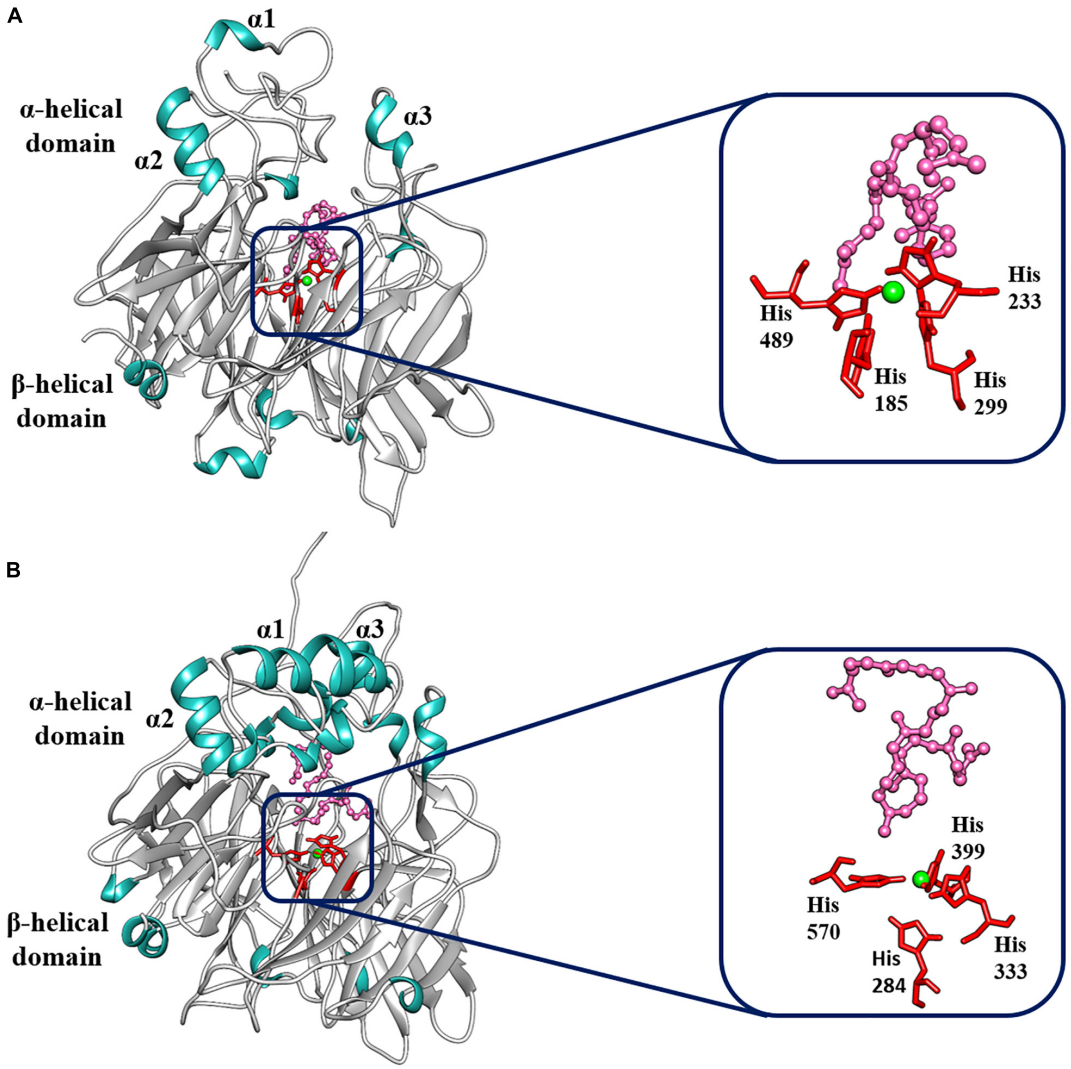


FIGURE 6 | Best-fit 3D model and docking analysis of **(A)** BoCCD1-4 and **(B)** BoCCD4-2 proteins. The α-helix (α1, α2, and α3) and the β-propeller domains are indicated on the models. Green shows the Fe^{2+} ion coordinated by the four histidine residues. The four histidine residues are shown in red, and the lycopene molecule is shown in pink.

Yucatecan variety of *B. orellana* allowed the identification of a new set of eight candidate *BoCCD* genes potentially involved in bixin biosynthesis. These genes were grouped into two subfamilies, four belong to the *BoCCD1* subfamily: *BoCCD1-1*, *BoCCD1-2*, *BoCCD1-3*, and *BoCCD1-4*; and four to the *BoCCD4* subfamily: *BoCCD4-1*, *BoCCD4-2*, *BoCCD4-3*, and *BoCCD4-4* (Cárdenas-Conejo et al., 2015). Until now, the biosynthesis of bixin aldehyde from lycopene had only been demonstrated for BoCCD1-1 and BoCCD4-3. However, based on previous research results, we hypothesized that the rest of the BoCCD1 and BoCCD4 proteins could have a possible role in bixin biosynthesis (Cárdenas-Conejo et al., 2015; Carballo-Uicab et al., 2019). Therefore, the main objective of this work was to determine if the proteins encoded by *BoCCD1-3*, *BoCCD1-4*, *BoCCD4-1*, *BoCCD4-2*, and *BoCCD4-4* genes (in addition to *BoCCD1-1* and *BoCCD4-3*) could produce bixin aldehyde.

The results of the analysis of the deduced amino acid sequence of the isolated *BoCCD1* (*BoCCD1-1*, *BoCCD1-3*, *BoCCD1-4*), and *BoCCD4* (*BoCCD4-1*, *BoCCD4-2*, *BoCCD4-3*, and *BoCCD4-4*) genes indicated that they were the same genes identified in the *B. orellana* transcriptome (similarity > 98%) (**Supplementary Table 2**; Cárdenas-Conejo et al., 2015). Sequence alignment also showed that they all contain the four conserved histidine residues of the active site as well as the glutamate and aspartate residues necessary for dioxygenase activity (**Supplementary Figure 4**; Harrison and Bugg, 2014; Daruwalla and Kiser, 2020). Additionally, 3D modeling showed that all BoCCD1 and BoCCD4 proteins obey the basic conformation of CCD enzymes, contain the hydrophobic α hélix domain predicted for the interaction with plastid membranes and the β-propeller domain with the catalytic site in the center of their structure, indicating that they encode functional carotene cleavage dioxygenases

(**Figure 6** and **Supplementary Figure 13**; Tan et al., 2001; Messing et al., 2010; Daruwalla and Kiser, 2020).

The functional activity of the BoCCD1 and BoCCD4 proteins in the production of bixin aldehyde from lycopene was evaluated using *in vivo* (*E. coli*) and *in vitro* approaches. The results of the mass spectrometry analysis of the *in vivo* and *in vitro* assays showed that the BoCCD1-3, BoCCD1-4, BoCCD4-1, BoCCD4-2, and BoCDD4-4 proteins also cleaved lycopene at the 5,6/5′,6′ double bonds producing bixin aldehyde (*m/z* 349). Identification was based on ion fragmentation (**Figure 3**, **Table 1**, and **Supplementary Figure 9**). These results are consistent with the ΔG° values from the docking analysis, which suggest that all BoCCD1 and BoCCD4 proteins can use lycopene as a substrate (**Supplementary Table 5**). In some extracts from the *in vivo* assay, bixin aldehyde was not detected as a prominent peak. Something similar has been observed in other works carried out in *E. coli*, where the dialdehydes were not detected in appreciable amounts (Schwartz et al., 2001; Vogel et al., 2008). It has been reported that diapocarotenoids or dialdehydes can be detected more efficiently in the *in vitro* assays, because they are generally found in low quantities and are rapidly catabolized (Ruch et al., 2005; Ilg et al., 2014; Mi et al., 2019; Wang et al., 2019).

In *Escherichia coli* Cells, Bixin Aldehyde Can Be Oxidized to Norbixin, the Second Compound of the Bixin Biosynthesis Pathway

Interestingly, we found that in the *in vivo* assay of BoCCD1-4, BoCCD4-1, BoCCD-4-2, and BoCCD4-4, bixin aldehyde was oxidized to norbixin (*m/z* 380), the second product of the bixin biosynthesis pathway (**Figure 4**). Norbixin was also identified by its ion fragmentation (**Supplementary Figure 11** and **Table 1**). One line of evidence that norbixin is being produced by BoCCD1-4, BoCCD4-1, BoCCD4-2, and BoCCD4-4 is the red/orange color of the pellets observed after their expressions. Norbixin has a natural red/orange to yellow color, depending on its concentration (Sharma et al., 2020). We suggest that *E. coli* must contain a non-specific aldehyde dehydrogenase capable of oxidizing bixin aldehyde produced by the BoCCD enzymes to norbixin. In *E. coli* cells, crocetin dialdehyde produced by the cleavage activity of CCD2 at the 7,8/7′,8′ double bonds of zeaxanthin is oxidized to crocetin by an unknown aldehyde dehydrogenase (Frusciante et al., 2014). A similar result was found by Jimenez-Lopez et al. (2016), reinforcing the idea of an *E. coli* aldehyde dehydrogenase capable of oxidizing the cleaved aldehyde product. Moreover, the signal vibration of the aldehyde group of bixin aldehyde and of the carboxyl acid groups of norbixin were detected in the extracts of the *in vivo* and *in vitro* assay of the BoCCDs (**Figure 5**).

To estimate the amount of lycopene cleaved and the amount of bixin aldehyde and norbixin produced, we analyzed the intensity level of the molecular ion of each compound. The values found suggest that in the *in vivo* assay: a) BoCCD1-1, BoCCD4-2, and BoCCD4-4 cleave lycopene more efficiently; b) BoCCD1-1, BoCCD4-3, and BoCCD4-2 produce more bixin aldehyde; and c) bixin aldehyde is more efficiently oxidized to norbixin by BoCCD4-4, followed by BoCCD4-2 (**Supplementary Table 6**). Even norbixin was only detected in the *in vitro* BoCCD4-4 assay (**Supplementary Figure 14**). The high production of bixin aldehyde by BoCCD1-1 and BoCCD4-2 is also related to the low values of the dissociation constant (Kd) toward lycopene of these enzymes (**Supplementary Table 5**). Although in the *in vitro* assay we used equal amounts of synthetic lycopene, we did not observe as obvious changes in coloration as in the *in vivo* assay; for this reason, we considered the values found in the *in vivo* assays, where the activity of BoCCD1 and BoCCD4 is more efficient. The activity of CCD enzymes can be greatly reduced during protein extraction and purification (Rodríguez-Concepción and Welsch, 2020).

BoCCD1 and CCD4 Enzymes Could Act in a Coordinated Way for the Biosynthesis of Bixin and Other Apocarotenoids in *Bixa orellana* Plants

Bixa orellana produces and accumulates bixin mainly in its seeds, but bixin has also been detected in other plant tissues such as roots, leaves, stems, floral buds, and flowers (Mahendranath et al., 2011; Rodríguez-Ávila et al., 2011a,b). We propose that *in planta*, all BoCCD1 and BoCCD4 enzymes participate in bixin biosynthesis in a tissue-specific manner. The presence of families of enzymes that perform the same enzymatic activity in the biosynthesis of carotenoids and apocarotenoids seems to be common, allowing the biosynthesis of these compounds to be tissue-specific and specific to a stage of development (Rubio-Moraga et al., 2008; Walter and Strack, 2011; Cárdenas-Conejo et al., 2015). Expression analyzes of *CCD1* and *CCD4* genes from different plant species show that they individually have tissue-specific expressions and that their expression is often regulated by the accumulation of carotenoids and apocarotenoids (Simkin et al., 2004; Rubio-Moraga et al., 2008, 2014; Rodríguez-Ávila et al., 2011b; Frusciante et al., 2014). *BoCCD1*, orthologous of *BoCCD1-1*, was expressed in root, stem, leaves, floral tissues, and seeds of *B. orellana*, but its expression was only highly induced in immature seeds when pigment production began (Rodríguez-Ávila et al., 2011b; Cárdenas-Conejo et al., 2015). Carballo-Uicab et al. (2019) also found that the expression of BoCCD1-1 and BoCCD4-3 was markedly high in bixin storage cells (BSCs) of the arils of *B. orellana* seeds, coinciding with the accumulation of bixin. In leaves of *B. orellana*, a positive correlation was observed between the content of bixin and the expression of the *BoCCD4-3* and *BoCCD4-2* genes (Faria et al., 2020). This suggests that more than one BoCCD enzyme can carry out bixin biosynthesis in the same tissue. According to the prediction of the subcellular location and our findings, bixin biosynthesis can be performed both in the cytosol and in plastids, probably by independent pathways, but acting in coordination. The prediction of the subcellular location of the *BoALDH* and *BoSABATH* gene families showed that several members can be located in the cytosol or plastids, supporting this proposal (Cárdenas-Conejo et al., 2015). BoCCD1 proteins probably anchor to plastid membranes to gain access to the lycopene, as predictions of the subcellular location indicated that

they are cytosolic (Floss and Walter, 2009; Frusciante et al., 2014). Conversely, BoCCD4 proteins have a plastid target sequence, which would give them direct access to lycopene (Rubio-Moraga et al., 2008; Vranová et al., 2013).

Cell decoloration after the expression of CCD enzymes in *E. coli* cells that produce pigmented carotenoids indicates that they are metabolized to new cleaved colorless compounds (Schwartz et al., 2001; Huang et al., 2009; Nawade et al., 2020). In BoCCD1-1, BoCCD1-3, and BoCCD4-3, we observed a decolored pellet after its expression in lycopene-producing *E. coli* cells. These enzymes are likely to cleave lycopene or its precursors at other bonds, generating mostly volatile compounds that were not detectable by LC-ESI-QTOF-MS/MS. CCD1 enzymes can cleave lycopene at the 5, 6 or 5′, 6′, 9, 10 or 9′, 10′, and 7, 8 or 7′, 8′ bonds, producing the volatile MHO (6-methyl-5-hepten-2-one), pseudoionone, and geranial, respectively (Vogel et al., 2008; Huang et al., 2009; Ilg et al., 2009; Lashbrooke et al., 2013; Nawade et al., 2020). CCD4 enzymes can also produce MHO and cleave lycopene at 7, 8 and 5, 6 bonds to produce apo-8′-lycopenal (C_{30}), and apo-10′-lycopenal (C_{27}), respectively (Lashbrooke et al., 2013; Rodrigo et al., 2013). Moreover, CCD1 and CCD4 can cleave C_{30} and C_{27} apolycopenals at various bonds producing pseudoionone and C_{15}, C_{17}, C_{19}, C_{20}, C_{22}, and C_{25} dialdehydes (Ilg et al., 2009; Rodrigo et al., 2013). CCD1 and CCD4 enzymes can also cleave lycopene precursors such as phytoene and ζ-carotene producing geranylacetone and farnesyl acetone (Rubio-Moraga et al., 2008; Vogel et al., 2008; Lashbrooke et al., 2013; Ilg et al., 2014). The volatile apocarotenoids produced by the CCD1 and CCD4 enzymes are responsible for flavor, aroma, pollinators attraction, and signaling (Floss et al., 2008; Rubio-Moraga et al., 2008; Floss and Walter, 2009; Walter et al., 2010; Brandi et al., 2011). A variety of linear and cyclic apocarotenoids have been found in *B. orellana*, as well as volatile aldehyde, ketones, esters, and carboxylic acids compounds derived from the polyene chain, whose biosynthesis is still unknown (Mercadante et al., 1996, 1997; Galindo-Cuspinera et al., 2002; de Oliveira Júnior et al., 2019). The presence of multiple BoCCD1 and BoCCD4 enzymes in *B. orellana* opens the possibility of CCDs having different substrate specificity, although further studies are needed to confirm this. CCD1 enzymes are considered more promiscuous due to their wide substrate preference, while CCD4 enzymes appear to be more substrate specific (Vogel et al., 2008; Huang et al., 2009; Walter and Strack, 2011).

The elucidation of the enzymes involved in the bixin biosynthesis pathway, as well as their regulatory mechanisms are essential to improve the production of this important apocarotenoid. Although the main objective of this work was to find the BoCCD1 and BoCCD4 enzymes involved in the first step of the bixin biosynthesis pathway, in this work, it was also possible to shed light on the second step of the bixin biosynthesis pathway that had not been evaluated for a long time. *B. orellana* contains several potential members of the ALDH family of enzymes that can carry out the oxidation of bixin aldehyde; thus future work will aim at elucidating the BoALDH enzymes responsible for norbixin biosynthesis. It is also possible that the BoCCD1 and BoCCD4 enzymes carry out bixin biosynthesis in a tissue-specific way, and that they could be involved in other biological processes through the biosynthesis of other apocarotenoids by cleaving different carotenoids. Although further biochemical and molecular studies are needed, the difference in the production of bixin aldehyde and norbixin seems to support our argument. Finally, the elucidation of methyltransferases (BoSABATHS) that catalyze the last step of bixin biosynthesis opens the possibility of scaling the industrial production of bixin in heterologous organisms such as bacteria and plants.

Recent studies carried out on *Nicothiana tabacum* and *Solanum lycopersicum* show plants as potential alternatives for the production of apocarotenoids of commercial interest such as crocin through the expression of a carotenoid-cleaving dioxygenase enzyme from *C. sativus* (CsCCD2) and *B. orellana* (BoCCD4-3), respectively (Martí et al., 2020; Frusciante et al., 2021). Performing metabolic engineering with the BoCCD1s and BoCCD4s identified in this work in plants such as *S. lycopersicum* and *Carica papaya*, which contain different carotenoids in addition to lycopene, would allow evaluating them as possible alternatives for the production of bixin and probably other apocarotenoids, given the presence of different carotenoids in these plants and the ability of CCD1 and CCD4 enzymes to cleave them at different double bond sites.

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: Protein accession numbers (10.6084/m9.figshare.18461612), FTIR data (10.6084/m9.figshare.18480581), and Mass spectrometry data (10.6084/m9.figshare.18482666).

AUTHOR CONTRIBUTIONS

RU-C conducted most of the experiments, conceived and designed the experiments, and drafted the manuscript. MA-E helped in data collection, pigment extraction, and sample preparation. JR-C helped in data collection and establishment of mass spectrometry methodology. AV-C helped in data collection, mass spectrometry, and FTIR spectroscopy analysis. VC-U helped in 3D structural modeling and docking analysis. HS-P helped in the supervision of 3D structural modeling and docking analysis. RR-M coordinated the project, conceived the experiments, and reviewed and edited the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by the Consejo Nacional de Ciencia y Tecnología (CONACYT) (Fronteras de la Ciencia No. 2016-01-1716). Rosa Yazmín Us Camas was financially supported by

Fronteras de la Ciencia and CONACYT postdoctoral grants (CVU No. 375027).

ACKNOWLEDGMENTS

We thank Jonatan Dzib C. and undergraduate students María Fernanda De la Cruz V. and Gabriela del Carmén Vazquez G. for their technical help in gene cloning and protein expression; Miguel Beltran G. for the Autoflex Speed Bruker Mass Spectrometry equipment support; Sergio Peraza S. for his valuable comments on the analysis of mass spectra; and Luis Carlos Rodriguez for their technical help in the project.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2022.829089/full#supplementary-material

REFERENCES

Akshatha, V., Giridhar, P., and Gokare, R. (2011). Food, ethanobotanical and diversified applications of *Bixa orellana* L.: a scope for its improvement through biotechnological mediation. *Ind. J. Fundament. Appl. Life Sci.* 1, 9–31.

Amorim, A. G. N., Souza, J. M. T., Santos, R. C., Gullón, B., Oliveira, A., Santos, L. F. A., et al. (2018). HPLC-DAD, ESI–MS/MS, and NMR of lycopene isolated from *P. guajava* L. and its biotechnological applications. *Eur. J. Lipid Sci. Technol.* 120:1700330. doi: 10.1002/ejlt.201700330

Auldridge, M. E., McCarty, D. R., and Klee, H. J. (2006). Plant carotenoid cleavage oxygenases and their apocarotenoid products. *Curr. Opin. Plant Biol.* 9, 315–321. doi: 10.1016/j.pbi.2006.03.005

Bouvier, F., Dogbo, O., and Camara, B. (2003). Biosynthesis of the food and cosmetic plant pigment bixin (Annatto). *Science* 300, 2089–2091. doi: 10.1126/science.1085162

Brandi, F., Bar, E., Mourgues, F., Horváth, G., Turcsi, E., Giuliano, G., et al. (2011). Study of 'Redhaven' peach and its white-fleshed mutant suggests a key role of CCD4 carotenoid dioxygenase in carotenoid and norisoprenoid volatile metabolism. *BMC Plant Biol.* 11:24. doi: 10.1186/1471-2229-11-24

Campbell, R., Ducreux, L. J. M., Morris, W. L., Morris, J. A., Suttle, J. C., Ramsay, G., et al. (2010). The metabolic and developmental roles of carotenoid cleavage dioxygenase 4 from potato. *Plant Physiol.* 154, 656–664. doi: 10.1104/pp.110.158733

Carballo-Uicab, V. M., Cárdenas-Conejo, Y., Vallejo-Cardona, A. A., Aguilar-Espinosa, M., Rodríguez-Campos, J., Serrano-Posada, H., et al. (2019). Isolation and functional characterization of two dioxygenases putatively involved in bixin biosynthesis in annatto (*Bixa orellana* L.). *PeerJ* 7:e7064. doi: 10.7717/peerj.7064

Cárdenas-Conejo, Y., Carballo-Uicab, V., Lieberman, M., Aguilar-Espinosa, M., Comai, L., and Rivera-Madrid, R. (2015). De novo transcriptome sequencing in *Bixa orellana* to identify genes involved in methylerythritol phosphate, carotenoid and bixin biosynthesis. *BMC Genomics* 16:877. doi: 10.1186/s12864-015-2065-4

Daruwalla, A., and Kiser, P. D. (2020). Structural and mechanistic aspects of carotenoid cleavage dioxygenases (CCDs). *Biochim. Biophys. Acta Mol. Cell Biol. Lipids* 1865:158590. doi: 10.1016/j.bbalip.2019.158590

de Oliveira Júnior, R. G., Bonnet, A., Braconnier, E., Groult, H., Prunier, G., Beaugeard, L., et al. (2019). Bixin, an apocarotenoid isolated from *Bixa orellana* L., sensitizes human melanoma cells to dacarbazine-induced apoptosis through ROS-mediated cytotoxicity. *Food Chem. Toxicol.* 125, 549–561. doi: 10.1016/j.fct.2019.02.013

Devi, V. N., V, N. A., and Prasad, N. (2013). Annatto: eco-friendly and potential source for natural dye. *Int. Res. J. Pharm.* 4, 106–108. doi: 10.7897/2230-8407.04623

Evans, W. (2000). Annatto: a natural choice. *Biologist (London)* 47, 181–184.

Faria, D. V., de Freitas Correia, L. N., Batista, D. S., Vital, C. E., Heringer, A. S., De-la-Peña, C., et al. (2020). 5-Azacytidine downregulates the SABATH methyltransferase genes and augments bixin content in *Bixa orellana* L. leaves. *Plant Cell. Tiss. Org.* 142, 425–434. doi: 10.1007/s11240-020-01857-8

Fiorilli, V., Wang, J. Y., Bonfante, P., Lanfranco, L., and Al-Babili, S. (2019). Apocarotenoids: old and new mediators of the arbuscular mycorrhizal symbiosis. *Front. Plant Sci.* 10:1186. doi: 10.3389/fpls.2019.01186

Floss, D. S., and Walter, M. H. (2009). Role of carotenoid cleavage dioxygenase 1 (CCD1) in apocarotenoid biogenesis revisited. *Plant Signal. Behav.* 4, 172–175. doi: 10.4161/psb.4.3.7840

Floss, D. S., Schliemann, W., Schmidt, J., Strack, D., and Walter, M. H. (2008). RNA Interference-mediated repression of MtCCD1 in mycorrhizal roots of *Medicago truncatula* causes accumulation of C27 apocarotenoids, shedding light on the functional role of CCD1. *Plant Physiol.* 148, 1267–1282. doi: 10.1104/pp.108.125062

Frusciante, S., Demurtas, O. C., Sulli, M., Mini, P., Aprea, G., Diretto, G., et al. (2021). Heterologous expression of *Bixa orellana* cleavage dioxygenase 4–3 drives crocin but not bixin biosynthesis. *Plant Physiol.* kiab583. doi: 10.1093/plphys/kiab583

Frusciante, S., Diretto, G., Bruno, M., Ferrante, P., Pietrella, M., Prado-Cabrero, A., et al. (2014). Novel carotenoid cleavage dioxygenase catalyzes the first dedicated step in saffron crocin biosynthesis. *Proc. Natl. Acad. Sci. U.S.A.* 111, 12246–12251. doi: 10.1073/pnas.1404629111

Galindo-Cuspinera, V., Lubran, M. B., and Rankin, S. A. (2002). Comparison of volatile compounds in water-and oil-soluble annatto (*Bixa orellana* L.) extracts. *J. Agric. Food Chem.* 50, 2010–2015. doi: 10.1021/jf011325h

Giuliano, G., Rosati, C., and Bramley, P. M. (2003). To dye or not to dye: biochemistry of annatto unveiled. *Trends Biotechnol.* 21, 513–516. doi: 10.1016/j.tibtech.2003.10.001

Gómez-Gómez, L., Diretto, G., Ahrazem, O., and Al-Babili, S. (2020). "Determination of in vitro and in vivo activities of plant Carotenoid Cleavage Oxygenases," in *Plant and Food Carotenoids: Methods and Protocols*, eds M. Rodríguez-Concepción and R. Welsch (New York, NY: Springer US), 63–74. doi: 10.1007/978-1-4939-9952-1_5

Gonzalez-Jorge, S., Ha, S.-H., Magallanes-Lundback, M., Gilliland, L. U., Zhou, A., Lipka, A. E., et al. (2013). Carotenoid cleavage dioxygenase4 is a negative regulator of β-carotene content in *Arabidopsis* seeds. *Plant Cell* 25, 4812–4826. doi: 10.1105/tpc.113.119677

Harrison, P. J., and Bugg, T. D. H. (2014). Enzymology of the carotenoid cleavage dioxygenases: reaction mechanisms, inhibition and biochemical roles. *Arch. Biochem. Biophys.* 544, 105–111. doi: 10.1016/j.abb.2013.10.005

Hou, X., Rivers, J., León, P., McQuinn, R. P., and Pogson, B. J. (2016). Synthesis and function of apocarotenoid signals in plants. *Trends Plant Sci.* 21, 792–803. doi: 10.1016/j.tplants.2016.06.001

Huang, F.-C., Horváth, G., Molnár, P., Turcsi, E., Deli, J., Schrader, J., et al. (2009). Substrate promiscuity of RdCCD1, a carotenoid cleavage oxygenase from *Rosa damascena*. *Phytochemistry* 70, 457–464. doi: 10.1016/j.phytochem.2009.01.020

Ilg, A., Beyer, P., and Al-Babili, S. (2009). Characterization of the rice carotenoid cleavage dioxygenase 1 reveals a novel route for geranial biosynthesis. *FEBS J.* 276, 736–747. doi: 10.1111/j.1742-4658.2008.06820.x

Ilg, A., Bruno, M., Beyer, P., and Al-Babili, S. (2014). Tomato carotenoid cleavage dioxygenases 1A and 1B: relaxed double bond specificity leads to a plenitude of dialdehydes, mono-apocarotenoids and isoprenoid volatiles. *FEBS Open Biol.* 4, 584–593. doi: 10.1016/j.fob.2014.06.005

Jako, C., Coutu, C., Roewer, I., Reed, D. W., Pelcher, L. E., and Covello, P. S. (2002). Probing carotenoid biosynthesis in developing seed coats of Bixa orellana (Bixaceae) through expressed sequence tag analysis. *Plant Sci.* 163, 141–145. doi: 10.1016/S0168-9452(02)00083-3

Jimenez-Lopez, J. C., Lopez-Valverde, F. J., Robles-Bolivar, P., Lima-Cabello, E., Gachomo, E. W., and Kotchoni, S. O. (2016). Genome-wide identification and functional classification of tomato (*Solanum lycopersicum*) aldehyde

dehydrogenase (ALDH) gene superfamily. *PLoS One* 11:e0164798. doi: 10.1371/journal.pone.0164798

Kloer, D. P., and Schulz, G. E. (2006). Structural and biological aspects of carotenoid cleavage. *Cell Mol. Life Sci.* 63, 2291–2303. doi: 10.1007/s00018-006-6176-6

Ko, M. R., Song, M.-H., Kim, J. K., Baek, S.-A., You, M. K., Lim, S.-H., et al. (2018). RNAi-mediated suppression of three carotenoid-cleavage dioxygenase genes. OsCCD1, 4a, and 4b, increases carotenoid content in rice. *J. Exp. Bot.* 69, 5105–5116. doi: 10.1093/jxb/ery300

Kumar, S., Stecher, G., Li, M., Knyaz, C., and Tamura, K. (2018). MEGA X: molecular evolutionary genetics analysis across computing platforms. *Mol. Biol. Evol.* 35, 1547–1549. doi: 10.1093/molbev/msy096

Lashbrooke, J. G., Young, P. R., Dockrall, S. J., Vasanth, K., and Vivier, M. A. (2013). Functional characterisation of three members of the *Vitis vinifera* L. carotenoid cleavage dioxygenase gene family. *BMC Plant Biol.* 13:156. doi: 10.1186/1471-2229-13-156

Ma, G., Zhang, L., Matsuta, A., Matsutani, K., Yamawaki, K., Yahata, M., et al. (2013). Enzymatic formation of β-citraurin from β-cryptoxanthin and zeaxanthin by carotenoid cleavage dioxygenase4 in the flavedo of citrus fruit. *Plant Physiol.* 163, 682–695. doi: 10.1104/pp.113.223297

Mahendranath, G., Venugopalan, A., Parimalan, R., Giridhar, P., and Ravishankar, G. A. (2011). Annatto pigment production in root cultures of Achiote (*Bixa orellana* L.). *Plant Cell. Tiss. Org.* 106, 517–522. doi: 10.1007/s11240-011-9931-9

Martí, M., Diretto, G., Aragonés, V., Frusciante, S., Ahrazem, O., Gómez-Gómez, L., et al. (2020). Efficient production of saffron crocins and picrocrocin in *Nicotiana benthamiana* using a virus-driven system. *Metab. Eng.* 61, 238–250. doi: 10.1016/j.ymben.2020.06.009

Mathieu, S., Terrier, N., Procureur, J., Bigey, F., and Günata, Z. (2005). A carotenoid cleavage dioxygenase from *Vitis vinifera* L.: functional characterization and expression during grape berry development in relation to C13-norisoprenoid accumulation. *J. Exp. Bot.* 56, 2721–2731. doi: 10.1093/jxb/eri265

Mercadante, A. Z., Steck, A., and Pfander, H. (1997). Isolation and identification of new apocarotenoids from annatto (*Bixa orellana*) Seeds. *J. Agric. Food Chem.* 45, 1050–1054. doi: 10.1021/jf960412k

Mercadante, A. Z., Steck, A., Rodriguez-Amaya, D., Pfander, H., and Britton, G. (1996). Isolation of methyl 9'Z-apo-6'-lycopenoate from Bixa orellana. *Phytochemistry* 41, 1201–1203. doi: 10.1016/0031-9422(95)00784-9

Messing, S. A., Gabelli, S. B., Echeverria, I., Vogel, J. T., Guan, J. C., Tan, B. C., et al. (2010). Structural insights into maize viviparous14, a key enzyme in the biosynthesis of the phytohormone abscisic acid. *Plant Cell* 22, 2970–2980. doi: 10.1105/tpc.110.074815

Mi, J., Jia, K.-P., Balakrishna, A., Feng, Q., and Al-Babili, S. (2019). A highly sensitive SPE Derivatization–UHPLC–MS approach for quantitative profiling of carotenoid-derived dialdehydes from vegetables. *J. Agric. Food Chem.* 67, 5899–5907. doi: 10.1021/acs.jafc.9b01749

Misawa, N., Satomi, Y., Kondo, K., Yokoyama, A., Kajiwara, S., Saito, T., et al. (1995). Structure and functional analysis of a marine bacterial carotenoid biosynthesis gene cluster and astaxanthin biosynthetic pathway proposed at the gene level. *J. Bacteriol.* 177, 6575–6584. doi: 10.1128/jb.177.22.6575-6584.1995

Moreira, P. A., Lins, J., Dequigiovanni, G., Veasey, E. A., and Clement, C. R. (2015). The domestication of annatto (*Bixa orellana*) from *Bixa urucurana* in amazonia. *Econ. Bot.* 69, 127–135. doi: 10.1007/s12231-015-9304-0

Nawade, B., Shaltiel-Harpaz, L., Yahyaa, M., Bosamia, T. C., Kabaha, A., Kedoshim, R., et al. (2020). Analysis of apocarotenoid volatiles during the development of Ficus carica fruits and characterization of carotenoid cleavage dioxygenase genes. *Plant Sci.* 290:110292. doi: 10.1016/j.plantsci.2019.110292

Nic-Can, G., Hernández-Castellano, S., Kú-González, A., Loyola-Vargas, V. M., and De-la-Peña, C. (2013). An efficient immunodetection method for histone modifications in plants. *Plant Methods* 9:47. doi: 10.1186/1746-4811-9-47

Ohmiya, A., Kishimoto, S., Aida, R., Yoshioka, S., and Sumitomo, K. (2006). Carotenoid cleavage dioxygenase (CmCCD4a) contributes to white color formation in chrysanthemum petals. *Plant Physiol.* 142, 1193–1201. doi: 10.1104/pp.106.087130

Priya, R., Sneha, P., Rivera Madrid, R., and Doss, C. G. P. (2017). Molecular modeling and dynamic simulation of *Arabidopsis thaliana* carotenoid cleavage dioxygenase gene: a comparison with Bixa orellana and *Crocus sativus*. *J. Cell. Biochem.* 118, 2712–2721. doi: 10.1002/jcb.25919

Ramamoorthy, S., Madrid, R. R., and Doss, C. G. P. (2020). *Biology, Chemistry and Applications of Apocarotenoids*. Boca Raton, FL: CRC Press.

Rivera, S. M., Christou, P., and Canela-Garayoa, R. (2014). Identification of carotenoids using mass spectrometry. *Mass Spectr. Rev.* 33, 353–372. doi: 10.1002/mas.21390

Rivera-Madrid, R., Aguilar-Espinosa, M., Cárdenas-Conejo, Y., and Garza-Caligaris, L. E. (2016). Carotenoid derivates in achiote (*Bixa orellana*) seeds: synthesis and health promoting properties. *Front. Plant Sci.* 7:1406. doi: 10.3389/fpls.2016.01406

Rivera-Madrid, R., and Ramamoorthy, S. (2020). "Apocarotenoid molecules are continuously being discovered in all taxa," in *Biology, Chemistry, and Applications of Apocarotenoids*, eds R. Rivera-Madrid, S. Ramamoorthy, and G. Priya Doss (Boca Raton, FL: CRC Press), 41–57.

Rivera-Madrid, R., Escobedo-Gm, R. M., Balam-Galera, E., Vera-Ku, M., and Harries, H. (2006). Preliminary studies toward genetic improvement of annatto (*Bixa orellana* L.). *Sci. Hortic.* 109, 165–172. doi: 10.1016/j.scienta.2006.03.011

Rodrigo, M. J., Alquézar, B., Alós, E., Medina, V., Carmona, L., Bruno, M., et al. (2013). A novel carotenoid cleavage activity involved in the biosynthesis of Citrus fruit-specific apocarotenoid pigments. *J. Exp. Bot.* 64, 4461–4478. doi: 10.1093/jxb/ert260

Rodríguez-Ávila, N. L., Narváez-Zapata, J. A., Ramírez-Benítez, J. E., Aguilar-Espinosa, M. L., and Rivera-Madrid, R. (2011b). Identification and expression pattern of a new carotenoid cleavage dioxygenase gene member from Bixa orellana. *J. Exp. Bot.* 62, 5385–5395. doi: 10.1093/jxb/err201

Rodríguez-Ávila, N. L., Narváez-Zapata, J. A., Aguilar-Espinosa, M., and Rivera-Madrid, R. (2011a). Regulation of pigment-related genes during flower and fruit development of *Bixa orellana*. *Plant Mol. Biol. Rep.* 29, 43–50.

Rodríguez-Ávila, N. L., Narváez-Zapata, J. A., Aguilar-Espinosa, M. L., and Rivera-Madrid, R. (2009). Full-length gene enrichment by using an optimized RNA isolation protocol in *Bixa orellana* recalcitrant tissues. *Mol. Biotechnol.* 42, 84–90. doi: 10.1007/s12033-008-9138-4

Rodríguez-Concepción, M., and Welsch, R. (2020). *Plant and Food Carotenoids*. Berlin: Springer.

Rottet, S., Devillers, J., Glauser, G., Douet, V., Besagni, C., and Kessler, F. (2016). Identification of plastoglobules as a site of carotenoid cleavage. *Front. Plant Sci.* 7:1855. doi: 10.3389/fpls.2016.01855

Rubio-Moraga, Á, Rambla, J., Santaella, M., Gomez, M., Orzáez, D., Granell, A., et al. (2008). Cytosolic and plastoglobule-targeted carotenoid dioxygenases from *Crocus sativus* are both involved in beta-ionone release. *J. Biol. Chem.* 283, 24816–24825. doi: 10.1074/jbc.M804000200

Rubio-Moraga, A., Rambla, J. L., Fernández-de-Carmen, A., Trapero-Mozos, A., Ahrazem, O., Orzáez, D., et al. (2014). New target carotenoids for CCD4 enzymes are revealed with the characterization of a novel stress-induced carotenoid cleavage dioxygenase gene from *Crocus sativus*. *Plant Mol. Biol.* 86, 555–569. doi: 10.1007/s11103-014-0250-5

Ruch, S., Beyer, P., Ernst, H., and Al-Babili, S. (2005). Retinal biosynthesis in eubacteria: in vitro characterization of a novel carotenoid oxygenase from *Synechocystis* sp. PCC 6803. *Mol. Microbiol.* 55, 1015–1024. doi: 10.1111/j.1365-2958.2004.04460.x

Sankari, M., Hemachandran, H., Anantharaman, A., Babu, S., Madrid, R. R., C, G. P. D., et al. (2016). Identifying a carotenoid cleavage dioxygenase 4a Gene and its efficient Agrobacterium-mediated genetic transformation in *Bixa orellana* L. *Appl. Biochem. Biotechnol.* 179, 697–714. doi: 10.1007/s12010-016-2025-8

Schwartz, S. H., Qin, X., and Zeevaart, J. A. D. (2001). Characterization of a novel carotenoid cleavage dioxygenase from plants. *J. Biol. Chem.* 276, 25208–25211. doi: 10.1074/jbc.M102146200

Schwartz, S. H., Tan, B. C., Gage, D. A., Zeevaart, J. A., and McCarty, D. R. (1997). Specific oxidative cleavage of carotenoids by VP14 of maize. *Science* 276, 1872–1874. doi: 10.1126/science.276.5320.1872

Shahid ul, I., Rather, L. J., and Mohammad, F. (2016). Phytochemistry, biological activities and potential of annatto in natural colorant production for industrial applications – a review. *J. Adv. Res.* 7, 499–514. doi: 10.1016/j.jare.2015.11.002

Sharma, P., Segat, A., Kelly, A. L., and Sheehan, J. J. (2020). Colorants in cheese manufacture: production, chemistry, interactions, and regulation. *Compr. Rev. Food Sci. Food Saf.* 19, 1220–1242. doi: 10.1111/1541-4337.12519

Simkin, A. J., Underwood, B. A., Auldridge, M., Loucas, H. M., Shibuya, K., Schmelz, E., et al. (2004). Circadian regulation of the PhCCD1 carotenoid

cleavage dioxygenase controls emission of beta-ionone, a fragrance volatile of petunia flowers. *Plant Physiol.* 136, 3504–3514. doi: 10.1104/pp.104.049718

Soares, V. L. F., Rodrigues, S. M., de Oliveira, T. M., de Queiroz, T. O., Lima, L. S., Hora-Júnior, B. T., et al. (2011). Unraveling new genes associated with seed development and metabolism in *Bixa orellana* L. by expressed sequence tag (EST) analysis. *Mol. Biol. Rep.* 38, 1329–1340. doi: 10.1007/s11033-010-0234-8

Tan, B. C., Cline, K., and McCarty, D. R. (2001). Localization and targeting of the VP14 epoxy-carotenoid dioxygenase to chloroplast membranes. *Plant J.* 27, 373–382. doi: 10.1046/j.1365-313x.2001.01102.x

Tan, B.-C., Joseph, L. M., Deng, W. T., Liu, L., Li, Q. B., Cline, K., et al. (2003). Molecular characterization of the *Arabidopsis* 9-cis epoxycarotenoid dioxygenase gene family. *Plant J.* 35, 44–56.

Tirimanna, A. S. L. (1981). Study of the carotenoid pigments of *Bixa orellana* L. Seeds by thin layer chromatography. *Microchim. Acta* 76, 11–16. doi: 10.1007/BF01197299

Trujillo-Hdz, J. A., Cárdenas-Conejo, Y., Turriza, P. E., Aguilar-Espinosa, M., Carballo-Uicab, V., Garza-Caligaris, L. E., et al. (2016). Functional polymorphism in lycopene beta-cyclase gene as a molecular marker to predict bixin production in *Bixa orellana* L. (achiote). *Mol. Breed.* 36:135. doi: 10.1007/s11032-016-0555-y

Us-Camas, R., Aguilar-Espinosa, M., and Rivera-Madrid, R. (2020). "An overview of the role of abscisic acid (ABA): from plants to humans," in *Biology, Chemistry, and Applications of Apocarotenoids*, eds R. Rivera-Madrid, S. Ramamoorthy, and G. Priya Doss (Boca Raton, FL: CRC Press), 149–172.

van Breemen, R. B. (2001). Mass spectrometry of carotenoids. *Curr. Protoc. Food Anal. Chem.* F2.4.1–F2.4.13. doi: 10.1002/0471142913.faf0204s00

Vega-Teijido, M., Cantero, J., Rodrigo, M. J., López, C., and Zunini, M. P. (2019). An in silico study of the citrus dioxygenases CCD4 family substrates. *J. Biomol. Struct. Dyn.* 37, 2086–2097. doi: 10.1080/07391102.2018.1477619

Vogel, J. T., Tan, B.-C., McCarty, D. R., and Klee, H. J. (2008). The carotenoid cleavage dioxygenase 1 enzyme has broad substrate specificity, cleaving multiple carotenoids at two different bond positions. *J. Biol. Chem.* 283, 11364–11373. doi: 10.1074/jbc.M710106200

Vranová, E., Coman, D., and Gruissem, W. (2013). Network analysis of the MVA and MEP pathways for isoprenoid synthesis. *Annu. Rev. Plant Biol.* 64, 665–700. doi: 10.1146/annurev-arplant-050312-120116

Walter, M. H., and Strack, D. (2011). Carotenoids and their cleavage products: biosynthesis and functions. *Nat. Prod. Rep.* 28, 663–692. doi: 10.1039/c0np00036a

Walter, M. H., Floss, D. S., and Strack, D. (2010). Apocarotenoids: hormones, mycorrhizal metabolites and aroma volatiles. *Planta* 232, 1–17. doi: 10.1007/s00425-010-1156-3

Wang, J. Y., Haider, I., Jamil, M., Fiorilli, V., Saito, Y., Mi, J., et al. (2019). The apocarotenoid metabolite zaxinone regulates growth and strigolactone biosynthesis in rice. *Nat. Commun.* 10:810. doi: 10.1038/s41467-019-08461-1

Ytterberg, A. J., Peltier, J.-B., and van Wijk, K. J. (2006). Protein profiling of plastoglobules in chloroplasts and chromoplasts. A surprising site for differential accumulation of metabolic enzymes. *Plant Physiol.* 140, 984–997. doi: 10.1104/pp.105.076083

Zheng, X., Zhu, K., Sun, Q., Zhang, W., Wang, X., Cao, H., et al. (2019). Natural variation in CCD4 promoter underpins species-specific evolution of red coloration in citrus peel. *Mol. Plant* 12, 1294–1307. doi: 10.1016/j.molp.2019.04.014

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.

frontiers in Plant Science

ORIGINAL RESEARCH
published: 14 February 2022
doi: 10.3389/fpls.2021.822198

Transcriptome Co-expression Network and Metabolome Analysis Identifies Key Genes and Regulators of Proanthocyanidins Biosynthesis in Brown Cotton

Zhenzhen Wang[1], Xiaomeng Zhang[1], Shoupu He[1,2,3], Abdul Rehman[2], Yinhua Jia[1], Hongge Li[1,2], Zhaoe Pan[1], Xiaoli Geng[1], Qiong Gao[1], Liru Wang[1], Zhen Peng[1,2,3] and Xiongming Du[1,2,3]**

[1] State Key Laboratory of Cotton Biology, Institute of Cotton Research, Chinese Academy of Agricultural Sciences, Anyang, China, [2] Zhengzhou Research Base, State Key Laboratory of Cotton Biology, Zhengzhou University, Zhengzhou, China, [3] National Nanfan Research Institute (Sanya), Chinese Academy of Agricultural Sciences, Sanya, China

OPEN ACCESS

Edited by:
Xiumin Fu,
South China Botanical Garden,
Chinese Academy of Sciences (CAS),
China

Reviewed by:
Vasundhara Thakur,
Panjab University, India
Liu Xiaofen,
Zhejiang University, China

***Correspondence:**
Zhen Peng
pengzhen01@caas.cn
Xiongming Du
duxiongming@caas.cn

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

Received: 25 November 2021
Accepted: 29 December 2021
Published: 14 February 2022

Citation:
Wang Z, Zhang X, He S, Rehman A, Jia Y, Li H, Pan Z, Geng X, Gao Q, Wang L, Peng Z and Du X (2022) Transcriptome Co-expression Network and Metabolome Analysis Identifies Key Genes and Regulators of Proanthocyanidins Biosynthesis in Brown Cotton. Front. Plant Sci. 12:822198. doi: 10.3389/fpls.2021.822198

Brown cotton fiber (BCF) is a unique raw material of naturally colored cotton (NCC). But characteristics of the regulatory gene network and metabolic components related to the proanthocyanidins biosynthesis pathway at various stages of its fiber development remain unclear. Here, the dynamic changes in proanthocyanidins biosynthesis components and transcripts in the BCF variety "Zong 1-61" and its white near-isogenic lines (NILs) "RT" were characterized at five fiber developmental stages (0, 5, 10, 15, and 20 days post-anthesis; DPA). Enrichment analysis of differentially expressed genes (DEGs), comparison of metabolome differences, and pathway enrichment analysis of a weighted gene correlation network analysis together revealed the dominant gene expression of flavonoid biosynthesis (FB), phenylpropanoid metabolisms, and some carbohydrate metabolisms at 15 or 20 DPA than white cotton. Eventually, 63 genes were identified from five modules putatively related to FB. Three R2R3-MYB and two bHLH transcription factors were predicted as the core genes. Further, *GhANS*, *GhANR1,* and *GhUFGT2* were preliminarily regulated by *GhMYB46*, *GhMYB6*, and *GhMYB3,* respectively, according to yeast one-hybrid assays *in vitro*. Our findings provide an important transcriptional regulatory network of proanthocyanidins biosynthesis pathway and dynamic flavonoid metabolism profiles.

Keywords: brown cotton, transcriptome, metabolome, flavonoid metabolism, yeast one-hybrid, R2R3-MYB genes

INTRODUCTION

Naturally colored cotton (NCC) is a variety of cotton with natural pigmentation in its fiber. The most common fiber colors are brown and green (Dutt et al., 2004). Compared with traditional cotton plants, colored ones provide significant benefits, namely a higher amount of tannin and phenols, and also resistance against diseases and insects. Because NCC fiber has a natural color, it does not undergo chemical bleaching and dyeing during textile processing. This advantage not only reduces environmental pollution and the harm to human health, but also decreases

cost and increases the revenue of cotton growers. Indeed, NCC is considered as ecologically and environment-friendly cotton products with broad prospects (Hua et al., 2007; Efe et al., 2009; Khatri et al., 2015). However, the quality of its colored cotton is not as good as that of white cotton. Likewise, single color and unstable pigment heredity have restricted its production, promotion, and end-use (Feng et al., 2011, 2015). In recent years, breeders have developed some commercial color cotton genotypes with good fiber quality. Yet, due to the lack of NCC genetic resources, traditional breeding programs are reluctant to change the color type (Sun et al., 2021). Therefore, genetic engineering or gene-editing technology is considered a promising tool to produce new types of fiber color.

Brown cotton is considered the most common type of NCC. The early identification of its chemical components showed that flavonoids were responsible for this brown color of cotton fiber (Xiao et al., 2007; Hua et al., 2009). Moreover, procyanidins (PAs) have proven to be the main component of pigment component in brown cotton fibers (BCFs; Li et al., 2013; Feng et al., 2014; Xiao et al., 2014). Additionally, *leucoanthocyanidin reductase* (*LAR*) plays a crucial role in the PA biosynthesis pathway of brown fibers, and its metabolites catechin and epicatechin are the main precursors of condensed tannins. *Anthocyanidin reductase* (*ANR*) is a key structural gene in the PA pathway (Feng et al., 2014). Recently, these two genes, *LAR* and *ANR*, have been verified for their transgene functionality, enabling that the development of new colored cotton germplasm has been developed (Gao et al., 2019). In the last decade, transcriptomic or proteomic analyses were used to uncover the PA precursor synthesis pathway involved in fiber pigmentation. It entails a group of flavonoid pathway genes, including those for the particular enzymes *LAR* and *ANR* of two PA precursors, all of which have been found upregulated or deposited in brown cotton (Xiao et al., 2007, 2014; Feng et al., 2013, 2014; Li et al., 2013; Gong et al., 2014; Hinchliffe et al., 2016; Peng et al., 2020).

The color of BCF is a genetically determined feature, which is produced by condensed tannins accumulating in the lumen of cotton fiber (Carvalho et al., 2014; Feng et al., 2014; Gong et al., 2014). Early work by Kohel (1985) tested the allele of the known lint color gene, which finds that the brown fiber color may be controlled by six loci, namely, *Lc1* to *Lc6*. Whereas *Lc1* and *Lc2* are responsible for lint that is medium-brown, *Lc3* is responsible for its dark brown appearance, and *Lc4*, *Lc5*, and *Lc6* are responsible for light-brown coloring (Wang et al., 2014). Simple sequence repeat (SSR) markers revealed that *Lc1* is located on chromosome A07 in upland cotton, whereas *Lc2* was detected on chromosome A06 (Wang et al., 2014). Later, Hinchliffe et al. (2016) found that *Lc1* is *GhTT2_A07*, which is a transcription factor (TF) similar to *Arabidopsis TRANSPARENT TESTA 2* (*TT2*). More recently, Wen et al. (2018) and Yan et al. (2018) elucidated the critical regulatory mechanisms of the *GhTT2* gene involved in the color of BCF *via* its fine mapping, gene expression difference analysis, and transgenic function verification. To date, only *Lc1* gene is well understood, whereas the mechanisms of the other five Lc genes are still unknown.

Few studies have investigated the relationship between proanthocyanidins biosynthesis and metabolism in cotton fiber and their effects on metabolism-related functional genes. In previous research, the metabolic pathways and key proteins involved in BCF's development and pigment biosynthesis were identified through proteomic and metabolomic analyses, for which selection criteria consisted of the fiber elongation and secondary wall thickening phases (Peng et al., 2020). In this study, we instead focused on early fiber development (0–20 DPA, different nodes of fiber initiation and elongation), using transcriptome and non-targeted metabolome analyses to identify key genes, metabolites, and coexpression networks involved in proanthocyanidin biosynthesis, and obtained pivotal TF genes *via* yeast one hybrid. The key TFs or genes active in the gene coexpression network of transcriptional regulation as related to brown color of cotton fiber, and likewise for other metabolic pathways related to fiber quality formation, are then discussed. The understanding of molecular mechanisms of pigmentation in brown cotton is still limited. Therefore, characterizing the metabolic pathways or transcriptome as related to PA biosynthesis in BCF by high-throughput sequencing technology is instrumental for mining prominent genes for the use in brown cotton breeding, and also for studying the corresponding relevant transcriptional regulation mechanisms.

MATERIALS AND METHODS

Plant Materials and Tissue Collection

Two upland cotton genotypes were used, Zong 1-61 (Abb. Z161), natural brown fiber cotton, and RT-white fiber (Abb. RT), whose cotton fibers are white. RT is a near-isogenic line (NIL) of Z161, as shown by Peng et al. (2020). The eight agronomic traits and four fiber color characters of Z161 and RT, namely, boll weight (BW), micronaire (MIC), fiber elongation (FE), seed-index (SI), lint percentage (LP), fiber length (mm) (FL), fiber strength (cN/tex) (FS), fiber length uniformity (LU), total chromatic aberration (ΔE), chromaticity difference (ΔC), chromaticity index a (Δa), and chromaticity index b (Δb) were determined from six environments, including Anyang of Henan Province; Alaer and Kuitun of Xinjiang Province in 2018 and 2019 with three replications. Young cotton bolls were labeled using tags on their day of anthesis, with their cotton ovule, and fibers were harvested at 0, 5, 10, 15, and 20 DPA for total RNA extractions. The harvested fibers (or ovules) were stripped and immediately frozen in liquid nitrogen and stored at −80°C for the RNA extraction.

Detection of Color Difference in Cotton Fiber

Fiber samples of the two accessions were collected from three experimental sites (cities of Anyang, Alaer, and Kuitun in China) at maturity in 2019. The whiteness colorimeter YT-48A (Yante Technology Co., Ltd., Hangzhou, China) was used to determine the color difference value of a given fiber. That instrument can measure the color and color difference reflected by the object (paper, fiber, etc.), the CIE brightness (i.e., Ganz brightness W10 and color deviation TW10), the white of Hunter System Lab and Hunter (Lab), yellow, opacity, transparency, light scattering

coefficient, and light absorption coefficient of a given fiber sample. In this study, the whiteness of cotton fiber hunt (L, a, b) was measured (Ibraheem et al., 2012; Karaarslan et al., 2013). The total chromatic aberration (ΔE), chromaticity difference (ΔC), chromaticity index a (Δa), and chromaticity index b (Δb) were used to evaluate the cotton fiber color difference.

RNA Extraction, cDNA Library Construction, and Sequencing

Total RNA of the two cotton genotypes from the 0, 5, 10, 15, and 20 DPA fiber samples were isolated with the RNAprep Pure Plant Plus Kit (DP441) (Tiangen, Beijing, China) following the manufacturer's protocol. The quality and concentration of each RNA sample were checked using an Agilent 2100 Bioanalyzer (Agilent Technologies, Palo Alto, CA, United States) with 28S/18S $\geq$ 1.5, A260/A280 $\geq$ 2.0, and RIN $\geq$ 7.5. High-quality RNA was prepared to construct a cDNA library, each with three biological replicates.

The 30 cDNA libraries' construction and sequencing were performed by the BGI Gene Technology Co. (Shenzhen, China). Finally, the cDNA libraries were sequenced on the BGISEQ-500 sequencing platform, and 150-bp paired-end (PE150) reads were generated.

Mapping Onto the Cotton Genome and Quantification of Gene Expression

The sequences of the adaptor, any unknown bases, and all low-quality reads were removed using the SOAP nuke (v1.4.0, parameters –l 5, -q 0.5, -n 0.1)[1]. Then, these clean reads were saved in a FASTQ format. Next, the clean reads of all the 30 samples were mapped onto the *Gossypium hirsutum* (TM-1) reference genome files downloaded from https://www.cottongen.org/data/download/genome_tetraploid/AD1 (CRI v1) by Bowtie2 (Langmead and Salzberg, 2012) and HISAT (Kim et al., 2015), respectively. Their gene expression levels were calculated as Fragments per kilobase million (FPKM), using RSEM (v1.2.8) software package (Li and Dewey, 2011).

Identification of Differentially Expressed Genes and Kyoto Encyclopedia of Genes and Genomes Enrichment Analysis of Differentially Expressed Genes

Based on the negative binomial distribution, DESeq2 was used to identify the differentially expressed genes (DEGs; Love et al., 2014). For this designation of DEGs, the criteria of $\log_2$ (fold change) are greater than or equal to 1 and *Q*-value (adjusted *p-value)* is less than or equal to 0.05. To assign specific biological pathways to DEGs, the Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway annotation was used based on the KEGG database (Kanehisa et al., 2008). False discovery rates (FDRs) were controlled using established methods (Benjamini and Hochberg, 1995), such that KEGG pathways with FDR < 0.05 were deemed significantly enriched categories.

[1]https://github.com/BGI-fexlab/SOAPnuke

Comprehensive Analysis of Transcriptome and Metabolome

In a previous study (Peng et al., 2020), LC-ESI-MS-based untargeted metabolomics was performed on the six samples, and these were taken at 0, 10, 20 DPA for the two cotton lines from the above transcriptome. Here, the differentially accumulated metabolites (DAMs), that is those metabolites that underwent significantly differential accumulation according to these parameters: (1) VIP $\geq$ 1, (2) fold change $\geq$1.2 or $\leq$0.833, (3) *Q*-value < 0.05, were used for comprehensive analysis with the corresponding transcriptome data (refer to **Supplementary Tables 1–3**). These DAMs' pathways were then searched against the online database of KEGG pathways.

Gene Network Construction and Visualization

The weighted gene coexpression network analysis (WGCNA) was performed to identify modules and networks of highly correlated genes based on the FPKM values of all DEGs with the help of an R package (v1.68) (Langfelder and Horvath, 2008). The following calculation parameters were at least 75% of the samples having an FPKM > 1, and soft thresholding power = 7, minimum module size = 30, and minimum height for merging modules = 0.3 were selected to analyze the RNA-seq data sets. After clustering with all samples FPKM, it is found that the two Z161_5-R1 and Z161_10-R2 samples are outlier samples, which can be removed. Finally, a total of 10,891 DEGs were used as input files to participate in the WGCNA. The dynamic tree-cutting algorithm was used to cut the hierarchical clustering tree, and the module was defined after decomposing or merging branches to obtain a stable number of clusters. The model characteristic gene (ME) is defined as the first principal component of a given model. It can be considered as the representative of the gene expression profile for that module. All the modules contained 8,305 genes.

Identification of the Sample or Stage- and Metabolites Modules

We determined the relationship between each module and cotton phenotypic traits of the samples (sample = 1, all other samples = 0) and metabolites (= 9 flavonoids' content). Later, the correlations between a module and sample specificity and flavonoids were determined, for which a correlation matrix was drawn in the R software package "ggplot2" (Wickham, 2011). A positive correlation indicated that the genes in a module were expressed preferentially in a specific sample or stage or flavonoids in all other samples.

Candidate Gene Selection

First, we collected all gene family members (2,285) related to the pathway and TF family members (5,011; refer. http://planttfdb.gao-lab.org/index.php?sp=Ghi) in upland cotton and then counted their numbers as distributed in the different modules. Given that R2R3-MYB, basic helix-loop-helix proteins (bHLH), and WD40 protein (MBW complexes) are related to fiber pigmentation, we focused on the MBW complexes and structural genes. The crucial module networks (kMEmagenta,

kMEcyan, and kMElightyellow) containing R2R3-MYB, bHLH TFs, and candidate target genes were visualized using Cytoscape (v3.7.1, United States).

Identification of Conserved Motifs in Promoter and Yeast One-Hybrid Assay

Conserved motifs of five structural genes in the flavonoid biosynthesis (FB) pathway were investigated using the website toolkit Multiple Expectation maximization for Motif Elicitation (MEME 5.3.3). The optimized parameters of MEME were as follows: site distribution, zero or one occurrence per sequence; number of motifs, 20; and motif width, between 6 and 50 (Bailey et al., 2009). Meanwhile, there were nine conserved motifs, of those five genes were further annotated with TBtools (Chen et al., 2020). Yeast one-hybrid (Y1H) experiments were conducted using the Matchmaker Gold Y1H library Screening System (Clontech, Code 630493). The open reading frame (ORF) of three upland cotton MYB genes was successfully cloned, which was named *GhMYB3 (Gh_A05G235400)*, *GhMYB6 (Gh_D13G138800)*, and *GhMYB46 (Gh_D09G124900)* (named after *Arabidopsis* homologs). Based on a structural analysis of the promoter of leucoanthocyanidin dioxygenase (*GhLODX/GhANS*, *Gh_D08G197600*), anthocyanidin reductase *(GhANR1/2*, Gh_A05G152700/*Gh_D05G168500*), and UDP-glucose: flavonoid 3-O-glucosyltransferase (*GhUFGT1/2*, *Gh_A02G169700/Gh_D03G054700*), 13 DNA motifs were assayed *in vitro*. Two oligonucleotides containing tandem copies of three *cis*-regulatory sequences were inserted into the upstream region of the AbAr reporter gene in the pAbAi vector and then transformed into *Saccharomyces cerevisiae* Y1H Gold strain to produce the bait reporter strain. We tested the bait reporter strains on minimal synthetic defined media (SD) lacking uracil with different concentrations of Aureobasidin A to determine the minimal inhibitory concentration of Aureobasidin A for this bait. The full-length coding sequence of *GhMYB3*, *GhMYB6*, and *GhMYB46* was cloned into pGADT7 AD and then transformed into the bait reporter strains to generate different pairwise combinations. After 3–5 days at 30°C, we assessed and scored growth on the minimum synthesis-limited medium (SD) lacking leucine and supplemented with the lowest inhibitory concentration of Aureobasidin A. If the strains grew normally on both media, there was an interaction between the prey plasmid and the corresponding bait gene. On the contrary, if there was normal growth on SD/-Leu medium, yet no growth on the SD/-Leu/ABA* medium indicates null interaction between them. All the primers are listed (**Supplementary Table 4**), and the results of each step of the yeast one-hybrid assays are shown in **Supplementary Figures 1**, **2**.

Quantitative Real-Time PCR Analysis

The expression levels of 20 genes involved in the proanthocyanidins biosynthesis were measured in Z161 and RT at 0,5,10,15, and 20 DPA by conducting a Quantitative Real-Time PCR (qRT-PCR) analysis. The first strand cDNAs for the qRT-PCR were synthesized in a 20-μL solution (containing 1 μg of RNA as the template) using the First-strand cDNA Synthesis SuperMix kit (No. E047-01B; Novoprotein, Shanghai, China) and following the manufacturer's instructions. Later, the qRT-PCR was performed using Novostar® SYBR qPCR SuperMix Plus (Code No. E096-01A; Novoprotein, Shanghai, China) in a LightCycler 480 (Roche, Mannheim, Germany). The housekeeping gene was *GhUBQ*. The gene-specific primers were designed using the Primer-BLAST online tool[2], and all the primers are listed in **Supplementary Table 5**. Each qRT-PCR analysis was performed in three biological replicates, each with three technical replicates. The relative expression values of these genes were calculated according to the $2^{-\Delta\Delta CT}$ method (Livak and Schmittgen, 2001).

Data Analysis

Statistically significant differences in significance for the agronomic traits data between the two cotton genotypes were determined by two-way ANOVA, and these results were presented with the help of GraphPad Prism version 9 (GraphPad Software, Inc., San Diego, CA, United States). The online analysis platform OmicShare[3] was used for the KEGG enrichment analysis. Dot charts were used to depict the number of genes and KEGG enrichment analysis of DEGs, drawn using the R package "ggplot2." The heatmap for gene expression level was drawn with TBtools (Chen et al., 2020).

Availability of Data and Materials

The RNA-seq raw data set used in this study has been uploaded to Sequence Read Archive[4] under BioProject PRJNA766762. The metabolome raw data set was submitted to the Metabolite database[5] and has the project ID MTBLS2715.

RESULTS

Characteristics of the Phenotypic and Transcriptomic Data of Brown and White Cotton Genotypes

As noted before (Peng et al., 2020), the two cotton lines used in this study are a pair of NILs. No significant differences were detected in the agronomic traits between the two genotypes at maturity, except for their fiber color (**Figure 1A**). Specifically, we found significant differences in four indexes of fiber color between the two lines. These findings pave the way to further analyze the pertinent genes functioning to confer brown cotton its fiber pigmentation.

BGI Tech (Shenzhen, China) sequencing technology was used for comparative transcriptome analysis of the Z161 and RT cotton plants at the fiber development stage. Metabolomics is a non-targeted strategy to monitor the multiple changes among different samples. **Figure 1B** provides an overview of the metadata analysis strategies used in this study. After conducting a quantitative and differential analysis of gene expression levels, we carried out an in-depth analysis of three aspects.

[2]http://www.ncbi.nlm.nih.gov/tools/primer-blast/

[3]http://www.omicshare.com/tools

[4]https://www.ncbi.nlm.nih.gov/sra

[5]https://www.ebi.ac.uk/metabolights/

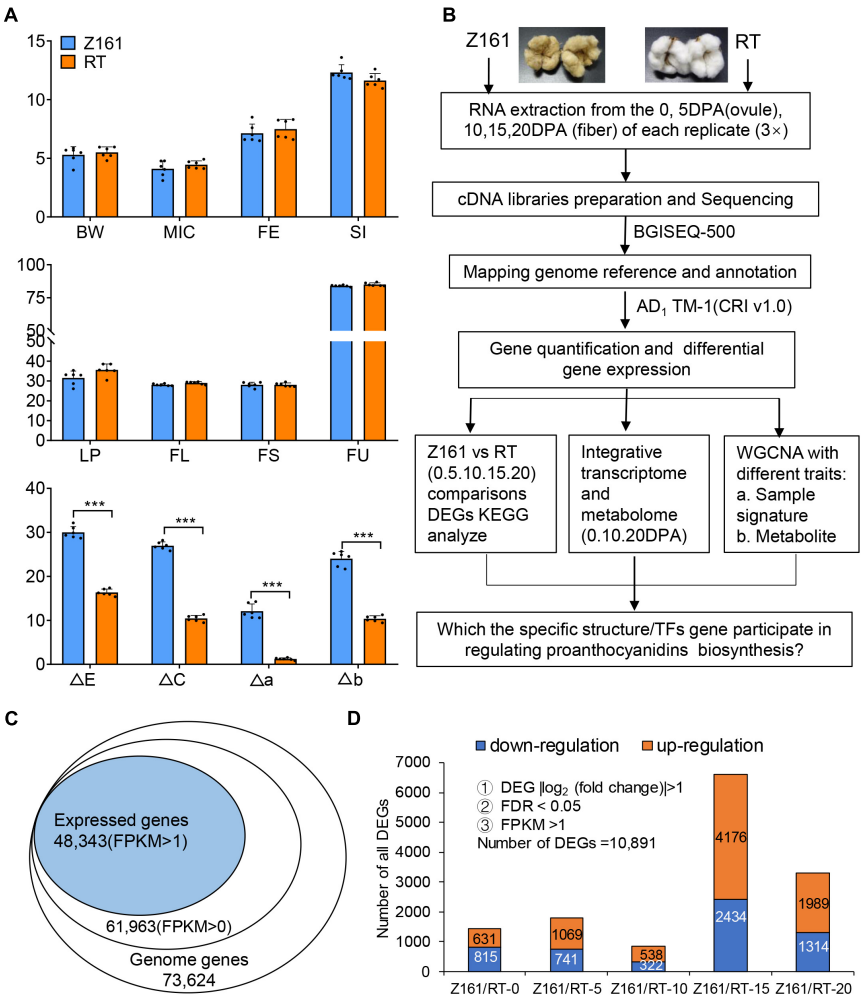


FIGURE 1 | Analysis of agronomic characters and transcriptome data of Z161 and RT at different cotton fiber development stages. **(A)** Analysis of eight agronomic characteristics, boll weight (BW), micronaire (MIC), fiber elongation (FE), seed-index (SI), lint percentage (LP), fiber length (mm) (FL), fiber strength (cN/tex) (FS), fiber length uniformity (FU), and fiber color traits (ΔE, ΔC, Δa, Δb), for 2 years (2018 and 2019) in three environments (Anyang, Alaer, and Kuitun). **(B)** Flow chart of the specific structure/TFs genes identified *via* transcriptomic and metabolic analyses. **(C)** Proportion of expressed genes identified in the *G. hirsutum* reference genome's gene database. **(D)** Number of DEGs between Z161 and RT at five fiber development stages. The significant difference is Student's *t* test; $^{*}p < 0.05$; $^{**}p < 0.01$; $^{***}p < 0.001$.

Finally, the specific TFs genes involved in the regulation of proanthocyanidins biosynthesis were screened out.

To further characterize the expression levels of pigmentation-related genes in BCF, the two cotton genotypes with contrasting fiber color at five developmental stages (0, 5, 10, 15, and 20 DPA) underwent a transcriptomic analysis. A total of 197.63-Gb high-quality base pairs were obtained, with an average of 6.59-Gb data per sample (**Supplementary Table 6**). Approximately 92.05–95.25% of the reads were mapped onto the *Gossypium hirsutum* genome, among them, of which 56.71–72.46% were uniquely aligned (**Supplementary Table 7**). A total of 61,963 genes (FPKM > 0) were detected in the two cotton genotypes sampled at 0–20 DPA. After filtering under the *a priori* specified conditions (FPKM > 1), a total of 48,343 (78.02%) of those genes were identified for further differentially expressed analysis (**Figure 1C**). These designated 10,891 DEGs related to the pigmentation of the two cotton genotypes were detected. The number of DEGs in either genotype was lowest at the 10 DPA stage and highest at 15 DPA, followed by 20 DPA (**Figure 1D** and **Supplementary Table 8**). Hence, by comparing gene expression differences between the two genotypes at the genome-wide level, the results indicated that 15 DPA is the key stage of brown pigment biosynthesis.

Comparative Transcriptome Analysis of Differential Gene Expression Profiles Between Brown and White Cotton Genotypes

Using the dual criteria of a 95% confidence level (FDR < 0.05) and a minimum 2-fold change, 10,891 (14.79% of 73,624) DEGs were identified in both Z161 and RT genotypes at five

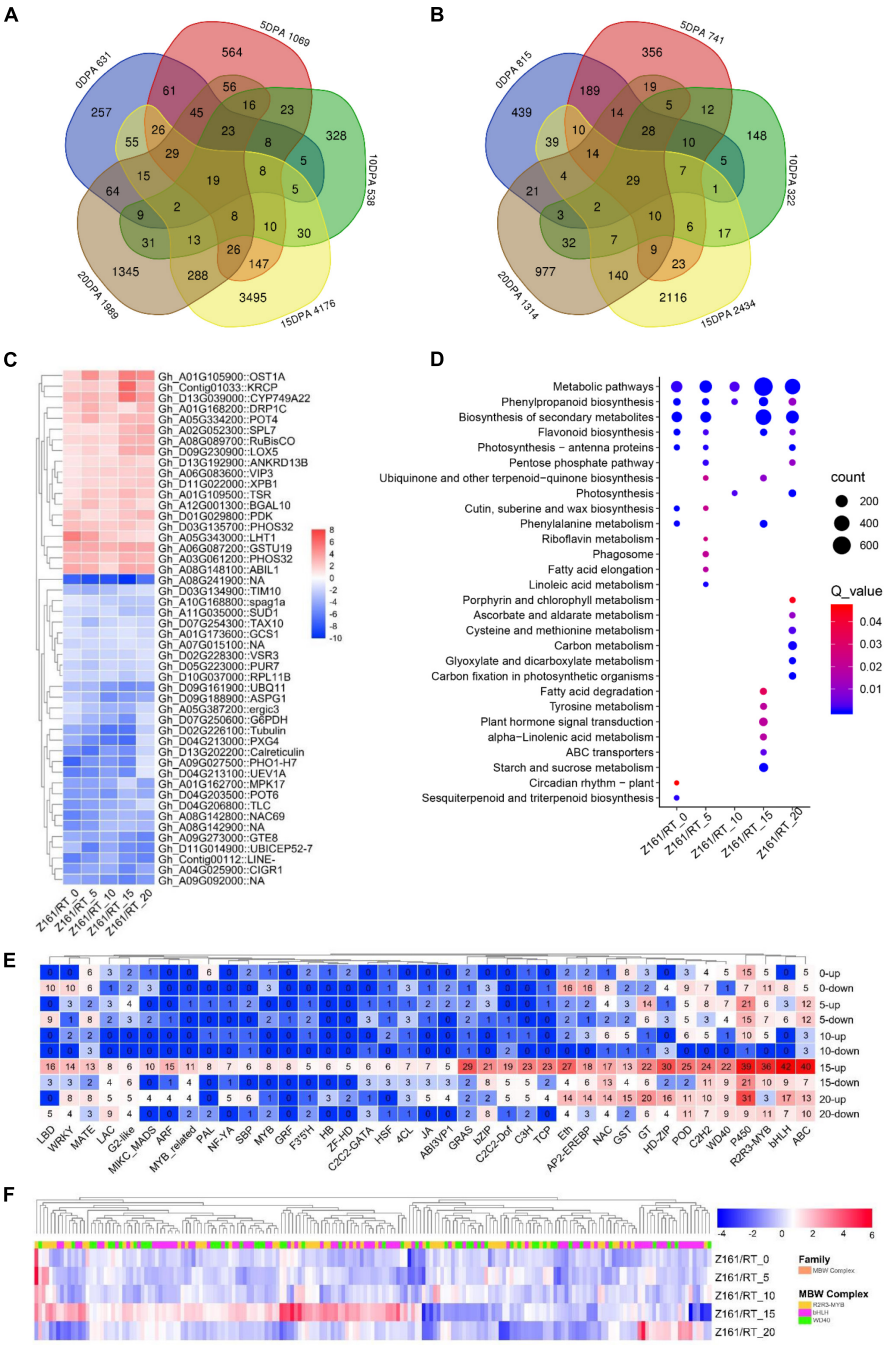


FIGURE 2 | Transcriptomic profiles in Z161 vs. RT and key genes identified in FB. Venn diagram showing the common upregulated **(A)** and downregulated **(B)** DEGs between Z161 and RT among the five stages. **(C)** Heatmap of 19 and 29 DEGs which were continuously upregulated or downregulated between the Z161 and RT during the fiber development stages. **(D)** Analysis of the KEGG enrichment of DEGs in different stages. **(E)** The number of upregulated and downregulated DEGs corresponding to different family members (including TF and structural genes) related to FB. **(F)** Heatmap of 185 MBW complexes of DEGs between Z161 and RT at the five stages. All information for these DEGs can be found in **Supplementary Tables 4**, **5**.

developmental stages (**Figure 1D**). A Venn diagram analysis was used to illustrate the distribution of upregulated and downregulated DEGs and their overlap between the Z161 and RT genotypes (**Figures 2A,B**). This revealed 19 and 29 common DEGs continuously upregulated and downregulated in the five stages, respectively (**Supplementary Table 9**). Among them, the upregulated genes were cytochrome P450 (Gh_D13G039000), squamosa promoter-binding-like protein (Gh_A02G052300), WD repeat-containing protein (Gh_A06G083600), glycosyltransferase protein (Gh_A01G105900), glutathione

S-transferase (Gh_A06G087200), ABI interactor-like protein (Gh_A08G148100), etc., whereas the downaccumulated genes were cysteine synthase (Gh_A08G241900), probable peroxygenase (Gh_D04G213000), tubulin (Gh_D02G226100) phosphate transporter (Gh_A09G027500), polyubiquitin (Gh_D09G161900), etc (**Figure 2C**).

To better understand the enriched pathways related to the biosynthesis and metabolism of DEGs, a KEGG pathway enrichment analysis was performed using an FDR < 0.05. In this way, 28 significantly enriched KEGG pathways (**Figure 2D** and**Supplementary Table 10**). Evidently, the top-five KEGG pathways were phenylpropanoid biosynthesis, FB, starch and sucrose metabolism, plant hormone signal transduction, and phenylalanine metabolism except for metabolic pathways and secondary metabolite biosynthesis, which suggested their involvement in fiber pigmentation and development at 15 DPA in both genotypes.

Next, we sought to identify and screen the specific members of the structural genes and regulatory genes in the pigment biosynthesis pathway of brown cotton during its fiber development stage. To do this, we first identified all the structural genes and MBW complexes in the proanthocyanidin biosynthesis pathway, and also the genome-wide TF family members. Then, we cross analyzed with the DEGs in the five stages between Z161 vs. RT, and the total of 1,568 DEGs' information was obtained (**Supplementary Table 11**). It was found that the top-ten families were bHLH, ABC, P450, R2R3-MYB, HD-ZIP, GRAS, Eth, POD, C2H2, and C3H, especially at the 15 and 20 DPA stages (**Figure 2E**). Multiple members of the MBW complex were observed. Other research has shown that the R2R3-MYB, bHLH, and WD40 TFs and MBW complexes play an important role in the biosynthesis of anthocyanins and proanthocyanidins (Xu et al., 2015). Compared with the white cotton RT, the MBW family members of brown cotton had more DEGs at 15 DPA, with some members having a greater abundance of DEGs at other stages (**Figure 2F**). These results confirmed that 15 and 20 DPAs during the process of BCF development were the key periods of active transcription of genes regulated by the pathway of pigment biosynthesis.

Comprehensive Analysis of the Transcriptome and Metabolome of Cotton

The untargeted metabolomic data were first used for principal component analysis (PCA). Six groups were clearly separated on the PC1 × PC2 score plot with neg- and pos-model (**Figures 3A,B**). Our results showed that the raw data from the comparisons of the two cotton lines' comparison at the same stage could reflect the significant differences between the samples. The PLS-DA method (Wen et al., 2017) was used for multivariate analysis of the six comparisons, to identify those DAMs with statistical and biological significance and thereby elucidate the related metabolic processes. The number of DAMs showed a similar trend in both two ion modes. At 10 and 20 DPA (fiber elongation phase), the accumulation level of metabolites decreased more than that at 0 DPA (fiber initiation phase). At the same time, the accumulation level of metabolites increased more among different genotypes at 20 DPA (**Figure 3C**). The results also uncovered the highest number of specific upregulated and downregulated metabolites in the overlapping phases of fiber elongation and secondary wall thickening (20 DPA), suggesting that this was the most active period for the deposition of various metabolites of brown cotton and white cotton (**Figure 3D**). To further investigate the DAMs with DEG-enriched pathways, a cross analysis of the DEG-enriched KEGG pathways (17) and the corresponding pathways of neg (90) and pos- (72) DAMs were mapped, with common 13 metabolic pathways (**Figure 3E** and **Supplementary Table 12**). As shown in **Figure 3F**, the top-five KEGG pathways were phenylpropanoid metabolism, FB, phenylalanine metabolism, carbon metabolism, and cysteine and methionine metabolism, which suggested that these DEGs in both genotypes were mainly enriched in five pathways compared with different development stages at the metabolomic and transcriptional levels. The heatmap for the differences in flavonoid content among the three comparisons indicated the greater levels of flavanols in the fiber of Z161 than in RT at 10 and 20 DPA (**Figure 3G** and **Supplementary Figure 3**). To sum up, these results indicated that gene expression and related metabolite production in flavonoid pathways were more active from the early stage of fiber initiation development stage.

Identification of Key Genes and Regulators Related to Proanthocyanidins Biosynthesis in Brown Cotton

To investigate the gene coexpression and regulatory network of proanthocyanidins biosynthesis in brown cotton, WGCNA was conducted using 10,891 filtered DEGs (**Supplementary Table 8**). Overall, 14 distinct coexpression modules corresponding to clusters of correlated genes were detected (**Figure 4A** and **Supplementary Table 13**). Each stage of the cotton samples and the fold-change of different flavonoids content of the three comparisons of metabolome behavior were used as trait data for a module–trait relationship analysis (**Figure 4B**). The MElightyellow module (109 genes) presented the strongest correlation with Z161-15 (r = 0.99, p = 2e-22) whereas the MEmagenta modules (239) were most corrected with Z161-20 (r = 0.98, p = 1e-19). Also, selected were MEdarkorange and MEsaddlebrown (82 genes), and MEcyan (499) and MEdarkgreen (95), based on the different stage samples of Z161 and the different differing abundance of flavonoids with *r-values* > 0.5. A summary profile for each module was based on the average value of FPKM from each sample, and that highly correlated traits are shown in **Figure 4C**. Furthermore, the KEGG pathway enrichment analysis of the above six modules showed that the FB and phenylpropanoid metabolism were significantly enriched in the MEcyan and MElightyellow module (**Figure 4D**). Except for the MBW complex, 34 TF genes of 20 families were screened, most of which were expressed more in 15 DPA than in other stages (**Figure 4E** and **Supplementary Table 14**). Among them, *GhBSD* (Gh_D08G073300) gene was highly expressed in all five stages of Z161 (0–20 DPA).

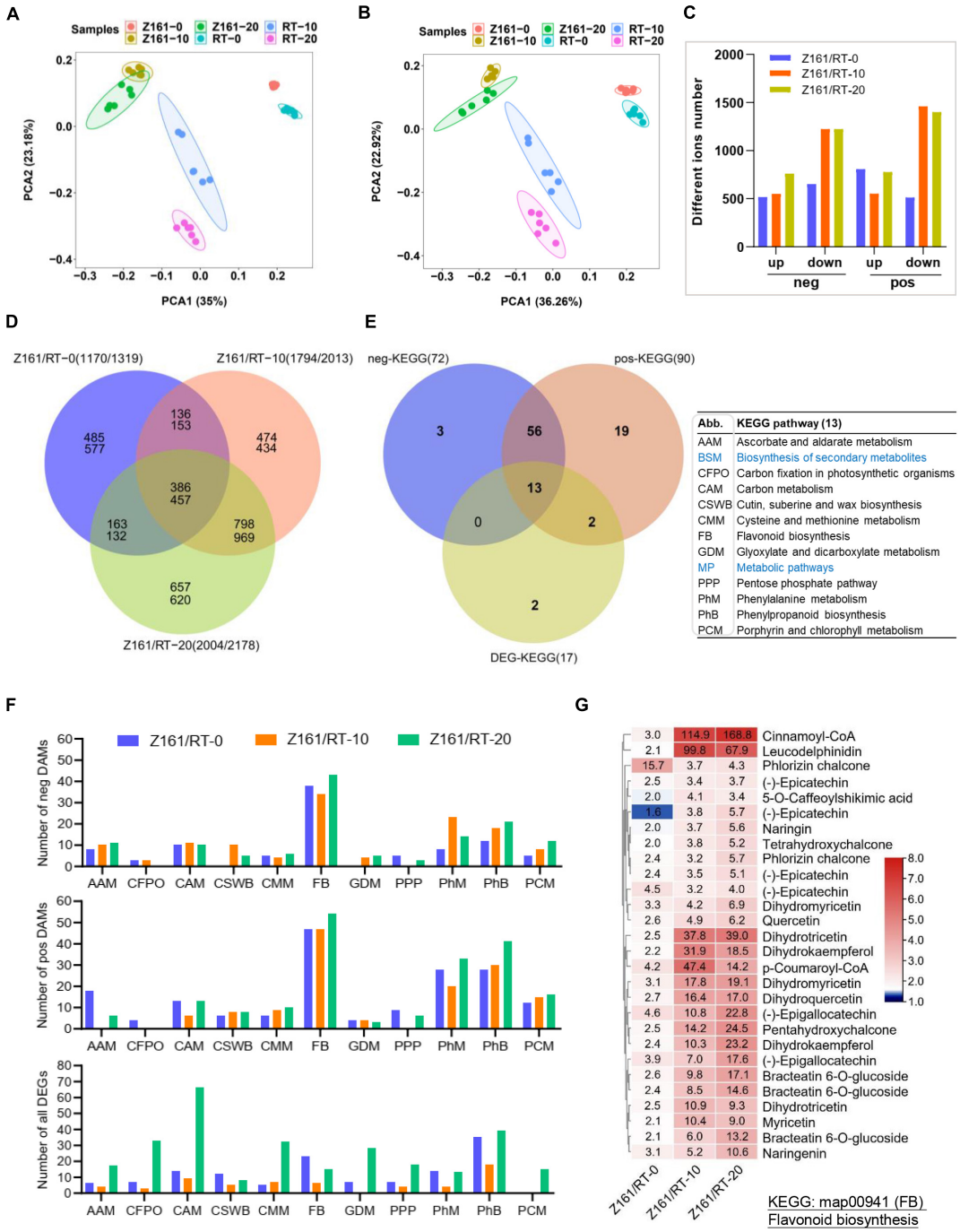


FIGURE 3 | Association analysis between transcriptome and metabolome during fiber development in Z161 vs. RT. PCA of samples (with six biological repetitions) at 0, 10, and 20 DPA, between Z161 and RT in the neg- **(A)** and pos- **(B)** ion mode. **(C)** Numbers of upregulated and downregulated DAMs in negative and positive ion mode between Z161 and RT at the same stage. **(D)** Venn diagram showing the common of upregulated and downregulated DAMs between Z161 and RT. **(E)** Venn diagram showing the unique and common KEGG pathways and all the corresponding pos- and neg-DAMs mapped to the KEGG pathways between Z161 and RT. The 13 pathways abbreviations are explained in the accompanying table. **(F)** Numbers of DAMs and DEGs from these 11 metabolic pathways in the three comparisons. **(G)** Heatmap of 28 representative DAMs in FB. Blue labeled metabolic pathways are the two largest classifications.

To understand the differential expression of structural genes and regulatory genes in the FB pathway, 290 genes from 15 catalytic enzyme or transporter families and MYB-bHLH-WD40 (total 62 genes) from the above five ME modules were identified (**Figure 5A** and **Supplementary Table 15**). It is worth mentioning that the two *GhANS* (Gh_A08G1593) and *GhANR*

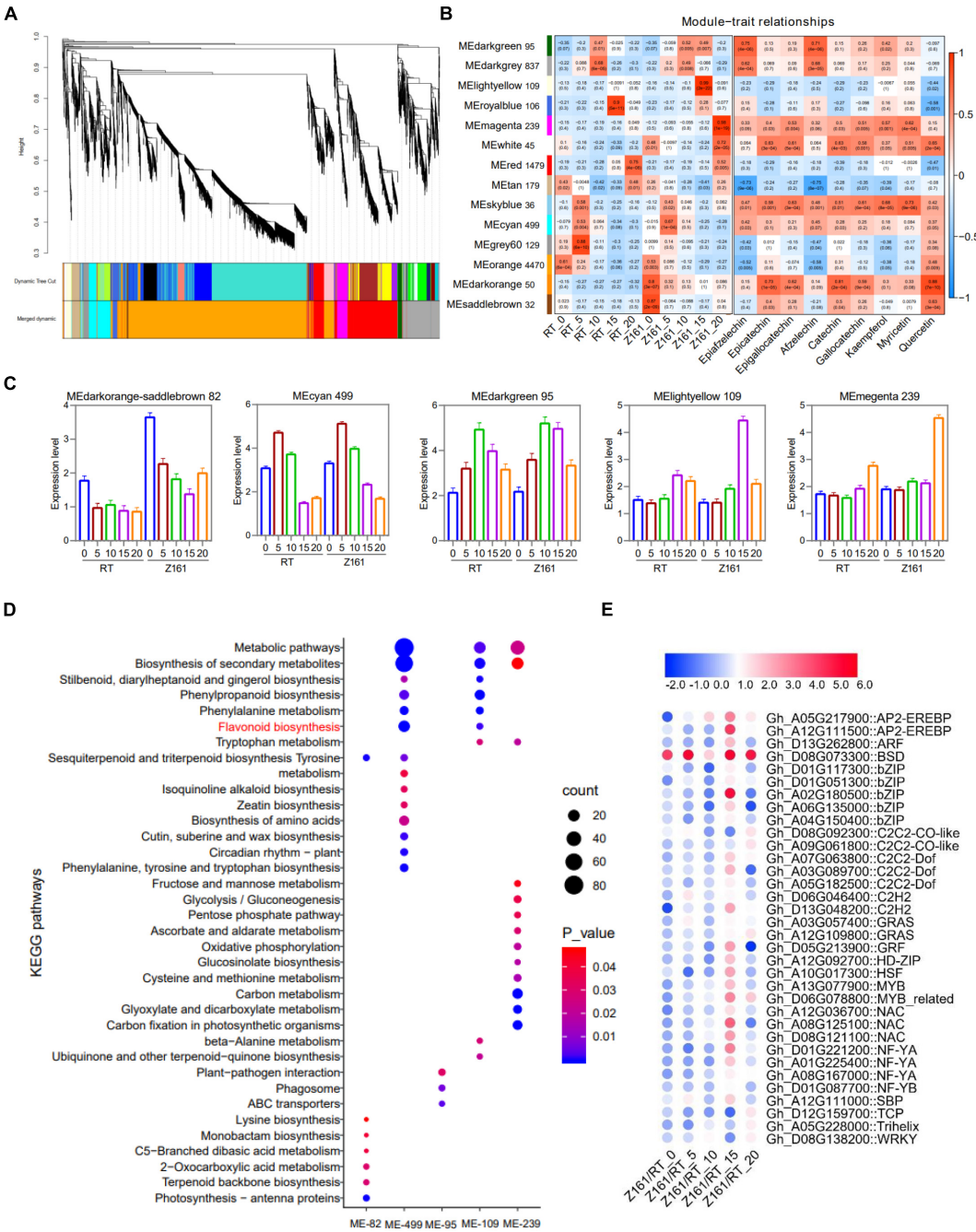


FIGURE 4 | Weighted gene coexpression network analysis (WGCNA) for DEGs. **(A)** Hierarchical cluster dendrogram. **(B)** Module–trait relationships with five stage samples of two varieties (Z161 and RT) and flavonoid contents. **(C)** Eigengene expression profile for the six modules in five-stage samples of two varieties (Z161 and RT). **(D)** KEGG enrichment for the six modules genes. **(E)** Expression pattern analysis of TF genes from the six modules. All information for these genes can be found in **Supplementary Tables 11**, **12**. The metabolic pathway marked by red is a metabolic pathway related to pigment formation.

(Gh_A05G1424) genes were also highly expressed in the Z161 proteome data (see Figure 5A from Peng et al., 2020). In tandem, we also identified some important glutathione-S-transferase (GST), multidrug and toxic compound extrusion (MATE), laccase (LAC), ABC transporter (ABC), and glucosyltransferase (GT) genes and they play a key role in the process of transporting proanthocyanidins to vacuoles for further condensation and polymerization. According to the above five modules, we tried searching for the coexpression networks of the R2R3-MYB and bHLH hub genes, thereby finding three important networks (**Figure 5B** and **Supplementary Table 16**). In MEmagenta, *GhMYB1* (Gh_D08G194300), *GhbHLH1* (Gh_A10G023700), and

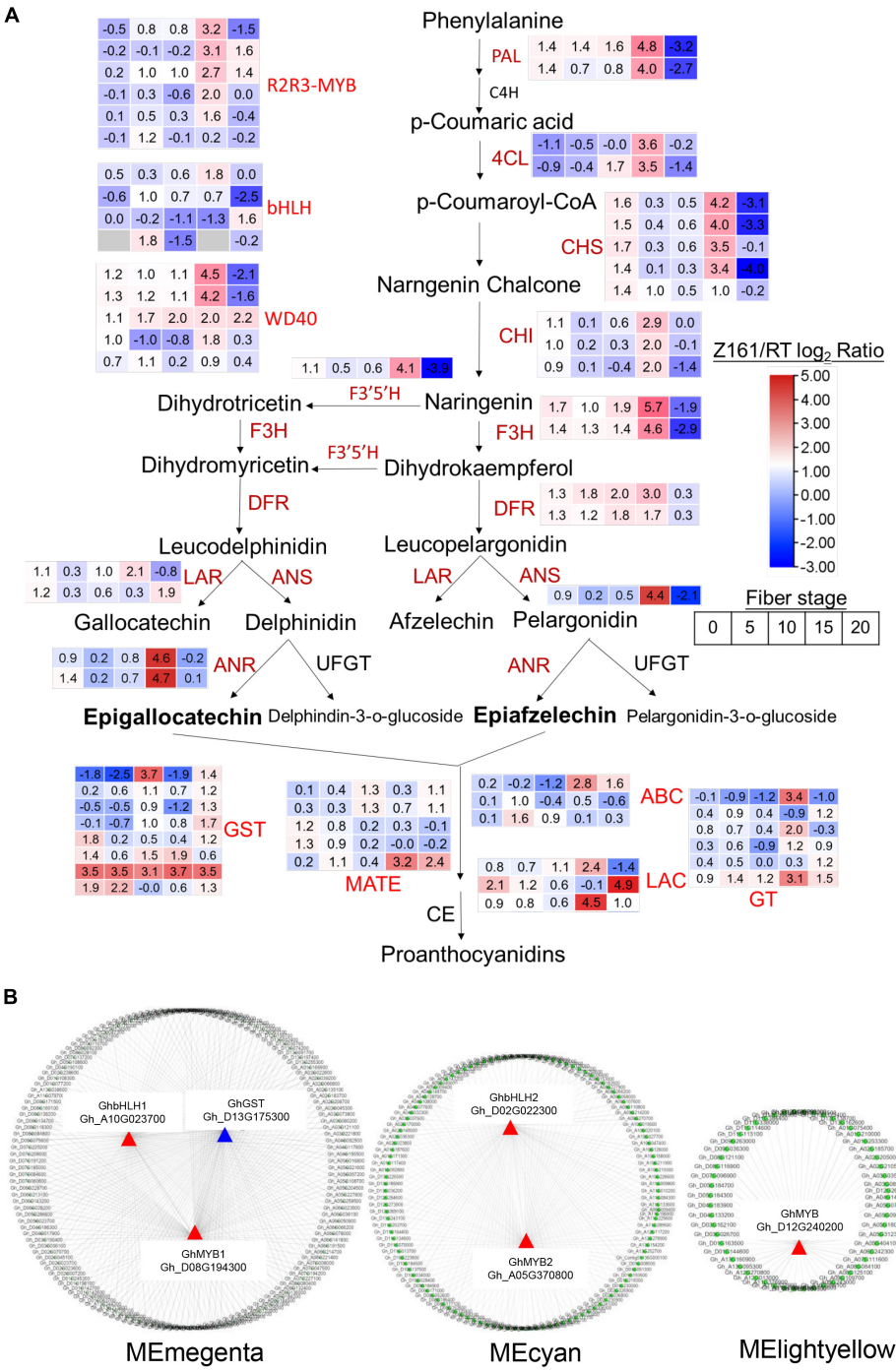


FIGURE 5 | Analysis of the structural and TF DEGs for FB in Z161 vs. RT. The pathways were constructed based on the KEGG pathways and literature sources. **(A)** Red fonts indicate upregulated genes (catalytic enzymes and TF) in Z161 during at least one stage of fiber development. Gene expression patterns are displayed in the box with a red–blue color model. FB (map00941). PAL, phenylalanine ammonia-lyase; C4H, cinnamate 4-hydroxylase; 4CL, 4-coumarate-CoA ligase; CHI, chalcone isomerase; CHS, chalcone synthase; F3′5′H, flavonoid 3′,5′-hydroxylase; F3H, flavanone 3-hydroxylase; F3′H, flavonoid 3′-monooxygenase; DFR, dihydroflavonol-4-reductase; LAR, leucoanthocyanidin reductase; ANS, anthocyanidin synthase; ANR, anthocyanidin reductase; UFGT, UDP flavonoid glucosyl transferase; GT, anthocyanidin 3-O-glucosyltransferase; GST, glutathione S-transferase, MATE, Multidrug and toxic compound extrusion or multiantimicrobial extrusion; ABC, ATP-binding cassette; LAC, laccase; GT, glycosyltransferase; CE, condensation enzyme. **(B)** Cystoscope representation of coexpressed genes with edge weight ≥0.20 in "magenta," "cyan," "lightyellow" modules. Coloring shows the node R2R3-MYB or bHLH genes. All the DEGs are shown in **Supplementary Table 14**.

GhGST (Gh_D13G175300) were the most correlated genes. In MEcyan, the *GhMYB2* (Gh_A05G370800) and *GhbHLH2* (Gh_D02G022300) genes were the most correlated genes. However, there was only one R2R3-MYB gene *GhMYB* (Gh_D12G240200) that correlated most with other genes in MElightyellow. To verify the reliability of our cotton RNA-seq data, we examined 20 genes possibly involved in the biosynthesis of phenylpropanoid and proanthocyanidins, and their results were also consistent with those of their qRT-PCR analysis (**Figure 6**).

To conclude, taken together, these results confirmed that the expression levels of some key regulatory genes and structural genes of the flavonoid pathway are far higher in Z161 than RT at 15 and 20 DPA, leading to rapid pigment biosynthesis and polymerization at the end stage of the fiber elongation or secondary wall thickening stage. The coexpression patterns of R2R3-MYB and bHLH with other genes at distinct stages may also play a key regulatory role in cotton fiber pigmentation.

Three R2R3-MYB Proteins Bind to the Promoters of *GhANS, GhANR1,* and *GhUFGT2* Genes

In our previous study, proteomics distinguished the key proteins *GhANR1* (Gh_A05G1424) and *GhANS* (Gh_A08G1593) as being highly abundant at 10 and 20 DPA in brown cotton Z161. Meanwhile, the qRT-PCR showed that *GhUFG11T2* was expressed from −1 to 30 DPA, and the expression level of 25 and 30 DPA in Z161 was higher than that in other stages (Peng et al., 2020). Likewise, in this study, the expression of three genes – *GhANR1* (Gh_A05G152700), *GhANR2* (Gh_D05G168500), and *GhANS* (Gh_D08G197600) – underwent more than a 16-fold change in Z161 vs. RT at transcriptional level at 15 DPA. Interestingly, both the identified genes (*GhANR1* and *GhANS*) exhibited the same expression trend as inferred from the proteome analysis (**Figure 5A**). To verify whether the five key genes of proanthocyanidin biosynthesis may be transcriptionally regulated by R2R3-MYB TF, the possible MYB-binding motifs in the promoter region of these genes were first predicted (i.e., the 2K sequence upstream of the gene transcript initiation sites). Later, the MYB recognition site [CNGTT(A/G)], MYB motif [T(C)AACCA], MYB-binding site (CAACAG), and MYB-like sequence (TAACCA) were identified in the promoter region (**Figure 7A**). The binding ability to the promoters of key structural genes *GhANR1/2*, *GhANS,* and *GhUFGT1/2* was then investigated using the yeast one-hybrid (Y1H) assay.

The ORF of GhMYB6 was inserted into MCS of pGADT7 vector to serve as the prey construct, which was cotransformed with the individual bait constructs harboring different MYB-binding motifs (these cloned from *GhANR1/2*, *GhANS*, and *GhUFGT1/2* promoter sequences) fusion proteins. This demonstrated that GhMYB6 could only physically interact with *GhANS* (**Figure 7B**), but no physical interactions occurred between GhMYB6 and *GhANR1*, *GhANR2*, or *GhUFGT1/2* (data not shown). In checking for possible interactions *via* different comparisons, it was finally determined that *GhMYB3* and *GhMYB46* can only physically interact with *GhUFGT2* and *GhANR1,* respectively (**Figure 7B**). Expression trends of the *GhANR1*, *GhANS*, and *GhUFGT2* genes and their corresponding regulatory genes —*GhMYB46*, *GhMYB6*, and *GhMYB3* —during fiber development were analyzed by qRT-PCR (**Figure 7C**). Three R2R3-MYB proteins could bind to the promoters of *GhANS*, *GhANR1,* and *GhUFGT2*. It is worth noting that GhMYB6 might negatively regulate the transcription of the *GhANS* gene, but this needs to be verified by further research.

DISCUSSION

The color of NCC results from the accumulation of pigments during the differentiation and development of fiber cells. Previous research has mainly focused on the fiber pigmentation (at ∼20 DPA; fiber elongation and secondary wall thickening). In this study, the data for the transcription and metabolism of white and brown fibers were analyzed in the early and late elongation stage and during the stage of secondary wall thickening. These studies are helpful for exploring the regulatory network mechanism of TFs of the proanthocyanidins biosynthesis pathway in brown cotton plants.

Key Period of Transcription Regulation of Pigmentation-Related Genes in Brown Cotton Fiber Development Occurs From 15 to 20 Days Post-anthesis

As is well known, both white cotton and colored cotton go through four stages of fiber development: differentiation and protrusion of fiber primordial cells, elongation of primary wall, thickening of the secondary wall, and dehydration and maturation. But colored cotton undergoes an additional process of pigment synthesis and deposition (Murthy, 2001). Many recent studies have shown that PAs are the main component of pigment deposition in BCF (Li et al., 2013; Feng et al., 2014; Xiao et al., 2014; Yan et al., 2018), and their content reaches typically peaks at 30 DPA (Feng et al., 2014). The biosynthesis of proanthocyanidins precursors regulates the FB pathway to mediate the gene expression of flavonoid metabolism pathway for the development of BCF (Hua et al., 2007; Xiao et al., 2007, 2014; Feng et al., 2013; Li et al., 2013; Tan et al., 2013; Gao et al., 2019). Earlier research demonstrated that the synthesis of proanthocyanidins in fiber is the main reason for the shortened elongation period, thus halting the early elongation of BCF (Qiu, 2004). The dominant expression of genes related to the biosynthesis of proanthocyanidins precursors will play a vital role in the fiber development stage. In this study, we compared the transcriptome of brown cotton and its white cotton NILs at 0, 5, 10, 15, and 20 DPAs. The results showed that 15 and 20 DPAs corresponded to the peak period for the dominant expression of genes related to the FB pathway during fiber development of brown cotton. These above said findings were substantiated by the analysis of differentially expressed structural genes and TFs (triple complex) related to FB and by the weighted gene coexpression network. Xiao et al. (2007) were able to detect

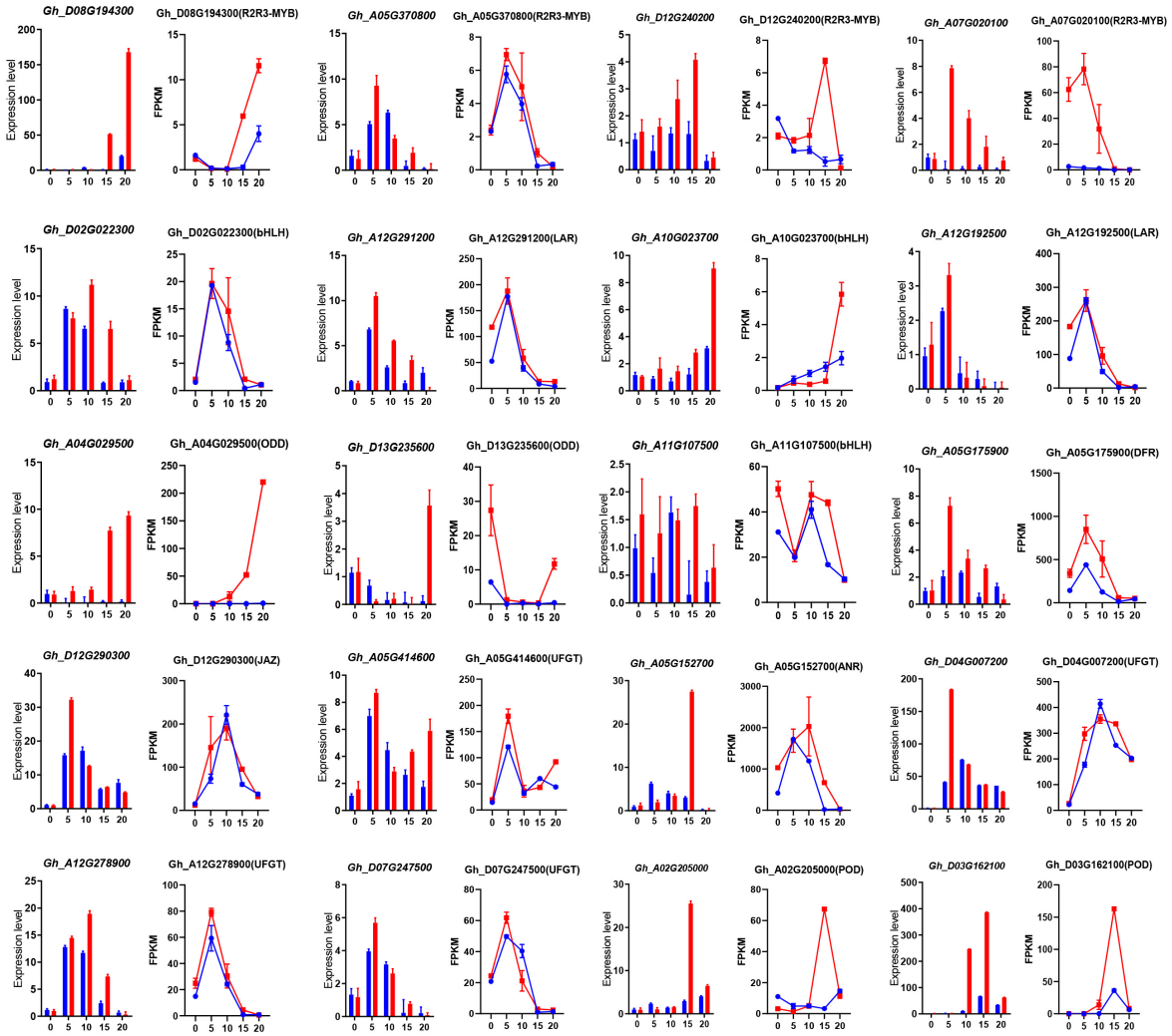


FIGURE 6 | The qRT-PCR analysis of the expression of 20 pigmentation-related candidate genes between Z161 and RT at the five stages. Relative expression levels are presented as mean ± SD, $n = 3$. The corresponding FPKM values were derived from the RNA-seq data. All the primers are shown in **Supplementary Table 5**. Red and blue pillars or lines represent Z161 and RT, respectively.

key genes in cotton by semiquantitative PCR, finding their greatest expression at 16 DPA and decreased at 20 DPA. Likewise, Canavar and Rausher (2021) recently reported that structural gene expression reached its peak at 14 DPA. Additionally, Peng et al. (2020) revealed that 20 DPA is the peak time for the accumulation of important proteins and metabolites in the biosynthesis pathway of flavonoids. Therefore, the proteins and metabolites accumulated at 20 DPA may be the genes highly expressed at 15 DPA in the posttranscription, translation, and posttranslation stages.

Gene Regulatory Network Related to Flavonoid Biosynthesis Pathway During Brown Cotton Fiber Development

Few reports exist that comparatively analyze main flavonoids metabolites in the fibrous tissues of brown cotton and white cotton at different developmental stages. In this study, all metabolites in the flavonoid pathway were identified and quantitatively expressed by a non-targeted metabolome analysis. Comparative analysis revealed that the main metabolites (i.e., flavanediol and flavane-3-alcohol) accumulated substantially in BCF. Previous studies have shown that the final pigment component in BCF is mainly PA. Flavane-3-ol is the precursor unit (flavonoids) of proanthocyanidins polymerization (Feng et al., 2014), and we further confirmed that a large number of pigment precursors (flavonoids) were accumulated during the BCF elongation stage of brown cotton (10–20 DPA). Previously, Feng et al. (2013) had shown that the accumulation of naringenin, quercetin, kaempferol, and myricetin in brown fiber at 12–21 DPA significantly exceeded that in white cotton fiber by HTLC analysis. That finding is consistent with our metabolome results.

Further, we explored differences in gene expression in five developmental stages *via* paired comparison of transcriptome

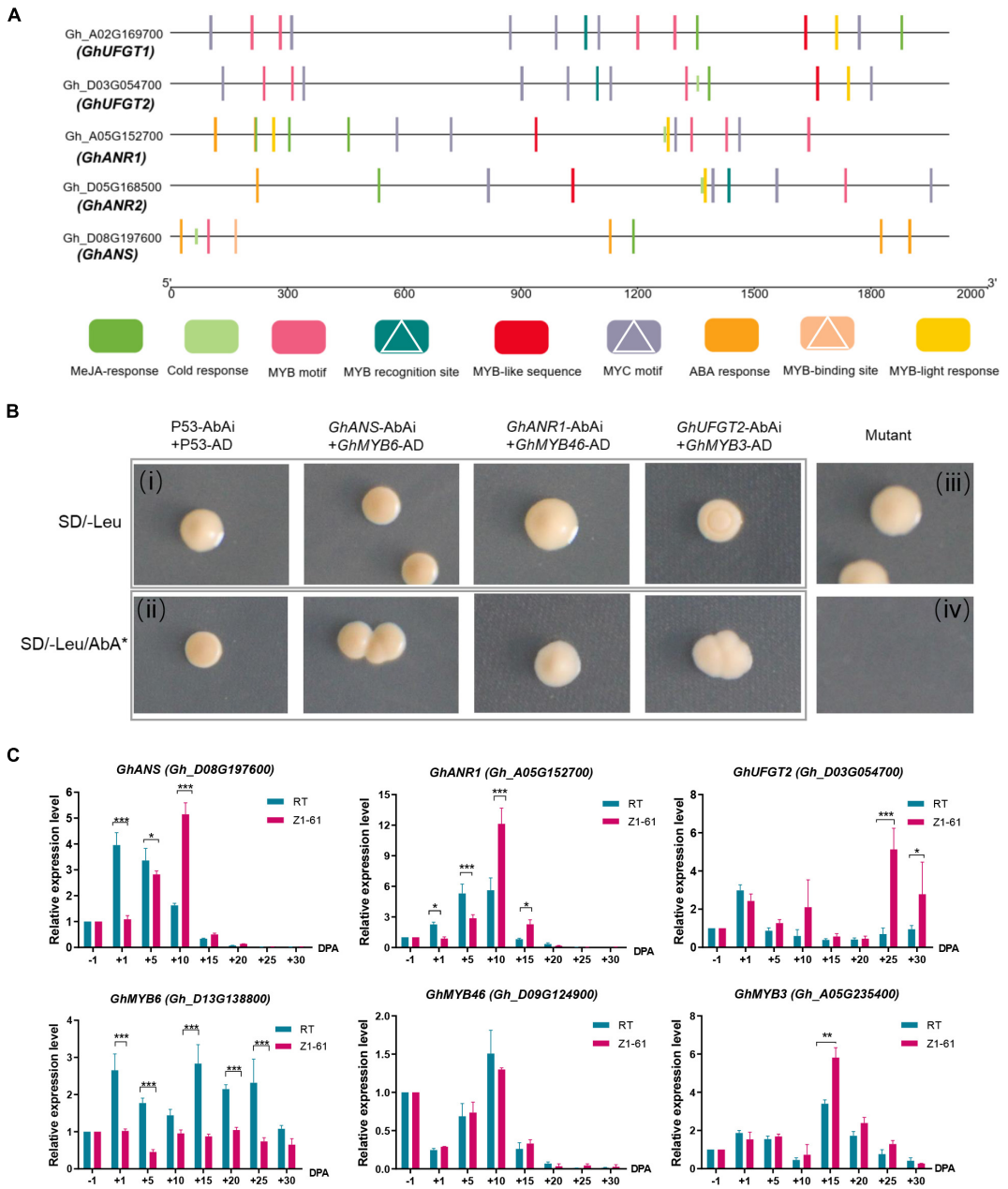


FIGURE 7 | Validation of the interaction of the transcriptional regulatory complex by yeast one-hybrid (Y1H) assays and qRT-PCR. **(A)** Schematic representation of MYB motifs in the promoter regions of the flavonoids biosynthetic genes (*GhANR1*, *GhANR2*, *GhUFGT1*, *GhUFGT2, and GhANS*) from cotton. MYB recognition sites, MYB-binding sites, and MYB motifs are indicated by white triangles. **(B)** Y1H analysis of *GhMYB6*, *GhMYB46*, *GhMYB3*, interactions with the canonical MYB-binding motif, respectively, identified in *GhANS*, *GhANR1,* and *GhUFGT2* promoter sequences, respectively. The **(i,ii)** are combined as a positive control, whereas **(iii,iv)** are combined as a negative control. SD, minimal media; Leu, leucine; AbA, Aureobasidin A. **(C)** qRT-PCR analysis of transcriptional levels of *GhMYB6*, *GhMYB46*, *GhMYB3, GhANS*, *GhANR1,* and *GhUFGT2* genes between Z161 and RT at these fiber development stages (–1, 1, 5, 10, 15, 20, 25, and 30 DPA). All the primers are shown in **Supplementary Table 4**. The significant difference is Student's *t* test; $^{*}p < 0.05$; $^{**}p < 0.01$; $^{***}p < 0.001$.

data. The results of the DEG analysis in different periods demonstrated that some specifically expressed genes involved in flavonoid synthesis pathway were enriched in the four periods at 0, 5, 15, and 20 DPA. It should be noted that the number of upregulated genes of different gene family members involved in the flavonoid synthesis pathway during fiber elongation (5–20 DPA) surpassed that of downregulated genes. In particular, the expression levels of structural genes involved in the flavonoid pathway (e.g., *GhF3H*, *GhLAR*, and *GhANR*) were significantly higher in the BCF samples. These patterns of gene expression patterns are consistent with the findings of Li et al. (2020) and Canavar and Rausher (2021).

Although there are many studies of flavonoid synthesis and its functional genes, the complex regulatory mechanism is still a mystery. Here, the WGCNA method was used to detect the gene interaction network involved in FB. We screened five R2R3-MYB and bHLH hub genes to better understand the *in vivo* relationships between genes during the development of BCF and flavonoid synthesis in the two cotton varieties. In this respect, *GhMYB1* (Gh_D08G194300) and *GhMYB3* (Gh_D12G240200) are homologous to *AtMYB42*, which can directly activate transcription inhibitors, thereby specifically inhibiting FB and directly activating lignin biosynthesis genes (Geng et al., 2020). However, in our study, the gene was upregulated at 15 and 20 DPA in brown cotton (Z161) and is thus speculated to harbor different functions. Building on our previous work (Peng et al., 2020), we also verified that the potential interacting TFs of GhANR1 (homologous gene *GhANR2*), *GhANS,* and *GhUFGT2* (homologous gene *GhUFGT1*) were *GhMYB46*, *GhMYB6,* and *GhMYB3*, respectively. According to their expression patterns, we may infer that *GhMYB6* is a negative regulator and *GhMYB3* is a positive regulator. We anticipate this study's compelling evidence, and findings can be used to reveal the complex transcriptional regulation mechanism of flavonoid synthesis in BCF in future work.

CONCLUSION

In this study, 10,891 DEGs related to the pigmentation of the brown cotton genotypes were detected. It was found that the top-three KEGG pathways at 15 DPA were phenylpropanoid biosynthesis, FB, and starch and sucrose metabolism. According to the DAMs and DEGs distinguishable in different fiber development stages, 13 common key KEGG pathways were identified, which included phenylpropanoid biosynthesis and FB. Compared with RT, the flavonoids contents in Z161 at different stages varied greatly, such as for leucodelphinidin, leucocyanidin, kaempferol, epiafzelechin, epicatechin, and epigallocatechin. Based on the WGCNA, three modules with five TF hub genes were found that were highly correlated with the FB pathway. Furthermore, the three genes *GhANS*, *GhANR1,* and *GhUFGT2* are preliminarily proven to interact with three R2R3-MYB TFs (*GhMYB46*, *GhMYB6*, and *GhMYB3*). The candidate genes and interaction patterns of PAs accumulation in brown cotton elucidated here can provide valuable clues for improving brown cotton's molecular breeding in the near future.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are publicly available. This data can be found here: https://www.ncbi.nlm.nih.gov/bioproject/PRJNA766762.

AUTHOR CONTRIBUTIONS

ZW analyzed and summed all the data and wrote the manuscript. XZ designed the yeast one-hybrid experiment and finished the operation. SH directed the methodology and software analysis. XG, QG, and LW managed and collected the plant tissues. YJ, HL, and ZhaP collected the data and evidence and project administration. AR revised the language. ZheP and XD conceptualized the research program and revised the manuscript. All authors read and approved the final manuscript.

FUNDING

This work was supported by the National Natural Science Foundation of China (31601353) and Agricultural Science and Technology Innovation Program of Chinese Academy of Agricultural Sciences (CAAS-ASTIP-ICR-2021-01).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2021.822198/full#supplementary-material

Supplementary Table 1 | The different abundance metabolites (DAMs) among negative and positive ions at 0 DPA.

Supplementary Table 2 | The DAMs among negative and positive ions at 10 DPA.

Supplementary Table 3 | The DAMs among negative and positive ions at 20 DPA.

Supplementary Table 4 | Primer sequences used in yeast one-hybrid analysis.

Supplementary Table 5 | Primer sequences used in qPCR analysis.

Supplementary Table 6 | RNA-seq data statistic.

Supplementary Table 7 | Mapping results of all 30 RNA-seq libraries to the *Gossypium hirsutum* reference genome.

Supplementary Table 8 | The detailed information of DEGs.

Supplementary Table 9 | The detailed information of 19 and 29 DEGs which were continuously upregulated or downregulated between the Z161 and RT.

Supplementary Table 10 | DEGs enriched in KEGG pathways.

Supplementary Table 11 | The detailed information of 1,568 DEGs (including TF and structure genes) related to FB.

Supplementary Table 12 | All DAMs and DEGs enriched in KEGG pathways.

Supplementary Table 13 | The detailed information of 8,305 genes of each sample/metabolite-specific module.

Supplementary Table 14 | The detailed information of TF genes (exclude MBW genes) of each sample/metabolite-specific module.

Supplementary Table 15 | The detailed information of 290 structural and TF genes related to FB of each sample/metabolite-specific module.

Supplementary Table 16 | The hub genes of "magenta," "cyan," "lightyellow" module.

REFERENCES

Bailey, T. L., Boden, M., Buske, F. A., Frith, M., Grant, C. E., Clementi, L., et al. (2009). MEME SUITE: tools for motif discovery and searching. *Nucleic Acids Res.* 37(Suppl._2), W202–W208. doi: 10.1093/nar/gkp335

Benjamini, Y., and Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. R. Stat. Soc. Ser. B* 57, 289–300. doi: 10.1111/j.2517-6161.1995.tb02031.x

Canavar, Ö, and Rausher, M. D. (2021). Molecular analysis of structural genes involved in flavonoids biosynthesis in naturally colored cotton. *Crop Sci.* 61, 1117–1126.

Carvalho, L. P. D., Farias, F. J. C., Lima, M. M. D. A., and Rodrigues, J. I. D. S. (2014). Inheritance of different fiber colors in cotton (*Gossypium barbadense* L.). *Crop Breed. Appl. Biotechnol.* 14, 256–260. doi: 10.1590/1984-70332014v14n4n40

Chen, C., Chen, H., Zhang, Y., Thomas, H. R., Frank, M. H., He, Y., et al. (2020). TBtools: an integrative toolkit developed for interactive analyses of big biological data. *Mol. Plant* 13, 1194–1202. doi: 10.1016/j.molp.2020.06.009

Dutt, Y., Wang, X., Zhu, Y., and Li, Y. (2004). Breeding for high yield and fibre quality in coloured cotton. *Plant Breed.* 123, 145–151. doi: 10.1016/j.tplants.2021.03.007

Efe, L., Killi, F., and Mustafayev, S. A. (2009). An evaluation of eco-friendly naturally coloured cottons regarding seed cotton yield, yield components and major lint quality traits under conditions of East Mediterranean region of Turkey. *Pakistan J. Biol. Sci.* 12:1346. doi: 10.3923/pjbs.2009.1346.1352

Feng, H., Guo, L., Wang, G., Sun, J., Pan, Z., He, S., et al. (2015). The negative correlation between fiber color and quality traits revealed by QTL analysis. *PLoS One* 10:e0129490. doi: 10.1371/journal.pone.0129490

Feng, H., Li, Y., Wang, S., Zhang, L., Liu, Y., Xue, F., et al. (2014). Molecular analysis of proanthocyanidins related to pigmentation in brown cotton fibre (*Gossypium hirsutum* L.). *J. Exp. Bot.* 65, 5759–5769. doi: 10.1093/jxb/eru286

Feng, H., Tian, X., Liu, Y., Li, Y., Zhang, X., Jones, B. J., et al. (2013). Analysis of flavonoids and the flavonoid structural genes in brown fiber of upland cotton. *PLoS One* 8:e58820. doi: 10.1371/journal.pone.0058820

Feng, H. J., Sun, J. L., Wang, J., Jia, Y. H., Zhang, X. Y., Pang, B. Y., et al. (2011). Genetic effects and heterosis of the fibre colour and quality of brown cotton (*Gossypium hirsutum*). *Plant Breed.* 130, 450–456.

Gao, J., Shen, L., Yuan, J., Zheng, H., Su, Q., Yang, W., et al. (2019). Functional analysis of GhCHS, GhANR and GhLAR in colored fiber formation of *Gossypium hirsutum* L. *BMC Plant Biol.* 19:455. doi: 10.1186/s12870-019-2065-7

Geng, P., Zhang, S., Liu, J., Zhao, C., Wu, J., Cao, Y., et al. (2020). MYB20, MYB42, MYB43, and MYB85 regulate phenylalanine and lignin biosynthesis during secondary cell wall formation. *Plant Physiol.* 182, 1272–1283. doi: 10.1104/pp.19.01070

Gong, W., He, S., Tian, J., Sun, J., Pan, Z., Jia, Y., et al. (2014). Comparison of the transcriptome between two cotton lines of different fiber color and quality. *PLoS One* 9:e112966. doi: 10.1371/journal.pone.0112966

Hinchliffe, D. J., Condon, B. D., Thyssen, G., Naoumkina, M., Madison, C. A., Reynolds, M., et al. (2016). The GhTT2_A07 gene is linked to the brown colour and natural flame retardancy phenotypes of Lc1 cotton (*Gossypium hirsutum* L.) fibres. *J. Exp. Bot.* 67, 5461–5471. doi: 10.1093/jxb/erw312

Hua, S., Wang, X., Yuan, S., Shao, M., Zhao, X., Zhu, S., et al. (2007). Characterization of pigmentation and cellulose synthesis in colored cotton fibers. *Crop Sci.* 47, 1540–1546.

Hua, S., Yuan, S., Shamsi, I. H., Zhao, X., Zhang, X., Liu, Y., et al. (2009). A comparison of three isolines of cotton differing in fiber color for yield, quality, and photosynthesis. *Crop Sci.* 49, 983–989.

Ibraheem, N. A., Hasan, M. M., Khan, R. Z., and Mishra, P. K. (2012). Understanding color models: a review. *ARPN J. Sci. Technol.* 2, 265–275.

Kanehisa, M., Araki, M., Goto, S., Hattori, M., Hirakawa, M., Itoh, M., et al. (2008). KEGG for linking genomes to life and the environment. *Nucleic Acids Res.* 36(Suppl._1), D480–D484. doi: 10.1093/nar/gkm882

Karaarslan, E. S., Mehmet, B., Ertan, E., Cebe, M. A., and Usumez, A. (2013). Assessment of changes in color and color parameters of light-cured composite resin after alternative polymerization methods. *Eur. J. Dentist.* 7, 110–116. doi: 10.1055/s-0039-1699004

Khatri, A., Peerzada, M. H., Mohsin, M., and White, M. (2015). A review on developments in dyeing cotton fabrics with reactive dyes for reducing effluent pollution. *J. Clean. Product.* 87, 50–57.

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

Kohel, R. (1985). Genetic analysis of fiber color variants in cotton 1. *Crop Sci.* 25, 793–797. doi: 10.2135/cropsci1985.0011183x0025000500017x

Langfelder, P., and Horvath, S. (2008). WGCNA: an R package for weighted correlation network analysis. *BMC Bioinform.* 9:559. doi: 10.1186/1471-2105-9-559

Langmead, B., and Salzberg, S. L. (2012). Fast gapped-read alignment with Bowtie 2. *Nat. Methods* 9, 357–359. doi: 10.1038/nmeth.1923

Li, B., and Dewey, C. N. (2011). RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinform.* 12:323. doi: 10.1186/1471-2105-12-323

Li, Y. J., Zhang, X. Y., Wang, F. X., Yang, C. L., Liu, F., Xia, G. X., et al. (2013). A comparative proteomic analysis provides insights into pigment biosynthesis in brown color fiber. *J. Proteomics* 78, 374–388. doi: 10.1016/j.jprot.2012.10.005

Li, Z., Su, Q., Xu, M., You, J., and You, C. (2020). Phenylpropanoid metabolism and pigmentation show divergent patterns between brown color and green color cottons as revealed by metabolic and gene expression analyses. *J. Cotton Res.* 3:27.

Livak, K. J., and Schmittgen, T. D. (2001). Analysis of relative gene expression data using real-time quantitative PCR and the 2- $\Delta\Delta$CT method. *Methods* 25, 402–408. doi: 10.1006/meth.2001.1262

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, 1–21. doi: 10.1186/s13059-014-0550-8

Murthy, M. (2001). Never say dye: the story of coloured cotton. *Resonance* 6, 29–35.

Peng, Z., Gao, Q., Luo, C., Gong, W., Tang, S., Zhang, X., et al. (2020). Flavonoid biosynthetic and starch and sucrose metabolic pathways are involved in the pigmentation of naturally brown-colored cotton fibers. *Ind. Crops Product.* 158:113045. doi: 10.1016/j.indcrop.2020.113045

Qiu, X. M. (2004). Research progress and prospects on naturally-colored cotton. *Cotton Sci.* 16, 249–254.

Sun, J., Sun, Y., and Zhu, Q. H. (2021). Breeding next-generation naturally colored cotton. *Trends Plant Sci.* 26, 539–542. doi: 10.1016/j.tplants.2021.03.007

Tan, J., Tu, L., Deng, F., Hu, H., Nie, Y., and Zhang, X. (2013). A genetic and metabolic analysis revealed that cotton fiber cell development was retarded by flavonoid naringenin. *Plant Physiol.* 162, 86–95. doi: 10.1104/pp.112.212142

Wang, L., Liu, H., Li, X., Xiao, X., Ai, X., Luo, C., et al. (2014). Genetic mapping of fiber color genes on two brown cotton cultivars in Xinjiang. *Springer Plus* 3, 1–5. doi: 10.1186/2193-1801-3-480

Wen, B., Mei, Z., Zeng, C., and Liu, S. (2017). metaX: a flexible and comprehensive software for processing metabolomics data. *BMC Bioinform.* 18:183. doi: 10.1186/s12859-017-1579-y

Wen, T., Wu, M., Shen, C., Gao, B., Zhu, D., Zhang, X., et al. (2018). Linkage and association mapping reveals the genetic basis of brown fibre (*Gossypium hirsutum*). *Plant Biotechnol. J.* 16, 1654–1666. doi: 10.1111/pbi.12902

Wickham, H. (2011). ggplot2. *Wiley Interdiscip. Rev. Comput. Stat.* 3, 180–185. doi: 10.1002/wics.147

Xiao, Y. H., Yan, Q., Ding, H., Luo, M., Hou, L., Zhang, M., et al. (2014). Transcriptome and biochemical analyses revealed a detailed proanthocyanidin biosynthesis pathway in brown cotton fiber. *PLoS One* 9:e86344. doi: 10.1371/journal.pone.0086344

Xiao, Y. H., Zhang, Z. S., Yin, M. H., Luo, M., Li, X. B., Hou, L., et al. (2007). Cotton flavonoid structural genes related to the pigmentation in brown fibers. *Biochem. Biophys. Res. Commun.* 358, 73–78. doi: 10.1016/j.bbrc.2007.04.084

Xu, W., Dubos, C., and Lepiniec, L. (2015). Transcriptional control of flavonoid biosynthesis by MYB–bHLH–WDR complexes. *Trends Plant Sci.* 20, 176–185.

Yan, Q., Wang, Y., Li, Q., Zhang, Z., Ding, H., Zhang, Y., et al. (2018). Upregulation of GhTT2-3A in cotton fibres during secondary wall thickening results in brown fibres with improved quality. *Plant Biotechnol. J.* 16, 1735–1747. doi: 10.1111/pbi.12910

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.

frontiers in Plant Science

ORIGINAL RESEARCH
published: 09 March 2022
doi: 10.3389/fpls.2022.765213

Insights Into the MYB-Related Transcription Factors Involved in Regulating Floral Aroma Synthesis in Sweet Osmanthus

Xin Yan[1,2], Wenjie Ding[1,2], Xiuyi Wu[1,2], Lianggui Wang[1,2], Xiulian Yang[1,2] and Yuanzheng Yue[1,2]**

[1] Key Laboratory of Landscape Architecture, Nanjing Forestry University, Nanjing, China, [2] Co-innovation Center for Sustainable Forestry in Southern China, Nanjing Forestry University, Nanjing, China

As an important member of the *MYB* transcription factor (TF) family, the *MYB-related* TFs play multiple roles in regulating the synthesis of secondary metabolites and developmental processes, as well as in response to numerous biotic and abiotic stressors in plants. However, little is known regarding their roles in regulating the formation of floral volatile organic compounds (VOCs). In this study, we conducted a genome-wide analysis of MYB-related proteins in sweet osmanthus; 212 *OfMYB-related* TFs were divided into three distinct subgroups based on the phylogenetic analysis. Additionally, we found that the expansion of the *OfMYB-related* genes occurred primarily through segmental duplication events, and purifying selection occurred in all duplicated gene pairs. RNA-seq data revealed that the *OfMYB-related* genes were widely expressed in different organs of sweet osmanthus, and some showed flower organ/development stage-preferential expression patterns. Here, three *OfMYB-related* genes (*OfMYB1R70/114/201*), which were expressed nuclearly in floral organs, were found to be significantly involved in regulating the synthesis of floral VOCs. Only, *OfMYB1R201* had transcriptional activity, thus implying that this gene participates in regulating the expression of VOC synthesis related genes. Remarkably, the transient expression results suggested that *OfMYB1R70*, *OfMYB1R114*, and *OfMYB1R201* are involved in the regulation of VOC synthesis; *OfMYB1R114* and *OfMYB1R70* are involved in accelerating β-ionone formation. In contrast, *OfMYB1R201* decreases the synthesis of β-ionone. Our results deepen our knowledge of the functions of *MYB-related* TFs and provide critical candidate genes for the floral aroma breeding of sweet osmanthus in the future.

Keywords: sweet osmanthus, MYB-related, transcription factor, GC-MS, floral aroma

OPEN ACCESS

Edited by:
Xiumin Fu,
South China Botanical Garden (CAS), China

Reviewed by:
Dengwei Jue,
Chongqing University of Arts and Sciences, China
Yujie Fang,
Yangzhou University, China

****Correspondence:***
Xiulian Yang
xly@njfu.edu.cn
Yuanzheng Yue
yueyuanzheng@njfu.edu.cn

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

Received: *26 August 2021*
Accepted: *03 January 2022*
Published: *09 March 2022*

Citation:
Yan X, Ding W, Wu X, Wang L, Yang X and Yue Y (2022) Insights Into the MYB-Related Transcription Factors Involved in Regulating Floral Aroma Synthesis in Sweet Osmanthus. Front. Plant Sci. 13:765213. doi: 10.3389/fpls.2022.765213

INTRODUCTION

MYB-related transcription factors (TFs), named for their highly conserved 52 amino acid MYB domain, are predominant and critical members of the MYB TF family in plants (Riechmann and Ratcliffe, 2000; Dubos et al., 2010). MYB TFs can be generally divided into four subfamilies based on the number of MYB domains present: 1R-MYB (MYB-related), R2R3-MYB, 3R(R1R2R3)-MYB,

and 4R-MYB proteins (Rosinski and Atchley, 1998). The 1R-MYB are collectively called MYB-related proteins and include proteins with a single or a partial MYB repeat (Dubos et al., 2010). Recently, many studies performing systematic identification and genome-wide characterization of the *MYB* family have reported genes in various plants, such as *Petunia hybrida*, potato (*Solanum tuberosum*), peach (*Amygdalus persica*), and Chinese jujube (*Ziziphus jujuba*) (Zhang et al., 2018; Ji et al., 2019; Liu et al., 2020; Chen et al., 2021). The MYB-related and R2R3-MYB subfamily contain the highest quantity of genes within the MYB family (Liu et al., 2020). To date, R2R3-MYB has been comprehensively characterized and reported to have diverse functions (Nesi et al., 2001; Koes et al., 2005; Ambawat et al., 2013), particularly in enhancing volatile aroma production, promoting anthocyanin accumulation, and responding to abiotic stresses. In contrast, several MYB-related proteins control diverse physiological processes, particularly the regulation of the anthocyanin biosynthesis pathway and the amelioration of multiple abiotic stresses. The first MYB-related TF identified in plants was in potato (*MYBSt1*) (Baranowskij et al., 1994), and since then, the MYB family has gradually gained more attention. However, to date, the function of *MYB-related* genes in regulating the synthesis of floral aroma remains largely unknown.

According to the most recent classification based on the highly conserved motif, MYB-related TFs are divided into three subgroups: the CPC-like subgroup of R3-type single-MYB proteins consisting of CPC (CAPRICE)-like, TRY (TRIPTYCHON)-like, ETC (Enhancer of TRY and CPC)-like and MYBL2; the CCA1-like/R-R subgroup, including the CCA1 (Circadian Clock Associated1)-like, TBP (Telomeric DNA-binding Protein)-like, RAD (RADIALIS)-like, TRF-like, LHY (Late Elongated Hypocotyl), and RVE (Early Phytochrome Responsive1)-like proteins; and the GARP subgroup, which is further classified into the LUX, MYB_CC, KANADI, ARR, HHO, GLK (GOLDEN2-LIKE), and GARP-like subfamilies. The *MYB-related* genes were subsequently confirmed to play fundamental roles as transcriptional regulators, associated repressors, circadian clock, and telomeric repeat-binding proteins in various biological processes. *AtMYBD*, a member of the CCA1-like/R-R subgroup in *Arabidopsis thaliana*, regulates anthocyanin biosynthesis in a circadian clock-related manner (Nguyen and Lee, 2016). *SlTRY*, an R3 MYB member of the CPC-like subgroup, has been confirmed to be involved in anthocyanin biosynthesis and is considered a core regulator of root-hair development and the plant trichome (Rumi et al., 2013). The grape TF AQUILO, a member of the GARP subgroup, increases cold tolerance by promoting the accumulation of raffinose family oligosaccharides (Sun et al., 2018). In recent years, *MYB-related* TFs have been found to be involved in regulating anthocyanin biosynthesis in various pathways and products. The proteins *MdLUX* and *MdPCL*-like, containing a single MYB-like repeat, promote anthocyanin accumulation through DNA hypomethylation in apple (*Malus domestica*) (Li et al., 2021). The functions of MYB-related proteins in metabolite biosynthesis have been characterized in various land plants, for instance, *Arabidopsis*, soybean (*Glycine max*), grapevine (*Vitis vinifera*), lily (*Lilium* spp.), and tiger lily (*Lilium lancifolium*) (Matsui et al., 2008; Niu et al., 2016; Sakai et al., 2019; Yong et al., 2019).

Sweet Osmanthus (*Osmanthus fragrans)* is an evergreen woody flowering member of the Oleaceae. Because of its pleasing floral fragrance and multitudinous flowers, sweet osmanthus is widely cultivated in China as a common landscape tree species (Huang et al., 2015). Sweet osmanthus is relevant to several industries, including tea, food, and perfume production (Li et al., 2020a). The floral fragrance of sweet osmanthus has been comprehensively studied in terms of compound identification and classification. Studies have identified that β-ionone is an essential substance that strongly affects the floral fragrance of sweet osmanthus, and *CCD4* is a vital enzyme regulating the formation of β-ionone (Han et al., 2016). Previous studies have confirmed that *OfCCD4* increases the production of β-ionone, based on the existence of β-carotene, and both *OfWRKY3* and *OfERF61* are positive regulators of the *OfCCD4* gene and enable the synthesis of β-ionone in petals of sweet osmanthus (Han et al., 2016, 2019). According to data on the sweet osmanthus genome, we found that the *OfCCD4* promoter region contains numerous *cis*-acting regulatory elements, thus indicating that *OfCCD4* might be regulated by other TFs (Yang et al., 2018). However, functional identification of related TFs and examination of the transcriptional regulation mechanism of floral volatile organic compounds (VOCs) have been limited. Therefore, studies on the MYB-related TFs' regulation of enhancing floral VOC production are needed to improve understanding of the mechanism of VOC production and the genetic quality of sweet osmanthus.

Here, we conducted a comprehensive analysis of sweet osmanthus' MYB-related TF family members, including their potential roles in regulating the formation of floral VOCs. In this study, 212 MYB-related members were identified from the genome of sweet osmanthus, and their chromosomal locations, conserved motifs, and expression patterns were analyzed and visualized. Additionally, we selected three candidate *MYB-related* genes that had flower organ/development stage-preferential expression patterns and then performed detailed gas chromatography-mass spectrometry (GC-MS) analysis of floral VOCs between transient expression plants and empty vector plants.

MATERIALS AND METHODS

Retrieval of Putative OfMYB-Related Family Genes

Sweet osmanthus genomic sequences were acquired from the genome database for *O. fragrans*. The sequence of the MYB DNA binding domain (DBD) was downloaded from the Pfam database (http://pfam.xfam.org/, accession PF00249) (Finn et al., 2016), and HMMER software (version 3.0) with a default *E*-value. The NCBI Batch Web CD-Search Tool (Marchler-Bauer and Bryant, 2004) was further used to assess the numbers of DBDs in the sequences and retain the genes with one DBD. Finally, the putative *MYB-related* genes, according to the number of DBDs and the position information extracted from the General Feature

Format (GFF) files, were identified and renamed according to their locations on the chromosomes. The online tool ExPASy was used to verify the physicochemical characteristics of these MYB-related protein sequences, such as the theoretical isoelectric point (pI) and molecular weight.

Phylogenetic Analysis and Classification of OfMYB-Related Proteins

A total of 132 AtMYB-related protein sequences[1] and 212 putative OfMYB-related proteins were selected to construct a phylogenetic tree through neighbor-joining method with a bootstrap score of 1,000 and p-distance parameter (MEGA7.0). The complete amino acid sequences of OfMYB-related proteins were used to obtain the DBDs and validated in the Pfam database. The alignment and visualization of 212 *OfMYB-related* domains and subsequent protein analysis were accomplished with ClustalX.

Gene Structures and Conserved Motifs of the OfMYB-Related Gene Family

The *OfMYB1R* gene structures and coding sequences (CDSs) were derived from the sweet osmanthus genomic GFF3 file. TBtools software was used to visualize the gene structures, CDS boundaries, and intron distribution (Chen et al., 2020). The MEME Suite webserver was used to screen the conserved motifs in *OfMYB-related* TFs, identifying a maximum of 20 motifs[2] (Bailey et al., 2009).

Gene Duplication and Synteny Analysis of OfMYB-Related Genes

The chromosomes corresponding to MYB-related genes were screened and mapped with MG2C.[3] The Multiple Collinearity Scan toolkit (MCscanX) was used to analyze the tandem and segmental duplication pattern of each *MYB-related* gene.[4] Ks (synonymous) and Ka (non-synonymous) substitution ratios of gene pairs were evaluated with TBtools software (Chen et al., 2020).

Quantitative Real-Time PCR Analysis

Total RNA of sweet osmanthus was extracted from the bud-pedicel (S1), bud-eye (S2), primary blooms (S3), full bloom (S4), and post-bloom flower (S5) stages. Three biological samples were collected and processed with the RNAprep Pure Plant Kit (Tiangen, Beijing, China). After removing the gDNA from the total RNA, cDNA was synthesized with SuperMix (Transgen, Beijing, China) under the following conditions: 42°C for 15 min and 85°C for 5 s (Yue et al., 2021). The qRT-PCR primers for *OfMYB-related* genes were designed in Primer Premier 5.0. The expression of *OfACTIN*, *OfRAN*, and *OfRBP2* genes served as average references for normalization of the expression levels in different flower stages (Zhang et al., 2015) (**Supplementary Table 1**). Each qRT-PCR assay was performed on three biological samples and three technical replicates. Data of comparative threshold cycle (Ct) values were used to calculate the relative gene expression levels (Yue et al., 2021).

Subcellular Localization and Transcriptional Activation Assays

To construct 35S:*OfMYBMYB1R70/114/201*:GFP, the Super 1,300 vector containing the CDS of *OfMYB-relared70/114/201*, without the stop codon was used. *Sma*I and *Kpn*I restriction sites were utilized. *Agrobacterium tumefaciens* (GV3101) containing 35S:*OfMYB1R70/114/201*:GFP and negative control vectors were infiltrated into tobacco (*Nicotiana benthamiana*) leaves for the analysis of subcellular localization. The tobacco plants were grown for 35 days in a growth chamber with 15/9 h light/dark (144 $\mu mol \cdot m^{-2} \cdot s^{-1}$) and 26 ± 2°C. After 3 days of growth, green fluorescent protein (GFP) fluorescence signals were confirmed in the infiltrated plants *via* DAPI staining with a LSM710 microscope (Zeiss, Germany).

Additionally, through the *Nde*I and *Eco*RI restriction sites, the CDS of *OfMYB-relared70/114/201*, lacking the termination codon, were transformed into the pGBKT7 vector. Then the *Saccharomyces cerevisiae* yeast strain AH109 (WeidiBio, Shanghai, China) was used to obtain the four vectors (three types of pGBKT7-*OfMYB1Rs* and one negative control). Finally, transformed yeast with overexpression vectors were grown in selective defined SD/-Trp, SD/-Trp-Ade, and SD/-Trp-Ade + X-α-gal media with incubation at 30°C for 3 days in the dark.

Aroma Compound Analysis of Plants With Transient Expression

To explore the regulatory function of *OfMYB1R70/114/201* genes in volatile metabolic components, we selected tobacco leaves to analyze floral VOCs by transient expression. The method to determine the transient expression of candidate genes in *Nicotiana benthamiana* leaves has been successfully used to explore the functions of genes involved in VOC synthesis, such as *OfTPS1/2* (Zeng et al., 2016), *AtTPS10/14* (Ginglinger et al., 2013), and *MpGPS.SSU* (Yin et al., 2017). GV3101 containing 35S:*OfMYB1R70/114/201*:GFP and empty vectors were used to infiltrate 35-day-old tobacco leaves. Six biological replicates of each OfMYB1R sample were injected after 48 h, and the empty vectors were injected in four biological replicates. After growth for 2 days (light/dark: 15/9 h, 144 $\mu mol \cdot m^{-2} \cdot s^{-1}$), headspace solid-phase microextraction (SPME), a precise method for measuring aroma, was used to collect the floral VOCs. In every 4-ml SPME vial, 0.4 g samples of fresh leaves were added to the bottom, and 1 μl of 100000-fold diluted ethyl caprate was added to the middle (Shanghai, Macklin Inc., China) (Yang et al., 2018). After exposure at the middle of the capped vials for 35 min at 42°C, the SPME fibers of DB-5 MS were injected into the heated syringe port for 3 min, and desorption was performed at 250°C (Bellefonte, Supelco Inc., United States). Subsequently, the desorbed floral VOCs were identified with the Trace DSQ GC-MS device (Ji, 2021), and the N-alkanes were used to calculate the RI value of the volatile compounds for comparison with the RIL value within DB-5MS (30 × 0.25 × 0.25) to confirm the aroma

[1] http://www.arabidopsis.org/

[2] http://meme-suite.org/tools/meme

[3] http://mg2c.iask.in/mg2c_v2.0/

[4] https://sourceforge.net/projects/mcscanx/

compound detection[5] (Ji, 2020, 2021). In the software SIMCA 13.0 (Umetrics AB, Umea, Sweden), divergence in the metabolite levels of plants with transient expression was detected with partial least squares discriminant analysis (PCA-X).

Total RNA Extraction and Semi-Quantitative RT-PCR of Plants With Transient Expression

To validate the empty vector and 35S:*OfMYB1R70/114/201*:GFP expression in transgenic tobacco plants, we snap-froze the leaves of 35-day-old tobacco with transient expression 48 h after injection of the vector (light/dark: 15/9 h, 144 $\mu mol \cdot m^{-2} \cdot s^{-1}$) in liquid nitrogen. According to the RNAprep Pure Plant Kit (Tiangen, Beijing, China) cDNA transcribed from total RNA isolated from the leaves of transgenic plants and empty vector plants was used as the template (Ding et al., 2019). Each *OfMYB1R* assay was performed on six biological samples, and empty vector plants were performed on four biological replicates. The primers of *OfMYB1Rs* for semi-quantitative RT-PCR (RT-PCR) were the same primers used in the qRT-PCR analysis. The gene *NbL25* was used as a reference to normalize the expression levels in tobacco (**Supplementary Table 1**). The primers for *NbCCD4s* and *NbLCYB*, which are homologs of *OfCCD4* and *OfLCYB*, were selected to verify the participation of candidate genes in regulating floral aroma synthesis (**Supplementary Table 1**).

RESULTS

Identification and Classification of MYB-Related Genes in Sweet Osmanthus

Utilizing the MYB DBD HMM profile (Pfam ID PF00249), MYB and MYB-like domains for the entire sweet osmanthus genome were screened. According to the results, verified by Pfam and NCBI-CDD analysis, a total of 212 non-redundant OfMYB1R protein sequences were identified, a number greater than that for most species (**Table 1**). Based on the chromosomal position of the MYB-related genes, these genes were successively renamed *OfMYB1R1* to *OfMYB1R212*. Basic information on *OfMYB-related* genes, such as gene ID, renamed name, chromosomal location, pI, and molecular weight, can be found in **Supplementary Table 2**. The protein sequences are provided in **Supplementary Table 3**. The total length of the OfMYB-related protein sequences ranged from 82 to 1,816 amino acids, the molecular weights ranged from 9.21 to 199.69 kDa, and the predicted pI values were 4.49–10.1.

Additionally, 132 AtMYB-related proteins were analogously identified from the *A. thaliana* genome (Dubos et al., 2010; Liu et al., 2020). The phylogenetic and evolutionary relationships between the 344 AtMYB- and OfMYB-related proteins were analyzed (**Figure 1**). In accordance with the classification of the MYB-related proteins of Arabidopsis, 212 OfMYB-related proteins were divided into three subgroups: CPC-like, GARP-like (MYB_CC, ARR, HHO, LUX, KANADI, and GRAP-related, considered an unknown subfamily), and a heterogeneous subgroup (consisting of CCA1-like/R-R, TRF-like, TBP-like, RAD-like, and atypical MYB-related proteins).

Structure and Motif Analysis of *OfMYB-Related* Genes

The three subgroups, comprising 212 OfMYB-related proteins, were further classified into 12 categories according to highly conserved motifs. The GARP-like subgroup was divided into six clades, whereas the CCA1-like/R-R subgroup consisted of five clades (**Figure 2**). Moreover, 22 OfMYB-related proteins in the TBP-like subgroup were split into three major clades whose sequence alignments implied high similarity within each clade, thus indicating that these genes may have similar functions.

TABLE 1 | Comparison of the *MYB-related* gene family size between *Osmanthus fragrans* and other species.

Type	Species	The number of *MYB-related* genes
Monocots	*Ananas comosus*	65
	Aegilops tauschii	97
	Zoysia japonica	99
	Oryza sativa subsp. japonica	106
	Eragrostis tef	143
	Zea mays	169
Dicotyledon	*Arachis hypogaea*	44
	Carica papaya	51
	Citrullus lanatus	54
	Coffea canephora	56
	Vitis vinifera	57
	Fragaria vesca	65
	Dianthus caryophyllus	66
	Cucumis melo	68
	Cannabis sativa	70
	Capsicum annuum	76
	Castanea mollissima	78
	Prunus mume	89
	Populus euphratica	91
	Cajanus cajan	94
	Cucumis sativus	94
	Ziziphus jujuba	96
	Cicer arietinum	99
	Citrus sinensis	117
	Actinidia chinensis	138
	Catharanthus roseus	144
	Prunus persica	145
	Malus domestica	146
	Nicotiana benthamiana	160
	Brassica oleracea	202
	Osmanthus fragrans	212
	Brassica napus	251
	Glycine max	342

[5]https://webbook.nist.gov/

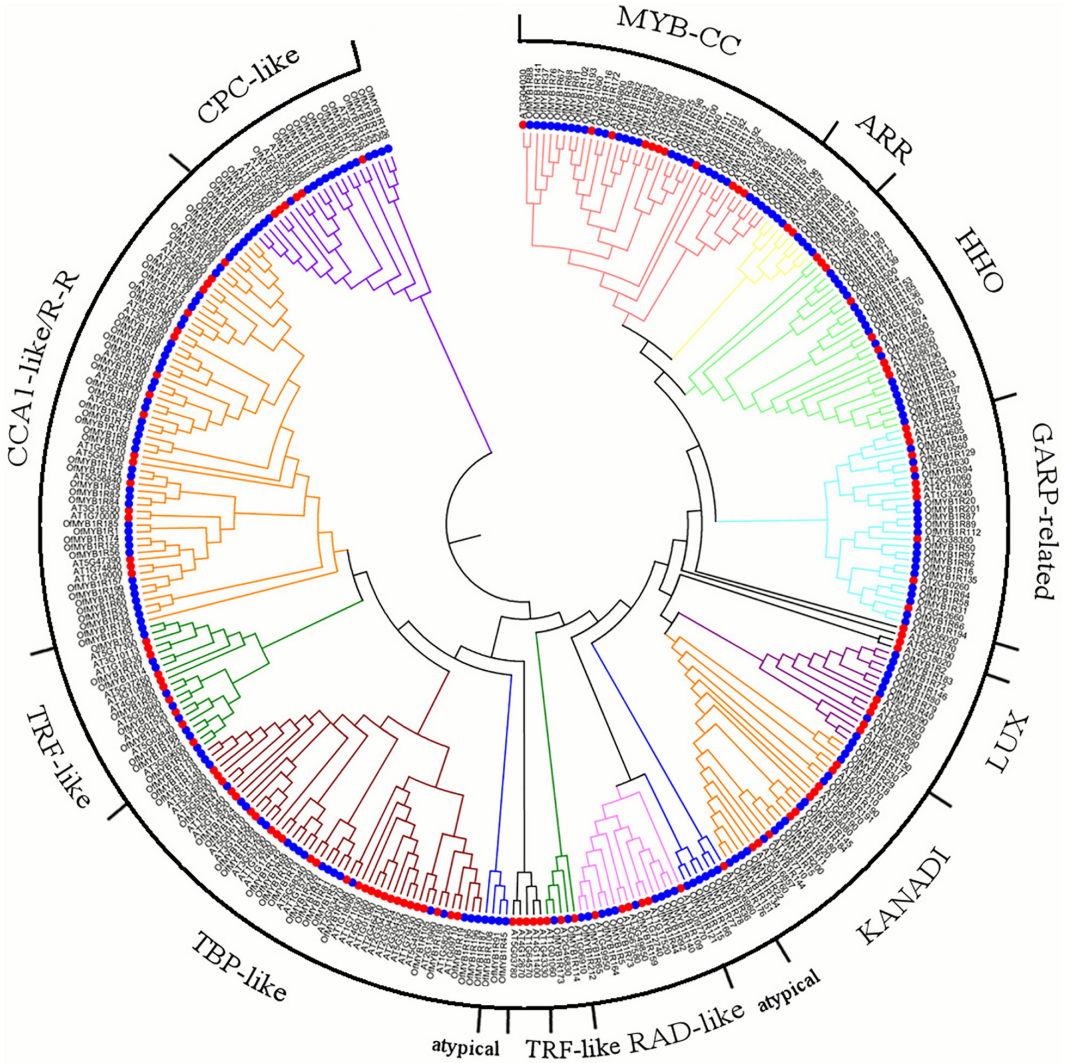


FIGURE 1 | Phylogenetic relationship between the 212 *O. fragrans* (blue circles) and 132 *Arabidopsis* (red circles) *MYB-related* proteins, illustrated by a maximum likelihood (ML) tree. Alignment was performed with the ClustalW program, and these proteins were imported to MEGA 7.0 to construct the phylogenetic tree with 1,000 bootstrap iterations. MYB1Rs represents OfMYB-related proteins.

The largest subgroup was GARP-like, which contained 94 OfMYB-related proteins, comprising MYB_CC, LUX, HHO, KANADI, ARR, and GARP-related, with diversely conserved motifs and CDS patterns (**Figure 2**); however, the GLK subfamilies did not have BLAST results. The GARP-like subgroup was categorized by the conserved motif Wx3LHx2Fx6L(G)Gx5PK/Sx5Mx4L and the DNA-binding helix 3 domain, harboring the SH(A/L)QK(F/Y) sequence, which is commonly found in diverse plant genomes (Antje et al., 2011). Apart from the common conserved motifs, several unique motifs were detected in different subfamilies, such as the LHEQLE motif in MYB_CC, which contained 28 members and was the largest clade. The smallest clade belonged to ARR, with seven proteins. The second-largest subgroup, with three W residues, consisted of the CCA1-like/R-R, TBP-like, TRF-like, RAD-like, and an atypical MYB-related clade. Notably, the TBP-like subgroup was classified into three distinct clades based on remarkably conserved motifs, and the most common conserved motif of TBP-like was LKDKW(R/K) (N/T) (**Supplementary Figure 1**). The smallest subgroup, containing only 44 OfMYB-related proteins, was CCA1-like/R-R, an R3-type MYB TF, which had a highly conserved SHAQK(Y/F) (F/Y) motif in the MYB domain third helix—the distinctive characteristic in this subgroup (**Supplementary Figure 1**).

Chromosomal Distribution and Gene Synteny Analysis of OfMYB-Related Gene Families

The 212 *OfMYB-related* genes were located on all of the 23 uneven sweet osmanthus chromosomes (**Figure 3**). Remarkably,

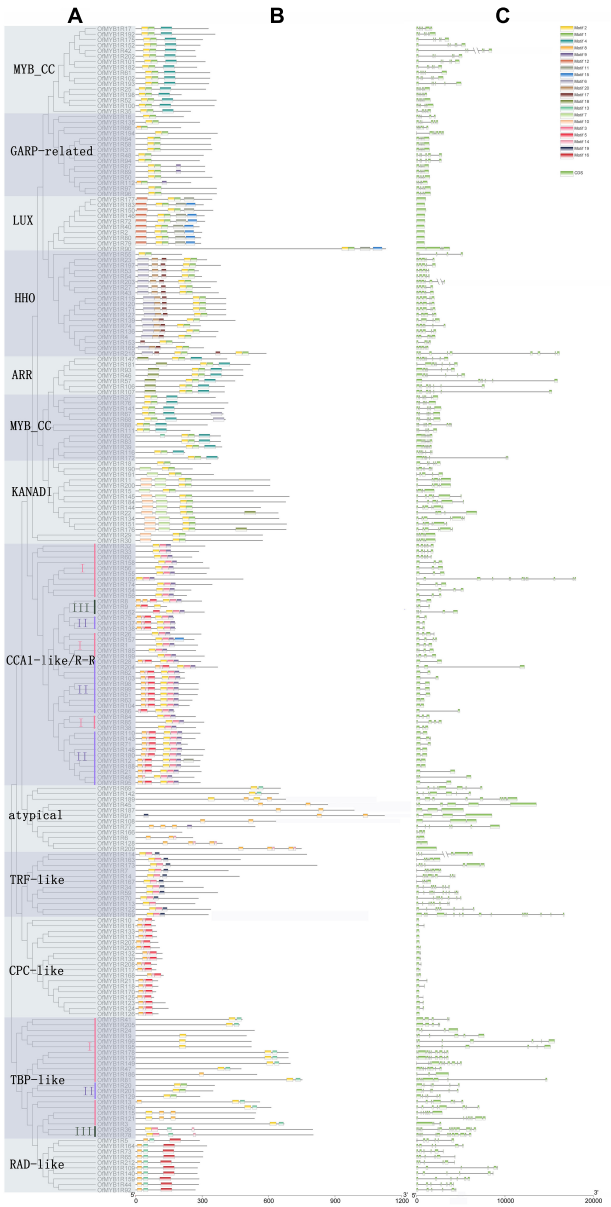


FIGURE 2 | Phylogenetic relationships of 212 OfMYB-related proteins, visualized by a ML tree **(A)**. Conserved motifs of OfMYB-related proteins **(B)**. Exon/intron structures of *OfMYB-related* genes. Exon(s) and intron(s) are shown with green boxes and gray lines, respectively **(C)**. The 20 conserved motifs in the OfMYB-related TFs were predicted by the MEME Suite web server. I–III indicate clades I–III.

chromosome (Chr) 03 contained the largest number of *OfMYB-related* genes (17), followed by Chr08 and Chr11 (each with 15 genes), and then Chr01 (14 genes). In contrast, only two genes were mapped to Chr23 and Chr06. The pattern of *OfMYB1Rs* followed a random distribution (sig < 0.05) (**Supplementary Figure 2**).

Gene duplication plays a crucial role in driving genetic diversity and is common in plant genomes. This research identified seven tandemly duplicated pairs of 12 *OfMYB-related* genes and 68 segmentally duplicated out of the 212 *OfMYB-related* genes. Tandemly duplicated gene pairs were found only in chromosomes 01, 04, 11, 18, and 22. Chr11 had three pairs of tandemly duplicated genes continuously, and segmentally duplicated genes were found on all chromosomes (**Figure 3**). The tandemly duplicated cluster on Chr11 provides interesting insight into the expansion pattern of sweet osmanthus expansion genes in this large gene family. Moreover, 68 segmental duplication events were distributed across all chromosomes in the synteny blocks (**Supplementary Figure 3**). Every tandemly duplicated *OfMYB-related* gene pair had a Ka/Ks value less than 1, which evolved in purifying selection; in particular, duplicated gene pairs of *OfMYB1R124* (Chr11) VS. *OfMYB1R125*(Chr11) and *OfMYB1R125* (Chr11) VS. *OfMYB1R126*(Chr11) were significantly different from 1 (**Supplementary Table 4**). Based on these results, we concluded that some *OfMYB-related* genes originated from gene duplication events, whereas the evolution of the *OfMYB-related* genes was influenced by segmental duplication.

Expression of OfMYB-Related Genes in Different Organs and Flower Developmental Stages

The expression patterns of *OfMYB-related* genes in different organs and flower developmental stages were analyzed with RNA-sequencing (**Supplementary Table 5**). Four sweet osmanthus organs were sequenced and used to acquire the reads per kilobase per million mapped reads (RPKM) values of 194 *OfMYB-related* genes, whereas the values of the remaining 17 genes were not found. Eight *OfMYB-related* genes were observed in a single stage: three genes were observed only in the root or primary-bloom stage, two genes (*OfMYB1R74* and *OfMYB1R156*) were observed only in the stem stage. The remaining 186 *OfMYB-related* genes showed different expression patterns, and 142 were present in all samples. Based on the of the conserved motif analysis results, 194 expressed *OfMYB-related* genes were classified into three clear subgroups. Most members of the CCA1-like/R-R subgroup had high expression in leaves as well as the primary- and full-bloom stage of flowers, whereas a wide range of CPC subgroup genes were expressed mainly in the root, stem, and flower fading (S5) stages. The members of the GARP subgroup showed particularly high expression in the root (**Figure 4**). Remarkably, in the CCA1-like/R-R subgroup, three *OfMYB-related* genes (*OfMYB1R70*, *OfMYB1R114*, and *OfMYB1R201*) had the highest expression in the flower organ/development stage, thus indicating a preferential expression pattern.

Quantitative Real-Time PCR Analysis

According to the expression patterns and RPKM values of *OfMYB-related* genes (**Figure 4**), three *OfMYB-related* genes (*OfMYB1R70/114/201*) of the CCA1-like/R-R subgroup were further selected for validation experiments in five flowering stages. The results were reflected those of the RNA-seq experiment, thus verifying the accuracy of the RNA-seq data. The expression of *OfMYB1R70* was similar to that of *OfMYB1R114*, which was generally up-regulated during flowering (S1 and

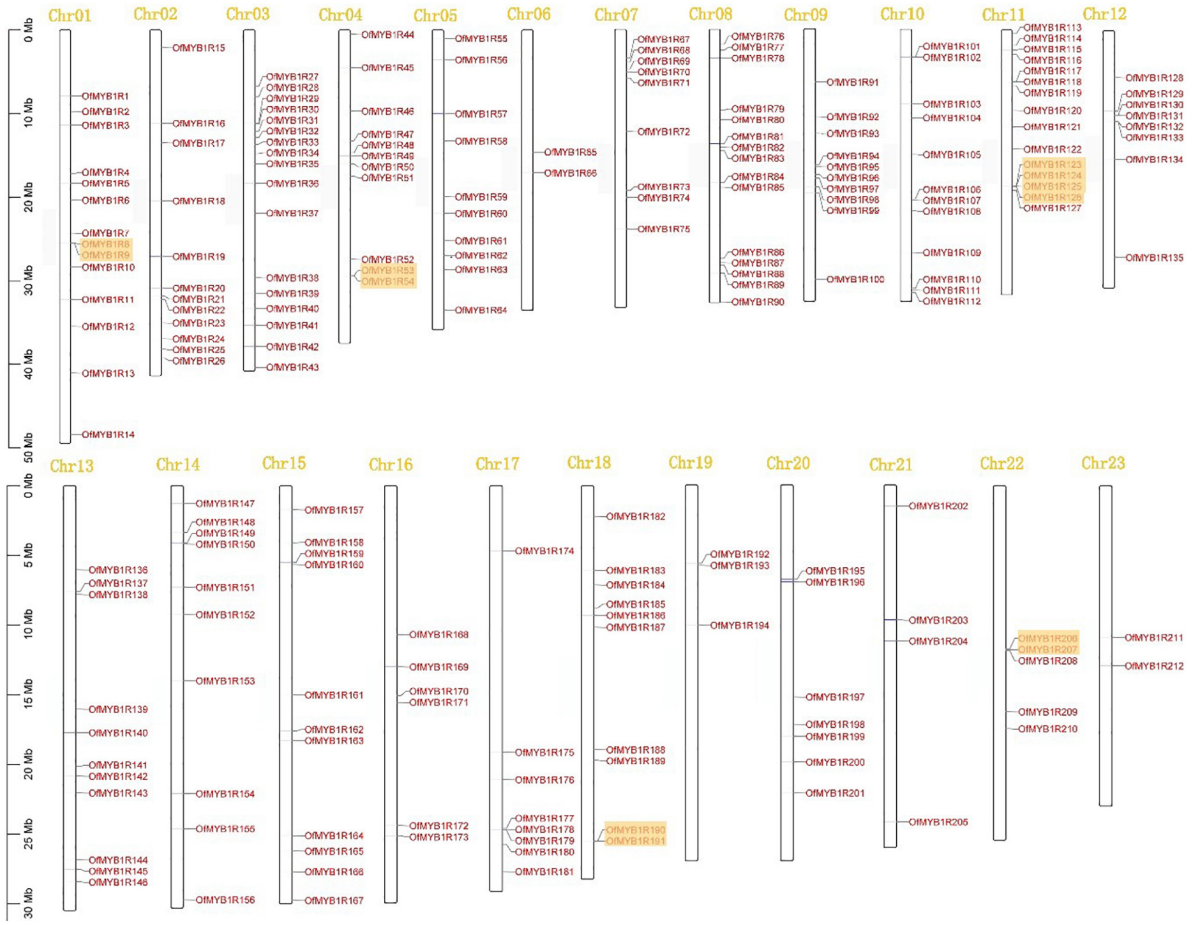


FIGURE 3 | Chromosomal locations of 212 *OfMYB-related* genes, mapped on the basis of the GFF files of *O. fragrans*. The numbers at the top represent each chromosome. Orange boxes represent clusters of tandemly duplicated genes.

S3). Meanwhile, S3 had the lowest expression, and other processes were down-regulated. Interestingly, the expression of *OfMYB1R70* in S3–S5 was similar to that of *OfMYB1R114*, but the expression of *OfMYB1R70/114* was considerably elevated in stages S3–S4, whereas the expression of *OfMYB1R*201 increased slightly in the same stages (**Figure 5**). For reliable normalization, we selected three reference genes, *ACTIN*, *RAN*, and *RBP2*, for the qRT-PCR analysis. The expression levels of *OfMYB1Rs* were obtained from the average of three reference genes (**Figure 5**), and the independent expression results of *ACTIN*, *RAN*, and *RBP2* are shown in **Supplementary Figure 4**.

Subcellular Location and Transcriptional Activation Activity of Candidate OfMYB-Related Genes

To thoroughly examine the potential function of *OfMYB-related* genes in transcriptional regulation, we transformed tobacco leaves to assess the subcellular localization. The GFP fluorescence of the p35S:*OfMYB1R70*:GFP fusion protein was intense in the nucleus, as confirmed by DAPI staining (**Supplementary Figure 5**). The fluorescence of p35S:*OfMYB1R114*:GFP, like that of the control GFP protein lacking *OfMYB-related* genes, was dispersed in the cytoplasm and nucleus (**Supplementary Figure 5**). In contrast, the fluorescence of the p35S:*OfMYB1R70/201*:GFP fusion protein was detected in the nucleus and was dispersed in the plasma membrane. To further identify the transcriptional activation functions of the *OfMYB-related* genes, we cloned the CDSs of selected *OfMYB-related* genes into the yeast expression vector pGBKT7. The negative control pGBKT7 and *OfMYB-related* genes showed comparable growth on SD/-Trp medium with serial dilutions (**Figure 6**). Nevertheless, except for *pGBKT7*–OfMYB1R201, no yeast cells containing *pDEST-GBKT7-OfMYB1R* plasmids survived and turned blue on SD/-Trp-Ade medium with X-α-Gal.

Verification and Determination of Volatile Aroma Compounds in Plants With Transient Expression

To investigate the volatile metabolic components, we obtained the GC-MS total ion current chromatograms of 22 tobacco samples from four types of plants with transient expression. Approximately 23 volatile compounds were identified in the

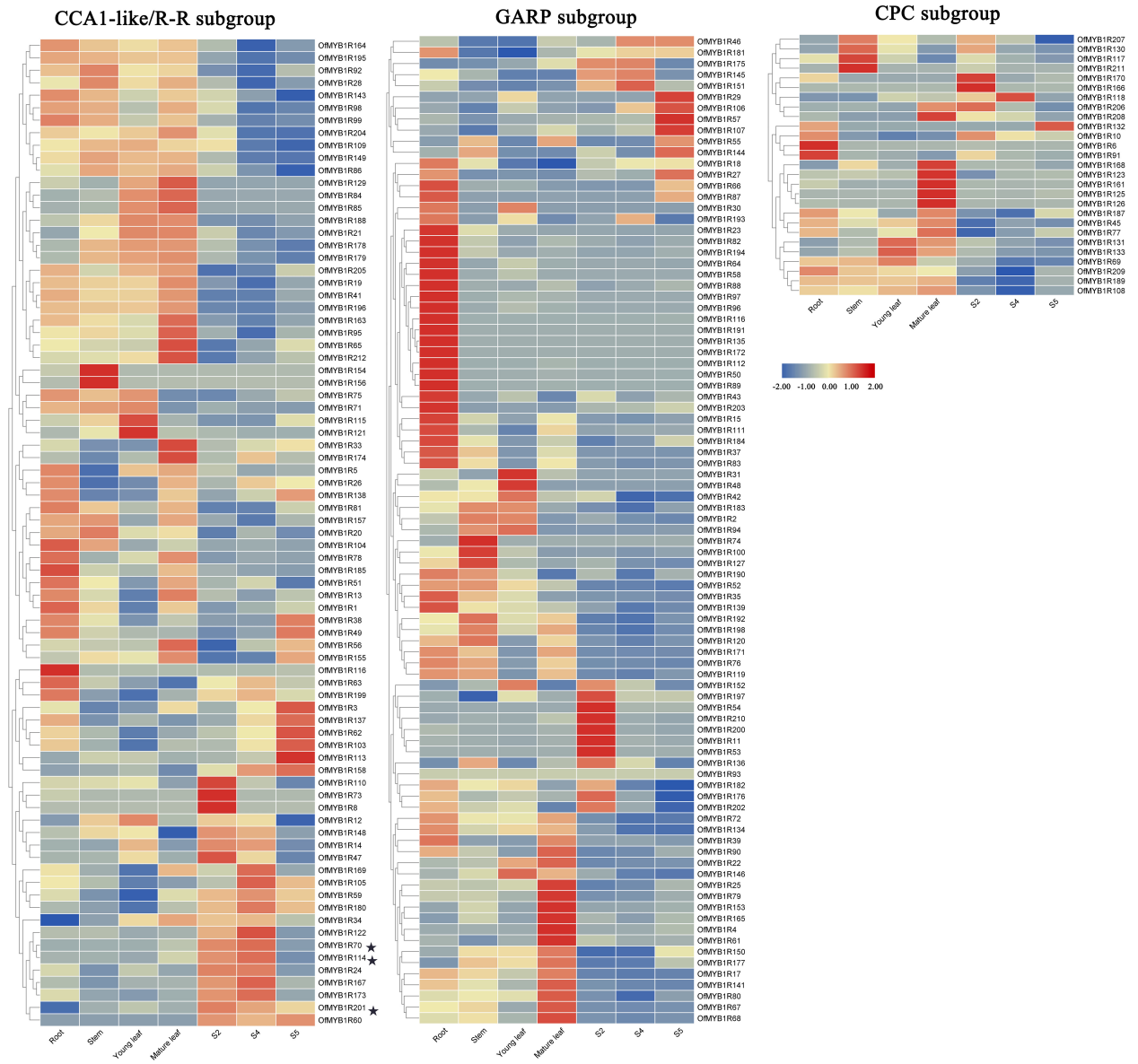


FIGURE 4 | Expression analysis of 194 *OfMYB-related* genes in seven organs. The RPKM values were $\log_2$ transformed. S2 represents the bud-eye stage, S4 the represents full blooming stage, and S5 represents the flower fading stage. Genes with no transcriptome data are not shown. The candidate genes in the biosynthesis of floral aroma are marked by asterisks.

plants with transient expression, including esters, phenolics, ketones, alkanes, and monoterpenes (**Supplementary Table 6**). PCA-X plots, generated by comparing three types of plants with transient expression and empty vector (NC), showed significant metabolic differences (**Figure 7**). Notably, according to two principal components, tobacco leaves infiltrated with 35S:*OfMYB1R70/114/201*:GFP differed from those with control vectors in the production of volatile aromatic compounds. *OfMYB1R70/114* and NC clustered separately in the PCA-X. *OfMYB1R201* was slightly scattered in the PCA (**Figure 7**), thus implying a role in regulating the production of volatile aromatic compounds. Among the three types of plants with transient expression, the plants with transient *OfMYB1R114* expression had a greater ability to regulate aromatics synthesis than the other two types. The comprehensive regulation pattern of *OfMYB1R114* was greater than *OfMYB1R70*, in view of the number of differentially present substances satisfying VIP $>$ 1 and $P < 0.05$.

Interestingly, further analysis revealed that *OfMYB1R114* regulates the synthesis of six esters, two ketones, one monoterpene, one alkane, and one aldehyde, which showed significant discrepancies in the GC-MS analysis of the aroma content with VIP $>$ 1 and SPSS with $P < 0.05$. The substances that slightly changed included floral VOCs such as β-ionone,

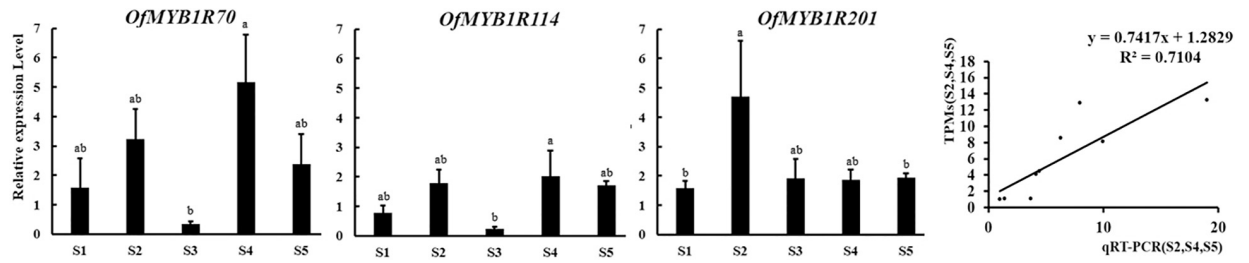


FIGURE 5 | Expression patterns of three candidate *OfMYB-related* genes using the average of three reference genes (*ACTIN*, *RAN*, and *RBP2*) in different stages of sweet osmanthus flowers. Three biological and technological replicate measurements were performed for each stage, and the data are shown as the mean ± SD ($n = 3$) at the top of each bar. The small letters above the error bars indicate significant differences obtained in SPSS 25 with Tukey's test at $P < 0.05$. S1–S5 represent the bud-pedicel stage, bud-eye stage, primary blooming, full blooming, and flower fading stage, respectively, on the abscissa.

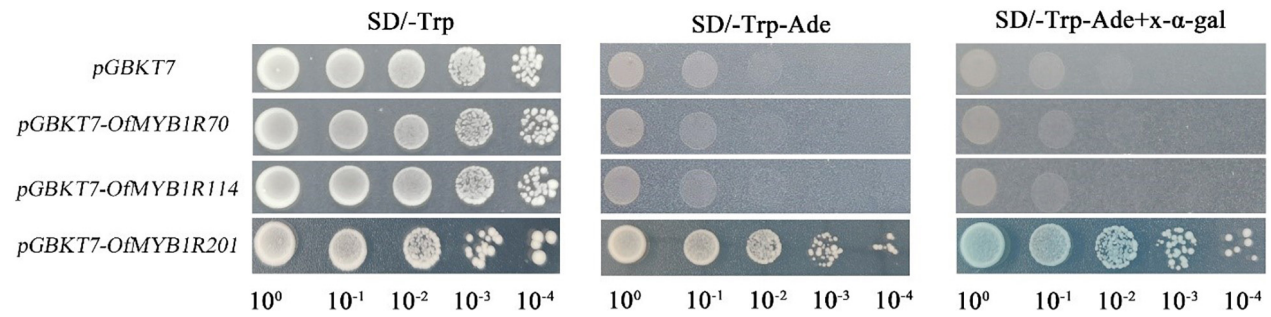


FIGURE 6 | Transcriptional activation analysis of candidate *OfMYB1Rs*. Growth of transformants containing *pGBKT7* and *pGBKT7-OfMYB1R70/114/201* on SD/-Trp medium with serial dilutions. Only *pGBKT7-OfMYB1R201* survived and turned blue on SD/-Trp-Ade medium in the presence of X-A-gal.

nonanal, and benzeneacetaldehyde (**Supplementary Table 6**). Notably, *OfMYB1R114* mainly altered the production of nonanal, which is common in rose and sweet orange, and caused a clear difference in β-ionone. Interestingly, in agreement with the results of correlation analysis, *OfMYB1R114* positively regulated the production of β-cyclocitral. *OfMYB1R70* is predominantly involved in the synthesizing of four esters and two ketones, and the most striking change in floral fragrance compounds was observed for β-ionone, a significant food flavor. In contrast, O*fMYB1R201* showed significantly negative regulation of four esters, one ketone, one aldehyde, and one naphthenic compound, and the production of floral VOCs (nonanal and β-ionone) was comparatively low. Simultaneously, the results of RT-PCR demonstrated that *OfMYB1R70/114/201* were transiently expressed in the plants. *OfMYB1R70/114/201* was detectable in 35S:OfMYB1R:GFP plants with transient expression, whereas *OfMYB1R70/114/201* expression was undetectable in transformed empty vector plants. According to the qRT-PCR results, *NbCCD4.1* was significantly up-regulated in plants with transient expression of *OfMYB1R70* and *OfMYB1R114*, and *NbCCD4.2* was markedly elevated in the plants with transient expression of *OfMYB1R114*. Interestingly, *NbCCD4.3* was significantly elevated in the plants with transient expression of *OfMYB1R70* and dramatically diminished in the plants with transient expression of *OfMYB1R201* (**Figure 7**). These results indicated that the transient overexpression of *OfMYB1R70* activated the expression of *NbCCD4.1* and *NbCCD4.2*; *OfMYB1R114* might have activated *NbCCD4.1* and *NbCCD4.3*; and *OfMYB1R201* repressed the transcription of *NbCCD4.3* in tobacco plants, thus leading to the changes in β-ionone content in the plants with transient expression of the three *OfMYB1Rs*.

DISCUSSION

To date, MYB TFs, mainly comprising MYB-related and R2R3 MYB, modulate the transcription levels of numerous essential genes in secondary metabolism biosynthetic pathways, thereby regulating the production of diverse floral VOCs (Jian et al., 2019). Several MYB members have been identified as a conserved module regulating terpene synthesis. *MsMYB*, a novel negative regulator in *Mentha spicata* associated with the promoter region of geranyl diphosphate synthase (GPPS), decreases the production of monoterpenes by inhibiting the activity of GPPS (Reddy et al., 2017). Recently, the function of R2R3MYBs in floral aroma formation in *Hedychium coronarium* has been extensively reported (Abbas et al., 2021). However, further studies of MYB-related proteins, particularly the close relationship between the function of *MYB-related* TFs and floral aroma synthesis, remain unclear.

The number of *OfMYB-related* genes was more significant than that observed in other plants reported over the recent years (**Table 1**; Jin et al., 2017), such as 109 genes in peach

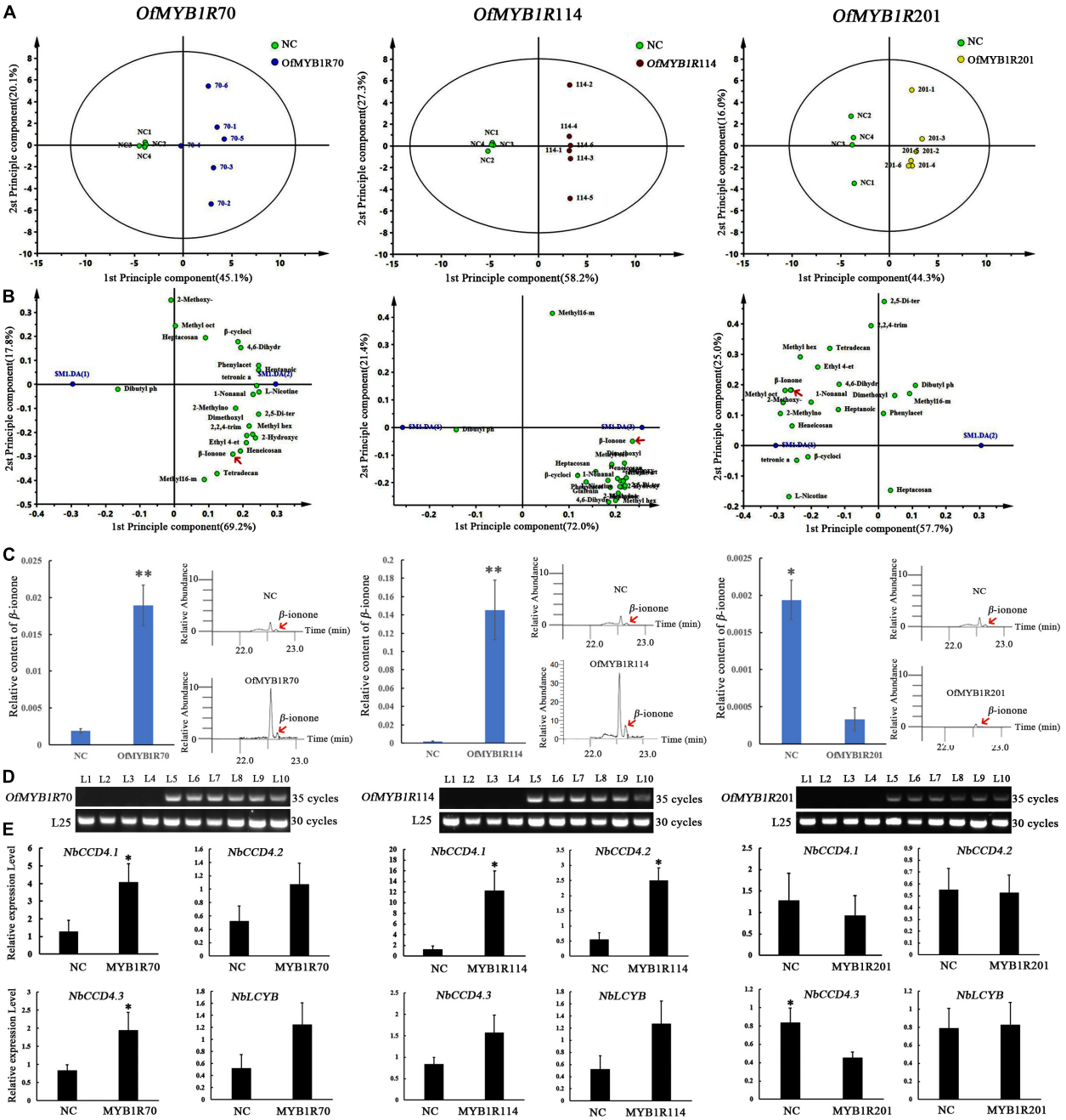


FIGURE 7 | Score scatter plot of the leaf GC-MS profiles of three types plants with transient expression of different *MYB1R* transcription factors in in tobacco, determined with PCA-X analysis in SIMCA13.0 **(A)**. Loading scatter plot analysis of floral VOCs in plants after infiltration of 35S:*OfMYBMYB1R70/114/201*:GFP and control vectors. The red arrow marks the floral aroma in plants with transient expression **(B)**. Floral composition analysis of β-ionone after infiltration with 35S:*OfMYBMYB1R70/114/201*:GFP and control vectors **(C)**. RT-PCR to explore the regulation of gene transcription *OfMYBMYB1R70/114/201* levels. A fragment of the *NbL25* gene was amplified as an internal control. L1–4 represent different plants infiltrated with empty vectors, and L5–10 represent different plants infiltrated with 35S:*OfMYB1R*:GFP vectors **(D)**. qRT-PCR analysis of the NbCCD4 transcript levels in control and transiently transformed tobacco leaves **(E)**. $^{*}P < 0.05$, $^{**}P < 0.01$.

(Zhang et al., 2018). In our study, we identified 212 *MYB-related* genes in *O. fragrans*, a value close to the number of *R2R3-MYB* genes in *O. fragrans* (243 genes) (Li et al., 2020a), thus representing the largest subfamily of *MYB* genes, which play critical roles in various processes of growth, development, and responses to abiotic or biotic stress. While high, the number of *R2R3-MYB* subfamily members did not exceed the number of *1R-MYB* subfamily members. The amount of *1R-MYBs* in *Ganoderma* was twice the number of *R2R3-MYBs*. Moreover, previous studies have performed genome-wide analyses of

MYB-related proteins from 16 species of flowering plants, moss, Selaginella, and algae (Du et al., 2013). In contrast to the detailed studies on the *R2R3-MYB* subfamily, little attention has been paid to the *MYB-related* (*1R-MYB*) subfamily, which has been reported only in a few common crops in recent years: potato, sunflower (*Helianthus annuus*), radish (*Raphanus sativus*), and oil palm (*Elaeis guineensis*) (Li et al., 2020b; Liu et al., 2020; Zhou et al., 2020; Muleke et al., 2021).

As an important part of the *MYB* family, *MYB*-related genes were initially classified into five subgroups: the CCA1-like, CPC-like, TBP-like, I-box-binding-like, R-R-type (Zhang et al., 2018). However, after detailed studies on the GARP family, *MYB*-related genes were divided into three subgroups: CCA1-like/R-R, CPC-like, and GARP-like (Liu et al., 2020). After extensive further studies, a merger occurred in the Circadian Clock Associated 1-like and R-R-type, equivalent to the CCA1-like/R-R subgroups in this study; the results are consistent with those of the most recent studies (Liu et al., 2020). In previous work, CCA1-like proteins have been found to be closely associated with R-R type MYB repeats, thus implying that R-R-type gene loss events may directly lead to CCA1-like proteins. In this classification, 44 members of the CCA1-like/R-R group contained the conserved motif SHAQK (motif 2 in **Supplementary Figure 1**) in the MYB repeat, in agreement with previous studies in Arabidopsis (Ji et al., 2019). The CAPR-like subgroups initially comprised maize G2-like and Arabidopsis ARR-B, which were newly derived in angiosperms. In agreement with previous research, the GARP-like subgroups were classified into six clades in this identification (Liu et al., 2020).

As a typical MYB member, the CPC-like subgroup has been reported to be involved in the basic function of regulating root-hair development and the plant trichome. It has been confirmed to regulate the synthesis of secondary metabolites in recent years. For instance, *AtCPC*, characterized as a repressor of anthocyanin production, has provided novel insight into the transcriptional regulation of flavonoid biosynthesis. *SlTRY* and *SlGL3*, *MYB-related* homologs of tomato, also influence synthesis of anthocyanin (Jian et al., 2019). In contrast, *AtMYBL2*, an R3-MYB protein, has been found to restrain flavonoid synthesis. Interestingly, *MYBL2* inhibits flavonoid biosynthesis and trichome development (Matsui et al., 2008). *LlMYB3*, a *MYB* related homolog from tiger lily, positively regulates the stress tolerance of plants and is involved in the anthocyanin biosynthesis pathway, thereby regulating stress tolerance (Yong et al., 2019).

In this expression analysis, *OfMYB1R70*, *OfMYB1R114*, and *OfMYB1R201*, which were preferentially expressed in floral organs, had higher RPKM values than other *OfMYB1Rs*, particularly *OfMYB1R70* and *OfMYB1R114*, which had the top two RPKM values. Moreover, subcellular localization prediction indicated that they were mainly located in the nucleus, thus suggesting that these candidate genes may play important roles in regulating floral aroma synthesis during the bloom-stage of sweet osmanthus; therefore, they were selected for follow-up experiments. However, no prior studies have focused on the functions of MYB-related TFs in floral aroma biosynthesis. The relationship with the synthesis of floral VOCs has been reported only for 2RMYB (Reddy et al., 2017), although *2RMYB* and *MYB-related* TFs jointly compose the main family in plants and are involved in many biological processes, particularly secondary metabolism, and responses to environmental factors, in divergent ways (Jian et al., 2019; Zhao et al., 2019).

Transcription factors have been reported to regulate the synthesis of floral VOCs by directly binding the promoter regions of target enzyme genes and controlling their expression levels. In *H. coronarium*, *HcMYB* genes directly bind the promoters of key structural volatile synthesis genes and regulate the biosynthesis of terpenoids and benzenoids (Abbas et al., 2021). In *O. fragrans*, *OfERF61* up-regulates the expression of *OfCCD4* and influences the synthesis of β-ionone (Han et al., 2019). In petunia, *PhMYB4* controls the production of floral volatile benzenoid/phenylpropanoid by suppressing the downstream gene transcription (Thomas et al., 2011). *PhERF6* binds the promoters of EOBI protein, which influences floral scent, thus negatively regulating the production of volatile compounds (Liu et al., 2017). In spearmint, *MsMYB* represses the expression of the *GPPS* gene, thus modifying secondary terpene metabolism (Reddy et al., 2017). In this study, we found that *OfMYB1R201* is nuclearly located with self-activation activity its transient overexpression decreases the synthesis of β-ionone (**Figure 7**) and other VOCs in tobacco leaves (**Supplementary Figure 6**). The expression of *NbCCD4.3* was significantly diminished in transformed plants with transient expression of the *OfMYB1R201* vector (**Figure 7**). These results indicated that *OfMYB1R201* could participate in negatively regulating VOC synthesis.

Remarkably, *OfMYB1R70* and *OfMYB1R114* were found to have nuclear localization (**Supplementary Figure 5**), and their expression during flower development was highly correlated (**Figure 4**). The transient transformation results showed that *OfMYB1R70* and *OfMYB1R114* were both involved in the regulation of aroma component synthesis, as confirmed by the markedly greater production of β-ionone and other emitted metabolites than that in the empty vector plants (**Supplementary Figure 6**). The expression of *NbCCD4.1* and *NbCCD4.3* was elevated in plants with transient expression of the *OfMYB1R114* vector, and the expression of *NbCCD4.1* and *NbCCD4.2* was markedly elevated in plants with transient expression of the *OfMYB1R70* vector (**Figure 7**). However, we found that *OfMYB1R70* and *OfMYB1R114* had no transcriptional activity, thus indicating that they do not directly activate the transcription of their downstream genes. Previous studies have reported that TFs interact with other proteins in regulating the synthesis of secondary metabolites by the MBW complex (Koes et al., 2005). For example, *Fh*MYBx, containing a bHLH-binding motif in the R3 domains, suppresses the formation of the MYB-bHLH-WD40 (MBW) complex via an association with bHLH proteins, thus down-regulating anthocyanin metabolism (Li et al., 2020c). *SmelMYBL1*, an R3 MYB repressor in eggplant, acts as an inhibitor of the MBW complex by competing with MYB activators (*SmelANT1* and *SmelAN2*) for the bHLH binding site involved in anthocyanin biosynthesis (Andrea et al., 2020). Hence, we concluded that *OfMYB1R70*/*114*/*201* regulates aroma component formation by interacting with other TFs or acting

as activators in competing pathways that alter the availability of precursors. The molecular mechanism of *OfMYB1R70/114/201* in regulating the aroma metabolite synthesis will be further explored in following studies.

CONCLUSION

In this study, to provide novel insight into the regulation of floral aroma by all *MYB-related* genes in sweet osmanthus, we comprehensively identified 212 putative *MYB-related* members by genome-wide analysis. We categorized them into three subgroups, including 12 categories related to 132 *MYB-related* members of *A. thaliana* in the phylogenetic tree. Additionally, the information on conserved motifs in genes was visualized in figures, and the expression patterns indicated that three candidate *OfMYB1Rs* might be associated with the development and growth of flowers. Notably, members of the CC1-like/R-R subgroup in sweet osmanthus, *OfMYB1R70/114/201*, localized primarily in the nucleus. Remarkably, transient expression studies demonstrated that all those genes were associated with the regulation of floral VOC synthesis, and *OfMYB1R114* had an apparent ability to positively regulate the synthesis of floral VOCs, such as β-ionone and nonanal, whereas *OfMYB1R201*, as a TF, negatively regulated floral VOCs. Together, our findings provide insight into *OfMYB-related* genes, explaining their roles in plants as regulators of the production of volatile aroma compounds. Our future efforts will focus on modulating the functions of *OfMYB1R70/114/201* in regulating the production of floral aroma and investigating the roles of upstream regulatory factors.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**, further inquiries can be directed to the corresponding authors.

AUTHOR CONTRIBUTIONS

YY and XnY conceived the original idea. XnY and XW performed the experimental work. XnY and WD contributed to the data analysis. XnY wrote the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by the National Natural Science Foundation of China (grant nos. 31870695 and 32071828) and the Priority Academic Program Development of Jiangsu Higher Education Institutions (PAPD).

ACKNOWLEDGMENTS

We thank Hannah Thomas for critical editing of the manuscript. We also thank Xiaoyue Ji, a lab master at the Advanced Analysis Testing Center (AATC), Nanjing Forestry University, China, for assistance in the detection of the plants with transient expression.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2022.765213/full#supplementary-material

Supplementary Figure 1 | Sequence logos of the MYB-related domains of subgroups.

Supplementary Figure 2 | The classification and frequency of each subgroup on each chromosome.

Supplementary Figure 3 | Tandemly and segmentally duplicated genes visualized in the *O. fragrans* genome.

Supplementary Figure 4 | The expression patterns of three candidate *OfMYB1Rs*, on the basis of three reference genes (*ACTIN*, *RAN*, and *RBP2*).

Supplementary Figure 5 | Subcellular localization of candidate OfMYB-related proteins.

Supplementary Figure 6 | Floral composition analysis after infiltration with 35S:*OfMYBMYB1R70/114/201*:GFP and control vectors (VIP $>$ 1 and $P < 0.05$).

Supplementary Figure 7 | OPLS-DA plots generated from comparisons among three types of plants with transient expression and negative control vector plants with transient expression (NC).

Supplementary Figure 8 | The process of transient expression in *N. benthamiana*.

REFERENCES

Abbas, F., Ke, Y. G., Zhou, Y. W., Yu, Y. Y., Waseem, M., Ashraf, U., et al. (2021). Genome-wide analysis reveals the potential role of MYB transcription factors in floral scent formation in *Hedychium coronarium*. *Front. Plant Sci.* 12:623742. doi: 10.3389/FPLS.2021.623742

Ambawat, S., Sharma, P., Yadav, N. R., and Yadav, R. C. (2013). MYB transcription factor genes as regulators for plant responses: an overview. *Physiol. Mol. Biol. Plants* 19, 307–321. doi: 10.1007/s12298-013-0179-1

Andrea, M., Francesco, E. F., Sergio, I., Alessandra, G., Maria, M. A., Cinzia, C., et al. (2020). Identification of a new R3 MYB type repressor and functional characterization of the members of the MBW transcriptional complex involved in anthocyanin biosynthesis in eggplant (*S. melongena* L.). *PLoS One* 15:e0232986. doi: 10.1371/journal.pone.0232986

Antje, F., Katja, M., Braun, E. L., and Erich, G. (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

Bailey, T. L., Bodén, M., Buske, F. A., Frith, M. C., Grant, C. E., Clementi, L., et al. (2009). MEME Suite: tools for motif discovery and searching. *Nucleic Acids Res.* 37: W202–W208. doi: 10.1093/nar/gkp335

Baranowskij, N., Frohberg, C., Prat, S., and Willmitzer, L. (1994). A novel DNA binding protein with homology to MYB oncoproteins containing only one repeat can function as a transcriptional activator. *EMBO J.* 13, 5383–5392. doi: 10.1002/j.1460-2075.1994.tb06873.x

Chen, C. J., Chen, H., Zhang, Y., Thomas, H. R., Frank, M. H., He, Y. H., et al. (2020). TBtools: an integrative toolkit developed for interactive analyses of big biological data. *Mol. Plant* 13, 1194–1202. doi: 10.1016/j.molp.2020.06.009

Chen, G. Q., He, W. Z., Guo, X. X., and Pan, J. S. (2021). Genome-wide identification, classification and expression analysis of the MYB transcription factor family in Petunia. *Int. J. Mol. Sci.* 22:4838. doi: 10.3390/IJMS22094838

Ding, W. J., Ouyang, Q. X., Li, Y. L., Shi, T. T., Li, L., Yang, X. L., et al. (2019). Genome-wide investigation of WRKY transcription factors in sweet osmanthus and their potential regulation of aroma synthesis. *Tree Physiol.* 40, 557–572. doi: 10.1093/treephys/tpz129

Du, H., Wang, Y. B., Xie, Y., Liang, Z., Jiang, S. J., Zhang, S. S., et al. (2013). Genome-wide identification and evolutionary and expression analyses of MYB-related genes in land plants. *DNA Res.* 20, 437–448. doi: 10.1093/dnares/dst021

Dubos, C., Stracke, R., Grotewold, E., Weisshaar, B., Martin, C., Lepiniec, L., et al. (2010). MYB transcription factors in Arabidopsis. *Trends Plant Sci.* 15, 573–581. doi: 10.1016/j.tplants.2010.06.005

Finn, R. D., Coggill, P., Eberhardt, R. Y., Eddy, S. R., Mistry, J., and Mitchell, A. L. (2016). The Pfam protein families database: towards a more sustainable future. *Nucleic Acids Res.* 44, D279–D285. doi: 10.1093/nar/gkv1344

Ginglinger, J., Boachon, B., Höfer, R., Paetz, C., Köllner, T., Miesch, L., et al. (2013). Gene co-expression analysis reveals complex metabolism of the monoterpene alcohol linalool in Arabidopsis flowers. *Plant Cell* 25, 4640–4657. doi: 10.1105/tpc.113.117382

Han, Y. J., Wang, H. Y., Wang, X. D., Li, K., Dong, M. F., Li, Y., et al. (2019). Mechanism of floral scent production in Osmanthus fragrans and the production and regulation of its key floral constituents, β-ionone and linalool. *Hortic. Res.* 6, 1359–1380. doi: 10.1038/s41438-019-0189-4

Han, Y. J., Wu, M., Cao, L. Y., Yuan, W. J., Dong, M. F., Wang, X. H., et al. (2016). Characterization of OfWRKY3, a transcription factor that positively regulates the carotenoid cleavage dioxygenase gene OfCCD4 in Osmanthus fragrans. *Plant Mol. Biol.* 91, 485–496. doi: 10.1007/s11103-016-0483-6

Huang, B., Chen, H. Q., and Shao, L. Q. (2015). The ethanol extract of *Osmanthus fragrans* attenuates *Porphyromonas gingivalis* lipopolysaccharide-stimulated inflammatory effect through the nuclear factor erythroid 2-related factor-mediated antioxidant signalling pathway. *Arch. Oral Biol.* 60, 1030–1038. doi: 10.1016/j.archoralbio.2015.02.026

Ji, Q., Wang, D. W., Zhou, J., Xu, Y. L., Shen, B. G., and Zhou, F. (2019). Genome-wide characterization and expression analyses of the MYB superfamily genes during developmental stages in Chinese jujube. *PeerJ* 7:e6353. doi: 10.7717/peerj.6353

Ji, X. Y. (2020). Comparative analysis of volatile organic compounds and bioactive compounds in typical coniferous and broad-leaved tree species. *J. Essent Oil Bearing Plants* 23, 1105–1117. doi: 10.1080/0972060X.2020.1854128

Ji, X. Y. (2021). Comparative investigation of volatile components and bioactive compounds in beers by multivariate analysis. *Flavour Fragr. J.* 36, 374–383. doi: 10.1002/FFJ.3649

Jian, W., Cao, H. H., Yuan, S., Liu, Y. D., Lu, J. F., Lu, W. J., et al. (2019). SlMYB75, an MYB-type transcription factor, promotes anthocyanin accumulation and enhances volatile aroma production in tomato fruits. *Hortic. Res.* 6:22. doi: 10.1038/s41438-018-0098-y

Jin, J. P., Tian, F., Yang, D. C., Meng, Y. Q., Kong, L., Luo, J. C., et al. (2017). PlantTFDB 4.0: toward a central hub for transcription factors and regulatory interactions in plants. *Nucleic Acids Res.* 45, D1040–D1045. doi: 10.1093/nar/gkw982

Koes, R., Verweij, W., and Quattrocchio, F. (2005). Flavonoids: a colorful model for the regulation and evolution of biochemical pathways. *Trends Plant Sci.* 10, 236–242. doi: 10.1016/J.TPLANTS.2005.03.002

Li, H. Y., Yue, Y. Z., Ding, W. J., Chen, G. W., Li, L., Li, Y. L., et al. (2020a). Genome-wide identification, classification, and expression profiling reveals R2R3-MYB transcription factors related to monoterpenoid biosynthesis in *Osmanthus fragrans*. *Genes* 11:353. doi: 10.3390/genes11040353

Li, J. J., Liu, H., Yang, C., Wang, J., Yan, G. J., Si, P., et al. (2020b). Genome-wide identification of MYB genes and expression analysis under different biotic and abiotic stresses in *Helianthus annuus* L. *Indust. Crops and Prod.* 143:111924. indcrop.2019.111924 doi: 10.1016/j

Li, Y. Q., Shan, X. T., Gao, R. F., Han, T. T., Zhang, J., Wang, Y. N., et al. (2020c). MYB repressors and MBW activation complex collaborate to fine-tune flower coloration in *Freesia hybrida*. *Commun. Biol.* 3:396. doi: 10.1038/s42003-020-01134-6

Li, W. F., Ning, G. X., Zuo, C. W., Chu, M. Y., Yang, S. J., Ma, Z. H., et al. (2021). MYB_SH[AL]QKY[RF] transcription factors MdLUX and MdPCL-like promote anthocyanin accumulation through DNA hypomethylation and MdF3H activation in apple. *Tree Physiol.* 41, 836–848. doi: 10.1093/TREEPHYS/TPAA156

Liu, F., Xiao, Z. N., Yang, L., Chen, Q., Shao, L., Liu, J. X., et al. (2017). PhERF6, interacting with EOBI, negatively regulates fragrance biosynthesis in petunia flowers. *New Phytol.* 215, 1490–1502. doi: 10.1111/nph.14675

Liu, Y. H., Zeng, Y. T., Li, Y. M., Liu, Z., Lin, W. K., Espley, R. V., et al. (2020). Genomic survey and gene expression analysis of the MYB-related transcription factor superfamily in potato (*Solanum tuberosum* L.). *Int. J. Biol. Macromol.* 164, 2450–2464. doi: 10.1016/j.ijbiomac.2020.08.062

Marchler-Bauer, A., and Bryant, S. H. (2004). CD-Search: protein domain annotations on the fly. *Nucleic Acids Res.* 32, W327–W331. doi: 10.1093/nar/gkh454

Matsui, K., Umemura, Y., and Ohme-Takagi, M. (2008). AtMYBL2, a protein with a single MYB domain, acts as a negative regulator of anthocyanin biosynthesis in Arabidopsis. *Plant J.* 55, 954–967. doi: 10.1111/j.1365-313X.2008.03565.x

Muleke, E. M., Wang, Y., Zhang, W., Xu, L., Ying, J., Karanja, B. K., et al. (2021). Genome-wide identification and expression profiling of MYB transcription factor genes in radish (*Raphanus sativus* L.). *J. Integr. Agric.* 20, 120–131. doi: 10.1016/S2095-3119(20)63308-1

Nesi, N., Jond, C., Debeaujon, I., and Caboche, M. (2001). The Arabidopsis TT2 gene encodes an R2R3 MYB domain protein that acts as a key determinant for proanthocyanidin accumulation in developing seed. *Plant Cell* 13, 2311–2322. doi: 10.1105/TPC.010098

Nguyen, N. H., and Lee, H. (2016). MYB-related transcription factors function as regulators of the circadian clock and anthocyanin biosynthesis in Arabidopsis. *Plant Signal. Behav.* 11:e1139278. doi: 10.1080/15592324.2016.1139278

Niu, T. Q., Gao, Z. D., Zhang, P. F., Zhang, X. J., Gao, M. Y., Ji, W., et al. (2016). MYBA2 gene involved in anthocyanin and flavonol biosynthesis pathways in grapevine. *Genet. Mol. Res.* 15:gmr15048922. doi: 10.4238/gmr15048922

Reddy, V. A., Wang, Q., Dhar, N., Kumar, N., Venkatesh, P., Rajan, C., et al. (2017). Spearmint R2R3-MYB transcription factor MsMYB negatively regulates monoterpene production and suppresses the expression of geranyl diphosphate synthase large subunit (MsGPPS.LSU). *Plant Biotechnol. J.* 15, 1105–1119. doi: 10.1111/pbi.12701

Riechmann, J., and Ratcliffe, O. (2000). A genomic perspective on plant transcription factors. *Curr. Opin. Plant Biol.* 3, 423–434. doi: 10.1016/S1369-5266(00)00107-2

Rosinski, J. A., and Atchley, W. R. (1998). Molecular evolution of the MYB Family of transcription factors: evidence for polyphyletic origin. *J. Mol. Evol.* 46, 74–83. doi: 10.1007/PL00006285

Rumi, T. W., Yuka, N., and Takuji, W. (2013). Tomato *(Solanum lycopersicum)* homologs of TRIPTYCHON (SlTRY) and GLABRA3 (SlGL3) are involved in anthocyanin accumulation. *Plant Signal. Behav.* 8:e24575. doi: 10.4161/psb.24575

Sakai, M., Yamagishi, M., and Matsuyama, K. (2019). Repression of anthocyanin biosynthesis by R3-MYB transcription factors in lily (*Lilium* spp.). *Plant Cell Rep.* 38, 609–622. doi: 10.1007/s00299-019-02391-4

Sun, X. M., Matus, J. T., Wong, D., Wang, Z. M., Chai, F. M., Zhang, L. L., et al. (2018). The GARP/MYB-related grape transcription factor AQUILO improves cold tolerance and promotes the accumulation of raffinose family oligosaccharides. *J. Exp. Bot.* 69, 1749–1764. doi: 10.1093/jxb/ery020

Thomas, A. C., Joo, Y. K., Ashlyn, E. W., Laura, A. L., Kyle, C. S., Robert, C. S., et al. (2011). PhMYB4 fine-tunes the floral volatile signature of *Petunia×hybrida* through PhC4H. *J. Exp. Bot.* 62, 1133–1143. doi: 10.1093/jxb/erq342

Yang, X. L., Li, H. Y., Yue, Y. Z., Ding, W. J., Chen, X., Shi, T. T., et al. (2018). Transcriptomics analysis of the candidate genes related to aroma formation in *Osmanthus fragrans*. *Molecules* 23:1604. doi: 10.3390/molecules23071604

Yin, J. L., Wong, W. S., Jang, I. C., and Chua, N. H. (2017). Co-expression of peppermint geranyl diphosphate synthase small subunit enhances monoterpene production in transgenic tobacco plants. *New Phytol.* 213, 1133–1144. doi: 10.1111/nph.14280

Yong, Y. B., Zhang, Y., and Lyu, Y. N. (2019). A MYB-related transcription factor from *Lilium lancifolium* L. (LlMYB3) is involved in anthocyanin biosynthesis pathway and enhances multiple abiotic stress tolerance in *Arabidopsis thaliana*. *Int. J. Mol. Sci.* 20:3195. doi: 10.3390/ijms20133195

Yue, Y. Z., Du, J. H., Li, Y., Thomas, H. R., Frank, M. H., Wang, L. G., et al. (2021). Insight into the petunia Dof transcription factor family reveals a new regulator of male-sterility. *Indust. Crops Prod.* 161:113196. doi: 10.1016/J.INDCROP.2020.113196

Zeng, X. L., Liu, C., Zheng, R. R., Cai, X., Luo, J., Zou, J. J., et al. (2016). Emission and accumulation of monoterpene and the key terpene synthase (TPS) associated with monoterpene biosynthesis in *Osmanthus fragrans* Lour. *Front. Plant Sci.* 6:1232. doi: 10.3389/fpls.2015.01232

Zhang, C. H., Ma, R. J., Xu, J. L., Yan, J., Guo, L., Song, J., et al. (2018). Genome-wide identification and classification of MYB superfamily genes in peach. *PLoS One* 13:e0199192. doi: 10.1371/journal.pone.0199192

Zhang, C., Fu, J. X., Wang, Y. G., Bao, Z. Y., and Zhao, H. B. (2015). Identification of suitable reference genes for gene expression normalization in the quantitative real-time PCR analysis of sweet osmanthus (*Osmanthus fragrans* Lour.). *PLoS One* 10:e0136355. doi: 10.1371/journal.pone.0136355

Zhao, P. C., Hou, S. L., Guo, X. F., Jia, J. T., Yang, W. G., Liu, Z. J., et al. (2019). A MYB-related transcription factor from sheepgrass, LcMYB2, promotes seed germination and root growth under drought stress. *BMC Plant Biol.* 19:564. doi: 10.1186/s12870-019-2159-2

Zhou, L., Yarra, R., Jin, L., and Cao, H. (2020). Genome-wide identification and expression analysis of MYB gene family in oil palm (*Elaeis guineensis* Jacq.) under abiotic stress conditions. *Environ. Exp. Bot.* 180:104245. doi: 10.1016/j.envexpbot.2020.104245

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.

frontiers | Frontiers in Plant Science

ORIGINAL RESEARCH
published: 25 April 2022
doi: 10.3389/fpls.2022.828042

A Comparative Study of Flavonoids and Carotenoids Revealed Metabolite Responses for Various Flower Colorations Between *Nicotiana tabacum* L. and *Nicotiana rustica* L.

Qinzhi Xiao[1,2,3], Yueyi Zhu[3], Guoxian Cui[2], Xianwen Zhang[4], Risheng Hu[1], Zhengyu Deng[1], Lei Lei[5], Liwen Wu[6] and Lei Mei[3,7]*

[1] Yongzhou Tobacco Monopoly Bureau of Hunan, Yongzhou, China, [2] College of Agronomy, Hunan Agricultural University, Changsha, China, [3] College of Agriculture and Biotechnology, Zhejiang University, Hangzhou, China, [4] Institute of Virology and Biotechnology, Zhejiang Academy of Agricultural Sciences, Hangzhou, China, [5] College of Plant Science and Technology, Huazhong Agricultural University, Wuhan, China, [6] College of Bioscience and Technology, Hubei Minzu University, Enshi, China, [7] Department of Plant Sciences, University of Cambridge, Cambridge, United Kingdom

Tobacco is a model plant for studying flower coloration. Flavonoids and carotenoids were reported to contribute to the flower color in many plants. We investigated the mechanism underlying flower color formation in tobacco by comparing the profiling flavonoids and carotenoids between various species *Nicotiana tabacum* L. and *Nicotiana rustica* L., as their flowers commonly presented red (pink) and yellow (orange), respectively. The metabolomes were conducted by UPLC–ESI–MS/MS system. The main findings were as follows: (1) A total of 31 flavonoids and 36 carotenoids were identified in all four cultivars involved in *N. tabacum* and *N. rustica*. (2) Flavonoids and carotenoids tended to concentrate in the red flowers (*N. tabacum*) and yellow flowers (*N. rustica*), respectively. (3) About eight flavonoids and 12 carotenoids were primarily screened out for metabolic biomarkers, such as the robust biomarker involving kaempferol-3-*o*-rut, quercetin-glu, rutin, lutein, and β-carotene. This is the first research of systematic metabolome involving both flavonoids and carotenoids in tobacco flower coloration. The metabolic mechanism concluded that flavonoids and carotenoids mainly contributed to red (pink) and yellow (orange) colors of the tobacco flowers, respectively. Our finding will provide essential insights into characterizing species and modifying flower color in tobacco breeding through genetic improvement or regulation of featured metabolic synthesis.

Keywords: floral color, *Nicotiana tabacum* L., *Nicotiana rustica* L., flavonoid, carotenoid

OPEN ACCESS

Edited by:
Xiumin Fu,
South China Botanical Garden (CAS), China

Reviewed by:
Bi Ma,
Southwest University, China
Yu Han,
Beijing Forestry University, China

***Correspondence:**
Lei Mei
meileihzruk@gmail.com

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

Received: *02 December 2021*
Accepted: *24 February 2022*
Published: *25 April 2022*

Citation:
Xiao Q, Zhu Y, Cui G, Zhang X, Hu R, Deng Z, Lei L, Wu L and Mei L (2022) A Comparative Study of Flavonoids and Carotenoids Revealed Metabolite Responses for Various Flower Colorations Between Nicotiana tabacum L. and Nicotiana rustica L. Front. Plant Sci. 13:828042. doi: 10.3389/fpls.2022.828042

INTRODUCTION

As one of the five major genera of the family Solanaceae, the genus *Nicotiana* consists of 64 species, of which 44 and 20 species are considered to be originated in the America and Oceania continent, respectively (Goodspeed, 1954). According to the morphologies, chromosomal number, cross-ability relationships, and features in interspecific hybrids, the species under genus *Nicotiana* were further classified into three sub-genera *tabacum*, *rustica*, and *petunioides*, among which *N. tabacum* and *N. rustica* were the main kinds of cultivars (Siva Raju et al., 2008; Huang et al., 2021). Generally, the red (pink) and yellow (orange) flowers were commonly phenotyped in *N. tabacum* and *N. rustica*, respectively. Currently, the mechanisms underlying the coloration in tobacco are not clear. Floral color can protect floral organs and attract pollinators. Notably, it was a crucial trait in resolving the evolutionary relationships between species. The coloration was determined by the types and amounts of pigments, which were also influenced by the internal or surface tissue structure of the petals (Zhao and Tao, 2015). The pigment components of colorful flowers have been studied for over 150 years, since the mid-19th century. Carotenoids and flavonoids are the two primary types of substances that give color to flowers (Maoka, 2020; Mekapogu et al., 2020).

Plants produce more than 1,000,000 metabolites in prediction (Afendi et al., 2012). Flavonoids are one of the major groups of specialized metabolites and by far the largest class of polyphenols. Currently, flavonoids are estimated to include more than 8,000 compounds (Yonekura-Sakakibara et al., 2019; Wen et al., 2020). The flavonoids process a common backbone of diphenylpropane (C6–C3–C6), where the two aromatic rings are linked *via* a three-carbon chain (Wen et al., 2020). One of the rings is originated from resorcinol or phloroglucinol produced by the acetate pathway (Perez et al., 2020). Correspondingly, another ring is formed from the shikimate pathway (Saito et al., 2013). Generally, flavonoids are divided into six subclasses, that is, flavones, isoflavones, flavonols, flavonols, flavanones, and anthocyanidins (Harbone, 1993). Flavonoids are the most critical pigments in plants, which exhibit the broadest spectrum of colors.

Carotenoids are a group of natural tetraterpenoid pigments synthesized by plants, bacteria, algae, and fungi (Fuentes et al., 2012; Avalos and Carmen Limón, 2015; Sun et al., 2018). Thus far, more than 1,100 natural carotenoids have been reported (Yabuzaki, 2017; Quian-Ulloa and Stange, 2021). In plants, carotenoids play vital roles in photoprotection and photosynthesis, as well as the precursors for phytohormones such as abscisic acids and strigolactones (Nambara and Marion-Poll, 2005; Domonkos et al., 2013; Niyogi and Truong, 2013; Al-Babili and Bouwmeester, 2015; Hashimoto et al., 2016). Furthermore, carotenoids also serve as signaling molecules in plant developments and respond to environments (Havaux, 2014; Tian, 2015; Hou et al., 2016). According to the structures, carotenoids can be divided into two classes, namely, carotenes (hydrocarbons without oxygen) and xanthophylls (hydrocarbons with oxygen). As the derivatives of tetraterpenes, carotenoids are molecules of a C40 polyene backbone with an ionone ring at the end. Based on this structure, carotenoids can absorb visible light of short wavelengths ranging from 400 to 550 nm (violet to green light). The number and properties of the double bonds determined the wavelength of light absorbed. Hence, the compounds are colored yellow, orange, or red.

Biomarkers are defined to predict phenotypical properties while these features are not apparent, which are valuable tools for life research (Meyer et al., 2007; Collard and Mackill, 2008; Geifman-Holtzman and Ober, 2008). Generally, the conventional biomarkers are defined as genetic markers, which are comprehensively applied to identify specific genetic lines with positive traits. However, to some genetic markers in plants, it's challenging for species with complex polyploid genomes. Metabolic profiles in a biological system are the endpoints in omics involving genome, transcriptome, and proteome (Khamis et al., 2017). Hence, the exploration of new metabolic biomarkers would provide a new approach to identifying such species due to not being based on genomics. Significantly, some tobaccos, such as allotetraploid *N. tabacum*, have complicated genomes.

Tobacco was considered a vital model plant for studying flower coloration. To date, the previous studies of flower coloration in tobacco commonly focused on and limited to a few genes involving anthocyanin biosynthesis (Li et al., 2017; Park et al., 2020), or transcriptome and metabolome engaged in the same *Nicotiana* species (Jiao et al., 2020). The color of flowers results from the combined action of various metabolites, while the close and same color of flowers may be determined by different metabolic components (Zhao and Tao, 2015). However, red (pink) and yellow (orange) are the two most common flower colors in tobacco. They are also the typical flower colors of different species, such as *N. tabacum* and *N. rustica*, respectively. Hence, systematically identifying the metabolites would provide insights into the mechanisms underlying the formation of differentiated flower color. Moreover, the exploration of related metabolic biomarkers would offer alternative tools for the identification of tobacco species.

MATERIALS AND METHODS

Plant Materials and Sampling

Species *N. tabacum*, including two cultivars No-2146 and Xiangyan-3, were employed in the experiments, and they presented red (pink) flowers. For convenience, the two cultivars were named RFA and RFB in this study, respectively. Parallelly, species *N. rustica,* involving two cultivars Laolaihuang and Xiaoyelanhua, showed both yellow (orange) flowers and called YFA and YFB in this study. Additionally, No-2146 were cultivated by the Hunan Tobacco Monopoly Bureau as lineage "{[(G28×Coker176)×(K326×Gexin-3)]×K326}×K32." G28, Coker176, K326, Gexin-3, and K32 are common *N. tabacum* species in China. Xiangyan-3 was the progeny of K394×Yunyan-87. Cultivars Laolaihuang and Xiaoyelanhua were *N. rustica* species and commonly used in many laboratories. The tobacco seeds that were employed in this study originated from the Yongzhou Tobacco Germplasm Resource Database (YTGRD, Yongzhou 425000, P. R. China). The plants were grown in a glasshouse at the Agricultural Experiment Station, Zhejiang

University. Seeds of the four varieties were germinated in glass Petri dishes before wet filter papers were paved. The dishes containing seeds were transferred to the program ray radiation incubator until the green shoots appeared. The conditions involved light intensity of 100 mol/m^2/s, light/dark 14 h/10 h, 22°C, and humidity 60%. Consequently, the germinated seeds transfer onto the pots filled with quartz sand. The pots were placed in a greenhouse under 300 mol/m^2/s, light/dark 14 h/10 h, 28°C, and humidity 60%. The pots were watered every 3 days and fertilized every 9 days. The liquid fertilizer was one-fourth of the nutrient solution concentration (Bovet et al., 2006). The redundant seedlings were removed and retained one seedling in each pot. On the 20th day after the first blooming flower appeared, the petals of all the four varieties were collected and kept in a −80°C refrigerator for use. Three biological replicates were taken from individual plants within every variety.

Extraction and Analysis of the Flavonoids and Carotenoids

For extraction of flavonoids, the freeze-dried samples were ground into powder by a mixer mill (MM,400, Retsch, Germany) for 2 min at 30 Hz. Next, the 100 mg powdered sample was mixed with 1.0 ml methanol/water/hydrochloric acid (500:500:1, v/v/v), and then overnight at 4°C. Consequently, ultrasound was applied to the extract for 8–10 min and centrifuged at 13,000 *g* at 4°C for 8 min. The supernatants were collected and filtrated with a membrane filter (0.22 μm pore size, Anpel, Shanghai, China). The final filtrate was kept for LC–MS/MS analysis. For extraction of carotenoids, the freeze-dried samples were homogenized and powdered in the same mill. A 100 mg powdered sample was extracted with 1 ml of compound solution of hexane/acetone/ethanol (2:2:1, v/v/v) and extra 0.1% butylated hydroxytoluene. Then the internal standards were added. The extract was vortexed for 15 min at room temperature and consequently centrifuged at 13,000 g at 4°C for 8 min. The supernatants were collected, and the residue was re-extracted at the same conditions. The final supernatant liquid was evaporated under a nitrogen gas stream. Next, the dried residues were dissolved by 200 μl of compound solution of (ethyl nitrile/methyl alcohol)/methyl tert-butyl ether (17(3/1)/3, v/v/v). Finally, the solution was filtered by more than 0.22-μm filter for the upcoming analysis.

Detection of the Flavonoids and Carotenoids

Flavonoids and carotenoids were detected based on the platform UPLC–ESI–MS/MS system (UPLC'ExionLC AD'https://sciex.com.cn/; MS'Applied Biosystems 6500 Triple Quadrupole, https://sciex.com.cn/), and UPLC–APCI–MS/MS system (UPLC'ExionLC AD https://sciex.com.cn/; MS, Applied Biosystems 6500 Triple Quadrupole, https://sciex.com.cn/), respectively. The analytical conditions of flavonoids were shown as HPLC: column, Waters ACQUITY UPLC HSS T3 C18 (1.8 μm, 100 × 2.1 mm); Solvent A, water, 0.04% acetic acid; Solvent B (acetonitrile, 0.04% acetic acid); Gradient program, 100:0, 5:95, 5:95, 95:5, and 95:5 of V(A)/V(B) at 0, 11.0, 12.0, 12.1, and 15.0 min, respectively; Flow rate, 0.40 ml/min; Temperature, 40 °C; Injection volume, 5 μl. Parallelly, the analytical conditions of carotenoids were followed, LC: column, YMC C30 (3 μm, 2.0 mm × 100 mm); Solvent A, methanol acetonitrile (1/3, v/v) with 0.1% formic acid and 0.01% butylated hydroxytoluene (BHT); Solvent B, methyl tert-butyl ether with 0.01% BHT; Gradient program, 0%, 70%, 95%, 0% Solvent B were used for 3, 2, 4, and 1 min, respectively. Flow rate, 0.8 ml/min; Temperature, 28°C; Injection volume, 2 μl. All 12 samples' reproducibility was evaluated *via* principal component analysis (PCA). Unsupervised PCA was performed by R (www.r-project.org). The data were unit variance scaled before unsupervised PCA.

Hierarchical Cluster Analysis and Screening of Metabolic Biomarkers

The hierarchical cluster analysis in terms of metabolites and samples was exhibited as heatmaps combined with dendrograms, which were conducted by *R* package heatmap involving algorithm pheatmap (m, scale = "row") and the normalized signal intensities of metabolites, as unit variance scaling are visualized as a color spectrum. In the experiments, the metabolic biomarkers were screened out based on the following criteria: (1) In the same sample, the content of metabolic biomarker is at least two times higher than that of other metabolites; therefore, the metabolic biomarker can be distinguished from other metabolites in the same sample; (2) For metabolic biomarkers, the highest content value was equal or more than two times of lowest content value among the differentiated groups aimed at, that is, the contents of metabolic biomarkers varied significantly among different samples.

Differential Metabolites Screened and KEGG Annotation and Enrichment Analysis

Based on the orthogonal projections to latent structures discrimination analysis (OPLS-DA) (Thevenot et al., 2015), the variable importance in projection (VIP) of the obtained multivariate analysis of OPLS-DA, was able to preliminarily screen out metabolites that differ in different varieties or tissues. At the same time, the Fold_Change of the univariate analysis further screened out the difference in metabolism. Combining Fold_Change and the VIP value of the OPLS-DA model, differential metabolites were filtered. Filter criteria are as follows: (1) Select metabolites with Fold Change $\geq$ 2 and Fold Change $\leq$ 0.5. If the difference of metabolites is more than 2 or less than 0.5 times between the control and experimental groups, the difference is considered significant; (2) Select metabolites with VIP$\geq$1, first. The VIP value represents the influence strength of the corresponding metabolite between groups, and it is generally considered that the metabolites with VIP $\geq$ 1 are significantly different. Schematics of flavonoids and carotenoid synthesis were constructed according to literature involving plants (Liu et al., 2020). The identified metabolites were mapped onto the two pathways and checked with the KEGG pathway database (http://www.kegg.jp/kegg/pathway.html) (Kanehisa and Goto, 2000).

RESULTS

Morphological Characterization of Species and the Quality Assessment of Metabolomes

The flower color and petal shape are vital traits to characterize *Nicotiana* germplasms. RFA (No-2146) and RFB (Xiangyan-3) were typical genera of *N. tabacum*, which both showed red (pink) flowers (shown in **Figure 1**). Parallelly, YFA (Laolaihuang) and YFB (Xiaoyelanhua) were representative genera of *N. rustica*, delivering yellow (orange) flowers. The petals of RFA and RFB showed with sharp tips, and the shapes of petals involved in YFA and YFB presented round or curved edges. The flower of RFA and RFB exhibited an obvious pentagram from the quarter view, whereas YFA and YFB gave a specious ring. As a whole, RFA and RFB demonstrated the more extensive leaf areas, whereas the smaller leaf area appeared within YFA and YFB. Compared with *N. tabacum*, there were more secondary branches within *N. rustica*.

Features of ion chromatogram on flavonoids and carotenoids demonstrate that the APCI–MS/MS and ESI–MS/MS produced reliable and high-quality data. The extracted ion chromatogram exhibited independent and sharp peaks of flavonoids (**Supplementary Figure 1-1A**). The targeted ingredients were isolated successfully. Additionally, overlapping spectra of total ion chromatogram of quality control samples in positive ion mode indicated the isolations were reliable (**Supplementary Figure 1-1B**). The normalized area of ion intensity reflected the individual ingredient content (**Supplementary Figure 1-1C**). All these indicated that the LC–MS Assay System was working correctly, and the high-quality data were produced perfectly. Parallelly, the assay system involved in carotenoids demonstrated reliable circumstances (**Supplementary Figure 1-2**).

FIGURE 1 | Morphology of the four *Nicotiana* cultivars. **(A)** and **(B)** *Nicotiana tabacum* cultivars RFA (No-2146) and RFB (Xiangyan-3) and **(C)** and **(D)** *Nicotiana rustica* cultivars YFA (Laolaihuang) and YFB (Xiaoyelanhua). The small patches placed onto each board at the right top were featured flowers, correspondingly. The red bar shows 1 cm.

Profiles of Components Involved in Flavonoids and Carotenoids Between the Species

To qualify the metabolites of flavonoids, 108 types of the standard solution were made involving the concentrations of gradients 0.01, 0.02, 0.05, 0.1, 0.5,1, 5, 10, 50, 100, 500, 1,000, 2,000, and 5,000 ng/ml. Corresponding quantitative signals of each concentration standard were obtained. The standard curves of different substances were drawn based on a horizontal coordinate of standard concentrations and longitudinal coordinates of peak area. The standard curve linear equations and related coefficients of the targeted substances were shown in **Supplementary Table 1**. All 108 coefficients *rs* were higher than 0.997. It is clear that the linear equations were reliable. Similarly, regarding carotenoids, the concentrations of gradients included 0.001, 0.005, 0.01, 0.05, 0.1, 0.5, 1, 5, 10, 50, 100, 250, and 400 μg/ml. The standard curves were drawn based on the horizontal coordinates involving the concentration ratios of external and internal standards, and the longitudinal coordinates in terms of the concentration ratios of the area ratio of external and internal standards. A total of 68 standard curve linear equations were obtained, whose coefficients *rs* were higher than 0.991. The *rs* indicated that the standard equations in terms of carotenoids were reliable, although the values of *r* were lower than those within flavonoids. From the results on the PCA analysis of all samples (**Supplementary Figure 2**), it is illustrated that the samples belonging to the same cultivar had good reproducibility of these samples as given flocking together. Overall, cultivars RFA and RFB exhibited better reproducibility than that between YFA and YFB.

Among 12 samples involved in the four species, 31 metabolites in flavonoids were identified and quantified as **Figure 2**. These metabolites are of seven types of anthocyanins, including cyanidin, delphinidin, malvidin, pelnidin, peonidin, and petunidin. Meanwhile, there were other specific flavonoids such as kaempferol, naringenin, and quercetin. In total, 13 substances were quantified in all four cultivars. As a whole, metabolites belonging to anthocyanins showed lower levels than those of other flavonoids. Remarkably, kaempferol-3-o-rut, quercetin-glu, and rutin contents were profoundly higher than those of the remaining metabolites. In the cultivars, RFA and RFB, the values of kaempferol-3-o-rut were as high as 939.84 and 683.11 μg/g, respectively. However, kaempferol-3-o-rut contents were 42.30 and 26.67 μg/g within YFA and YFB, respectively. Regarding quercetin-glu, the values were 282.16, 270.20, 87.51, and 100.30 μg/g in the above cultivars, respectively. Additionally, the means of rutin involved in the four cultivars were 1,014.90, 1,008.40, 347.69, and 399.78 μg/g. Overall, the values of metabolites in terms of anthocyanins fell into the range from 0.03 to 18.39 μg/g in all samples, which showed low levels. In anthocyanins, more sub-types of metabolites including cyanidin, delphinidin, pelnidin, peonidin

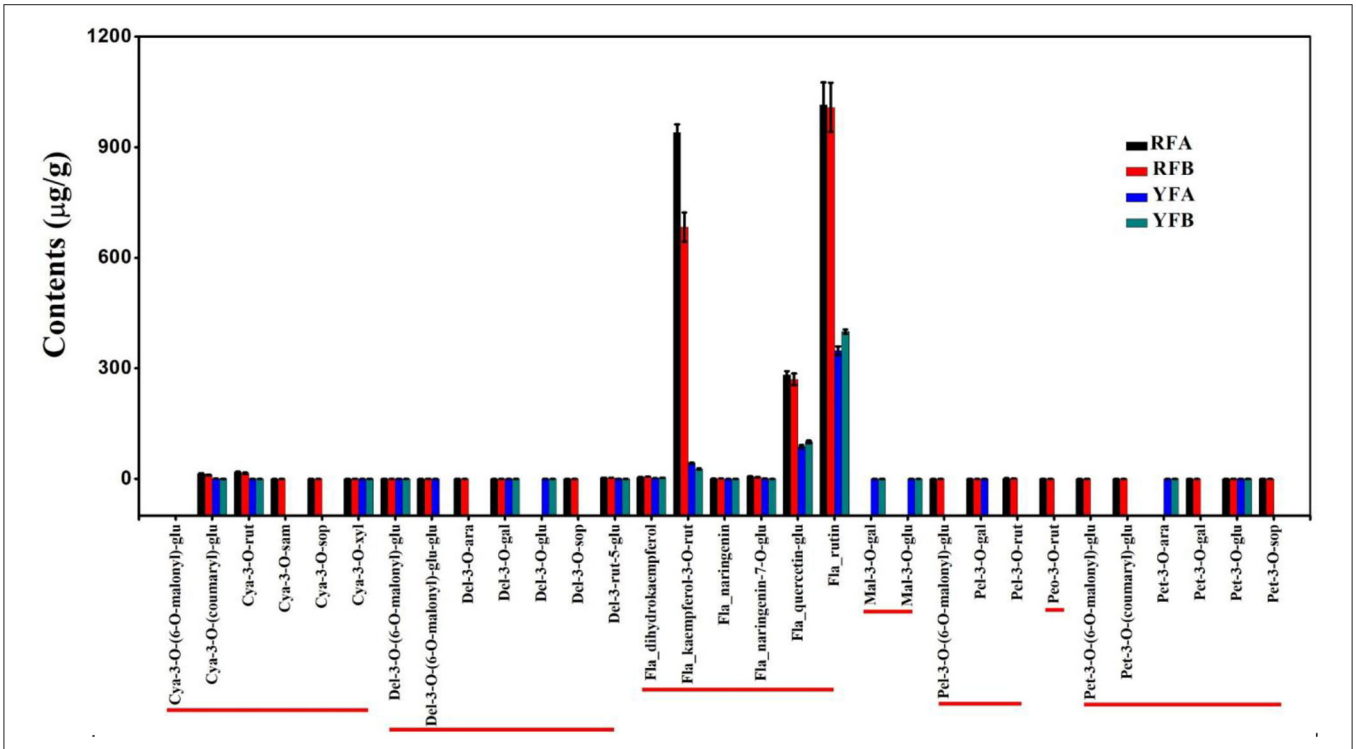


FIGURE 2 | Histogram of flavonoids detected in the four cultivars. RFA and RFB denote the *Nicotiana tabacum* cultivars No-2146 and Xiangyan-3, respectively. YFA and YFB denotes the *Nicotiana rustica* cultivars Laolaihuang and Xiaoyelanhua, respectively. Cya, Del, Fla, Mal, Pel, Peo, and Pet are abbreviated from cyanidin, delphinidin, flavonoid, malvidin, pelnidin, peonidin, and petunidin, respectively.

and petunidin, except for malvidin, were identified in cultivars RFA and RFB, successfully. Overall, cultivars RFA and RFB presented more and higher accumulation of flavonoids than cultivars YFA and YFB.

Similarly, 36 carotenoids were finally identified from the four cultivars, shown in **Figure 3**. A total of 24 metabolites were quantified entirely in all four cultivars. Significantly, the lutein and β-carotene demonstrated comparatively super-higher values. Within the cultivars, FRA, FRB, YFA, and YFB, the means of lutein were 151.00, 177.00, 524.00, and 473.00 μg/g, respectively. Moreover, the contents of β-carotene in related cultivars were 25.60 27.60, 90.10, and 101.00 μg/g. A total of 11 carotenoids were successfully identified in the flowers of YFA and YFB, while there were no values in those of RFA and RFB. They were lycopene, phytoene, lutein–laurate, γ-carotene, lutein–oleate, rubixanthin–laurate, violaxanthin–dioleate, violaxanthin–laurate, violaxanthin–myristate–oleate, β-cryptoxanthin–laurate, and β-cryptoxanthin–oleate. Moreover, the phytoene presented higher values with 67.90 and 78.10 μg/g in the cultivars YFA and YFB, respectively. As a whole, more types and higher levels of carotenoids were identified from cultivars YFA and YFB, compared with those in cultivars RFA and RFB.

Hierarchical Cluster Analysis Among Species

A total of 31 flavonoids and 36 carotenoids in 12 samples were hierarchically clustered, as shown in **Figure 4**. The samples exhibited the same topological structures involving both flavonoids and carotenoids. In detail, replicates FRA-1 and FRA-2, FRB-1 and RFB-2, YFA-1 and YFA-2, plus YFB-1 and YFB-2 showed intimate hierarchical relationships. The replicates belonging to the same cultivars were clustered together. RFA and RFB were close to each other in the four cultivars, as the same case appeared on YFA and YFB. The two groups of species exhibited the same hierarchical level, that is, the metabolisms rarely diverged between the species RFA and RFB or YFA and YFB based on flavonoids and carotenoids. However, the group involving RFA and RFB was differentiated from the YFA and YFB in terms of flavonoids and carotenoids. Metabolites that appeared to have a higher and varied accumulation between the two groups can be screened out as markers, under the premise of consistent accumulation in the individual group. Regarding flavonoids, Del-3-rut-5-glu showed a higher level in the species of RFA and RFB, yet a lower level in YFA and YFB. Consequently, it can be chosen as a marker. Similarly, another marker belonging to anthocyanin was Pet-3-o-glu. Furthermore, the other flavonoid markers included quercetin-glu, dihydrokaempferol, kaempferol-3-*o*-rut, naringenin, naringenin-7-*o*-glu, and rutin. As shown in **Figure 4B**, the carotenoids are of three types, namely, carotenes, carotenoid esters, and xanthophylls. The metabolites involved in the sort of xanthophylls demonstrated more uniform and significant accumulation between two species groups. Overall, the markers in terms of carotenoids included α-carotene, β-carotene, lutein–dimyristate, lutein—palmitate, violaxanthin–palmitate, apocarotenal, echinenone, lutein, neoxanthin, violaxanthin, α-cryptoxanthin, and β-cryptoxanthin.

Comparative Analysis of the Accumulation of Metabolites

Differential metabolites screened in cultivars are shown in **Supplementary Figure 3**. The comparative analysis of flavonoids and carotenoids was established as a Venn diagram (**Figure 5**). Regarding flavonoids, only a single metabolite (Pel-3-O-rut and lutein–dimyristate, respectively) exhibited a different accumulation between cultivars RFA and RFB. Similarly, there was no significant enrichment of metabolite between YFA and YFB. Parentally, there was rare variation between the cultivars in the same species. On the contrary, 11 metabolites in terms of flavonoids had a varied accumulation between RFB and YFB. Correspondingly, the numbers of metabolites variously enriched between RFA and YFA were 11, too. A total of 10 metabolites were diverging between RFA_*vs*_YFA and RFB_*vs*_YFB. The data illustrated that the flavonoids of cultivars belonging to different species were various. Only one carotenoid presented the same accumulation between cultivar RFA and RFB. Three carotenoids were significantly enriched between cultivars YFA and YFB. Analogously, there were fewer carotenoids that can be accumulated between the cultivars belonging to the same group. About 19 and 20 differentiated abundant carotenoids appeared within RFA_*vs*_YFA and RFB_*vs*_YFB, respectively. Seven essential carotenoids were significantly accumulated between the species.

Pathways of Flavonoid and Carotenoid Synthesis

Initiated from *ρ*-coumaroyl-CoA and malonyl-CoA, flavonoids were synthesized, involving key intermediates chalcones, flavanones, dihydroflavonols, and leucoanthocyanidins. Meanwhile, flavonols were derived from the branch of dihydroflavonols (**Figure 6**). Three classes of flavonoids, including cyanidins, delphinidins, and petunidins, were mapped onto the pathway. cya-3-*o*-(coumaryl)-glu and cya-3-rut represented exactly similar features in violin plot among four cultivars, and they showed the close higher and lower values within the same species. In other words, cyanidins quickly accumulated in the cultivars RFA and RFB, compared with YFA and YFB. Because of pet-3-*o*-glu, it seemed that petunidins kept in line with cyanidin involved in species in the anthocyanins synthesis pathway of tobacco. The delphinidins del-3-*o*-(6-*o*-malony)-glu and del-3-rut-5-glu showed similar profiles with other above-mentioned cyanidins in this pathway. However, del-3-*o*-gal demonstrated the specific values in cultivars RFA and YFA, as higher accumulated in YFA, nor RFA. Three flavanones, naringenin, naringenin-7-*o*-glu, and dihydrokaempferol, were mapped on the pathway. Naringenin was rarely enriched in cultivars YFA and YFB, while those in RFA and RFB showed higher values and variations. Compared with that in RFA, naringenin in sample replicates of RFB presented closer maximum and minimum. Naringenin-7-*o*-glu were highly enriched in the cultivars RFA and RFB, while strictly

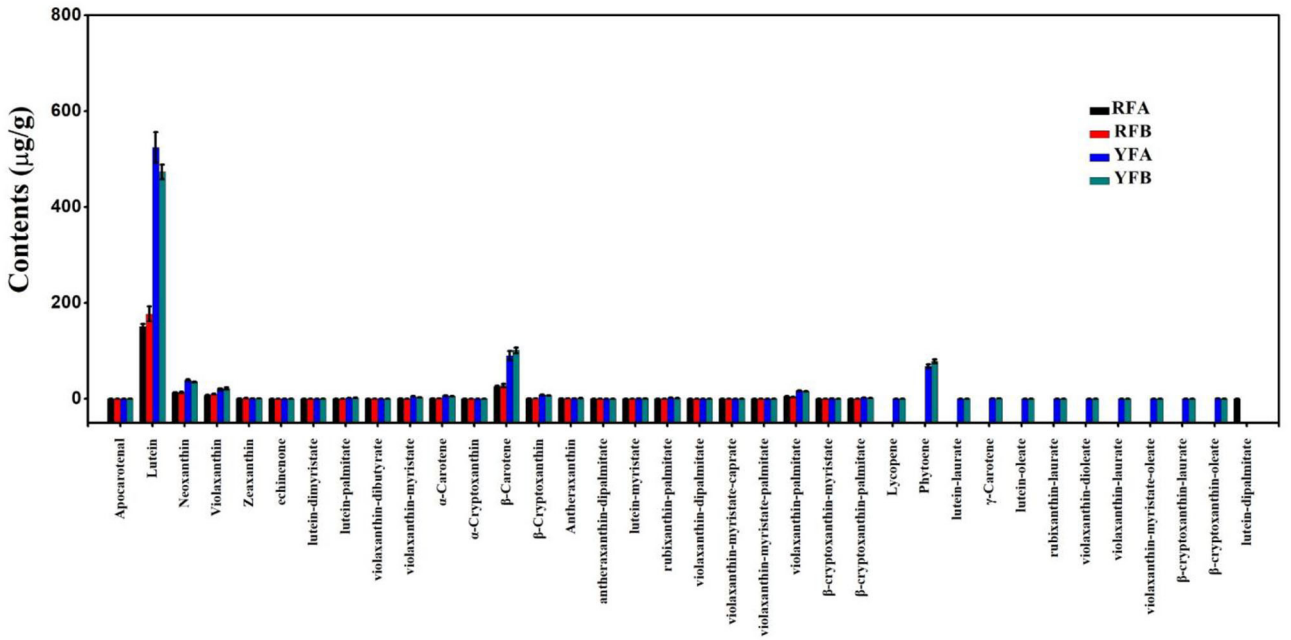


FIGURE 3 | Histogram of carotenoids detected in the four cultivars. RFA and RFB denote the *N. tabacum* cultivars No-2146 and Xiangyan-3, respectively. YFA and YFB denotes the *Nicotiana rustica* cultivars Laolaihuang and Xiaoyelanhua, respectively.

lowly accumulated in YFA and YFB. Three flavonols involving kaempferol-3-*o*-rut, quercetin-glu, and rutin consistently demonstrated accumulation in all cultivars. Again, the higher and lower values were detected in the group of RFA and RFB and the group of YFA and YFB, respectively.

In the carotenoid synthesis pathway, two main branches involving δ-carotene and γ-carotene have diverged from lycopene. The δ-carotene routine included α-carotene, zeinoxanthin, lutein, and 5, 6-epoxylutein. Correspondingly, the γ-carotene routine involved β-carotene, zeaxanthin, antheraxanthin, violaxanthin, and neoxanthin. Totally, eight carotenoids were mapped onto the carotenoid synthesis pathway. As a whole, the carotenoids were highly synthesized in the cultivars YFA and YFB. δ-Carotene was enriched more than five times in the group of YFA and YFB, compared with those in the group of species RFA and RFB. Four kinds of luteins, including lutein, lutein–dimyristate, lutein–myristate, and lutein—palmitate, were detected in the pathway. These four types of lutein showed more than double accumulation in the cultivars YFA and YFB, compared with RYA and RYB. Regarding another branch, there was no apparent increase or decrease in γ-carotene between the species RFA and YFA, given the bars of the replicates. However, their next-step derivative β-carotene presented more than two times increase in the cultivars YFA and YFB, compared with those in the group of species RFA and RFB. Interestingly, the synthesis of zeaxanthin showed very close values in the four cultivars, which agreed with γ-carotene. Two types of antheraxanthin were detected in this pathway, and they demonstrated slight decline in YFA and YFB, compared with those in RFA and RFB. Seven types of violaxanthin were mapped onto the routine of γ-carotene. Except for violaxanthin–dipalmitate and violaxanthin–myristate–caprate, apparent increases were observed in the remaining five types of violaxanthin in terms of the cultivars YFA and YFB. Regarding neoxanthin, the profiles generally kept agreeing with other carotenoids in this pathway.

DISCUSSION

The Red Color of Flowers That Appeared in Species *N. tabacum* Are Mainly Contributed by Flavonoids

In the study, the flavonoids were ultra-highly accumulated in the red flowers of species *N. tabacum*, which involved cultivars RFA (No-2146) and RFB (Xiangyan-3). On the contrary, the flavonoids were rarely enriched in the yellow flowers of species *N. rustica*, which included cultivars YFA (Laolaihuang) and YFB (Xiaoyelanhua). Flavonoids are the most common pigment group and present a broad spectrum of colors from pale yellow to blue-pure, and it is reported the most vital pigments in petals (Zhao and Tao, 2015). Flavonoids were detected in various plant flowers, such as groundcover rose, herbaceous peony, dahlia, violet, and so on (Schmitzer, 2010; Tatsuzawa et al., 2012; Thill et al., 2012; Zhao et al., 2012). Chalcone isomerase (CHI, EC 5.5.1.6) is a vital enzyme in the pathway of plant flavonoids biosynthesis, as it catalyzes the cyclization of chalcones and consequently produces flavanones. It is a practical approach to change the composition of flavonoids *via* downregulating or enhancing the expression of *CHI*. For example, tomato flavonols

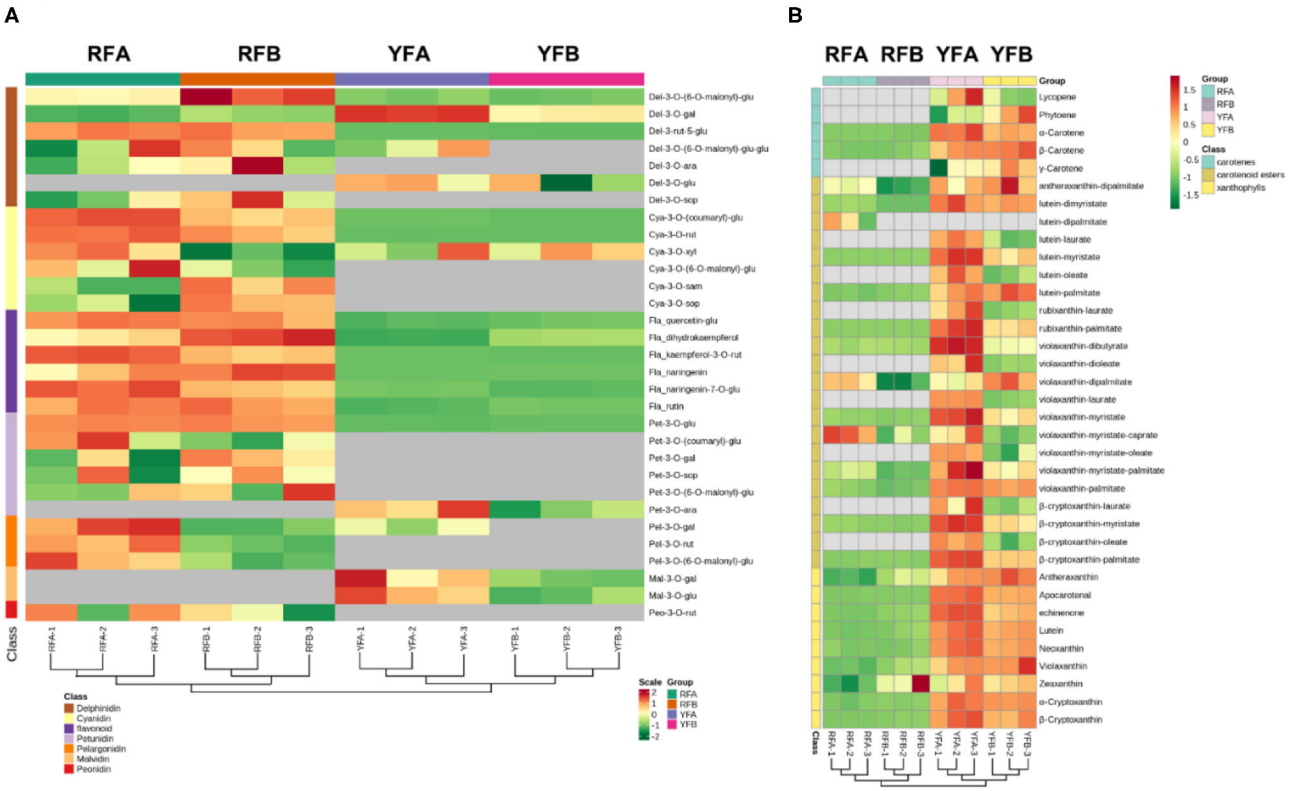


FIGURE 4 | Cluster analysis on the samples and metabolites depending on classes. **(A)** and **(B)** demonstrate features of clusters in terms of flavonoids and carotenoids. RFA and RFB denote the *Nicotiana tabacum* cultivars No-2146 and Xiangyan-3, respectively. YFA and YFB represents the *Nicotiana rustica* cultivars Laolaihuang and Xiaoyelanhua, respectively.

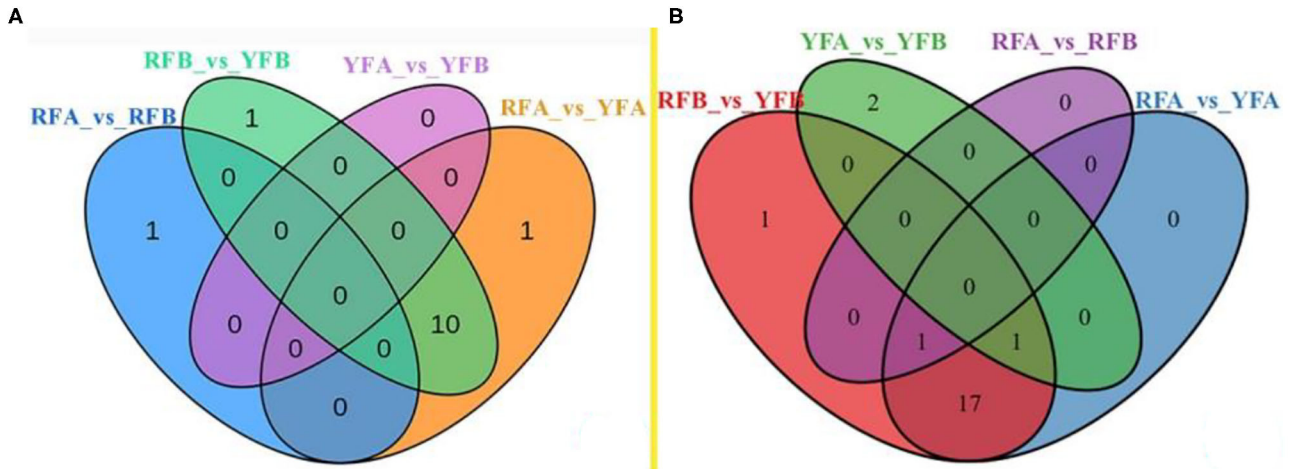


FIGURE 5 | Venn diagram on the overlapping and species-specific differential metabolites. **(A)** and **(B)** were involved in flavonoids and carotenoids, respectively. RFA and RFB, and YFA and YFB represent *Nicotiana tabacum* cultivars No-2146 and Xiangyan-3, and *Nicotiana rustica* cultivars Laolaihuang and Xiaoyelanhua, respectively.

contents were increased after the chi was overexpressed (Muir et al., 2001). Moreover, the flower colors could be modified *via* the changes of flavonoids based on other vital enzymes such as chalcones synthase (CHS) (Mol et al., 1999; Forkmann and Martens, 2001). In *Nicotiana*, the altered composition of flavonoids resulted in the change of flower color. The transgenic tobacco (*Nicotiana tabacum* L. cv. SR1) by RNAi transformation causes the suppression of CHI and consequently disappearance

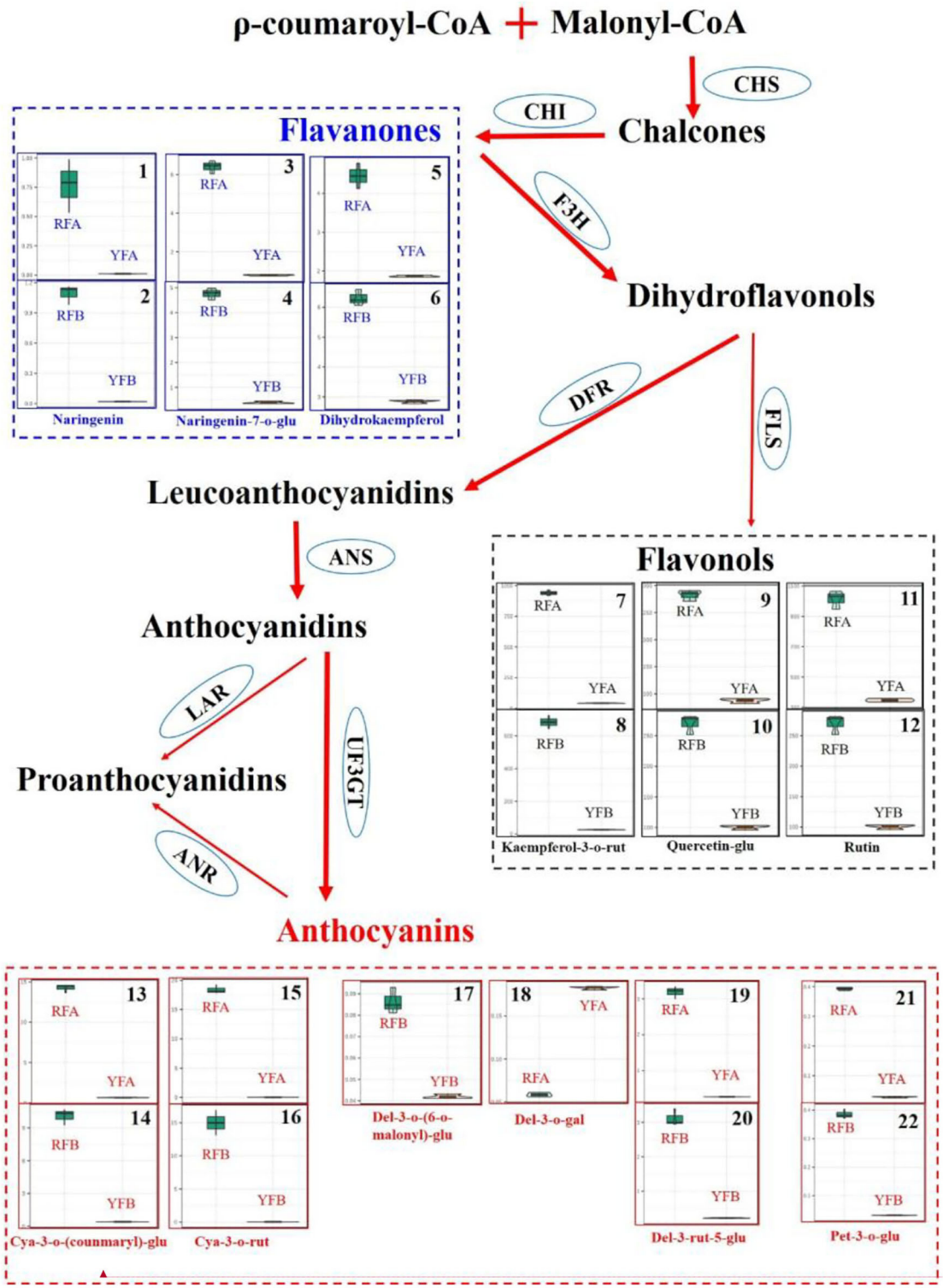


FIGURE 6 | Differentiate metabolisms of flavonoids synthesis between *Nicotiana tabacum* and *Nicotiana rustica*. Each square involving violin plots represents a metabolite. The metabolites were compared between cultivars RFA (No-2146) or RFB (Xiangyan-3) and YFA (Laolaihuang) or YFB (Xiaoyelanhua). The squares adhering without separation represent the same metabolites. The ellipses denote enzymes. CHS, Chalcone synthase; CHI, Chalcone isomerase; F3H, Flavanone 3-hydroxylase; FLS, Flavonol synthase; DFR, Dihydroflavonol reductase; ANS, Anthocyanidin synthase; UF3GT, Flavonoid 3-o-glucosyltransferase; LAR, Leucoanthocyanidin reductase; ANR, Anthocyanidin reductase.

of the red color of flowers (Nishihara et al., 2005). This study systematically compared three main classes of flavonoids, regarding flavanones, flavonols, and anthocyanins, in *Nicotiana* species. The red flowers contained higher these three classes of flavonoids, whereas the yellow flowers did not accumulate much on these three classes. The compositions of flavonoids posed crucial effects on the formation of flower color. In other plants, the compositions of flavonoids could be changed in the same species. For instance, the pink flowers contained anthocyanins, flavones, and flavonols in chrysanthemums, while the white flower had flavones and flavonols merely (Chen et al., 2012). In *Lycoris longituba*, less common anthocyanins were identified among the various-colored flowers (He et al., 2011). However, there were no apparent alternatives of flavonoids between cultivars belonging to the same *Nicotiana* species in our studies. The cultivars in the same species were probably with close genetic background. It was considered that anthocyanin belongs to the red series and determines the flower colors ranging from pink to blue. In contrast, the other flavonoids belong to the pure yellow series, where flavones, flavanones, and flavonols are light yellow and even close to colorless, and chalcones and aurones are deep yellows (Zhao and Tao, 2015). Our results did not strictly agree with this and inferred they were complex processes within the coloration determined by flavonoids. Jiao et al. (2020) discovered some anthocyanins' down-accumulation in white flower mutants of *N. tabacum*, compared with those in the red flowers of wild type, of which Cya-3-*o*-glu and Pel-3-*o*-glu in our research had an analogous accumulation in the red nor yellow flowers.

The Yellow Color of Flowers Appeared in Species *N. rustica* With Higher Carotenoid Accumulation

Carotenoids were more enriched in yellow flowers, in species *N. rustica*, including cultivars YFA (Laolaihuang) and YFB (Xiaoyelanhua). However, there was less carotenoid accumulation in the red flowers of species *N. tabacum* involving RFA (No-2146) and RFB (Xiangyan-3). In plants, carotenoids exist in photosynthetic and non-photosynthetic organs (Yuan et al., 2015). Mainly, carotenoids provide bright colors in non-photosynthetic tissues. The compositions of carotenoids are not the same among species, despite the flower representing the close color in the petals (Zhao and Tao, 2015). In the yellow flowers of species *N. rustica*, α-carotenes, luteins, β-carotenes, zeaxanthins, antheraxanthins, violaxanthins, and neoxanthins were much more enriched. In the carotenoid synthesis (shown in **Figure 7**), ζ-carotene is yellow, and lycopene is red. Lycopene cyclase produces carotenoids with cyclic terminal end groups such as α-carotene and β-carotene, which appeared with orange-yellow and yellow colors, respectively (Maoka, 2020). In *Osmanthus*, minor accumulation of α-carotene resulted in golden yellow, while significant enrichment of β-carotene produced butter yellow (Han et al., 2014). Regarding luteins and their epoxides, the higher accumulation leads the color ranging from yellow to orange in varieties such as *Marigold* and *Chrysanthemum* (Kishimoto and Ohmiya, 2006; Zhu et al., 2014). Minor accumulation of luteins showed yellow and orange flowers in species *Oncidium* and *Osmanthus* (Chiou et al., 2010; Han et al., 2014). Minor antheraxanthin accumulated in the flower of *Lili* var. Saija and Tiger showed red color (Yamagishi et al., 2010; Jekni et al., 2012). However, the flowers were yellow in *Lily* var. Connecticut King, when antheraxanthins were detected combing with violaxianthins and luteins simultaneously (Yamagishi et al., 2010). The plant flowers contain a single type of carotenoids such as antheraxanthins rarely, and thereby flowers enriched with carotenoids usually show yellow color. It is reported that only violaxanthins concentrated profoundly in *Oncidium* var. Gower Ramsey and subsequently presented yellow flowers. Additionally, the combing accumulation on violaxanthins and β-carotene generated orange flowers in the *Oncidium* var. Sunkist (Chiou et al., 2010). In general, carotenoids lead to yellow flowers in plants, especially those containing more than two classes. In *N. rustica*, the yellow-color of the flowers was probably due to the appearance of multiple types of carotenoids (shown in **Figure 8**).

Exploring Novel Metabolic Marker–Involved Metabolites Is an Effective Strategy to Identify Traits

The metabolome allowed exploration on metabolic biomarkers, besides dissecting the metabolic mechanisms underlying divergence between red and yellow flowers of *Nicotiana*. Some specifically distinct and ultra-higher accumulated metabolites were screened out in this study, such as kaempferol-3-*o*-rut, quercetin-glu, rutin, lutein, β-carotene (**Figures 2**, **3**). Generally, the biomarker should be employed in populations, and it should not be influenced by different environments in plant breeding (Steinfath et al., 2010). Indeed, typical marker-assisted selection has been proved to be very successful for many diploid crops. However, applying genetic markers is still problematic for species with a complex polyploid genome, which is prevailing in many crop plants. Another difficulty for genetic markers is that they fail to apply for polygenic traits directly. A batch of metabolites, relating to tobacco flower coloration were screened out for biomarkers in our study. Besides the traditional biomarker selection, these metabolic biomarkers must enrich the means to improve tobaccos.

CONCLUSIONS

The main findings were concluded as follows:

- A total of 31 flavonoids and 36 carotenoids were identified in all four cultivars involved in *N. tabacum* and *N. rustica*.
- Flavonoids and carotenoids tended to concentrate in the red flowers (*N. tabacum*) and yellow flowers (*N. rustica*), respectively.
- Eight flavonoids and 12 carotenoids were primarily screened out for metabolic biomarkers, such as the robust biomarker involving kaempferol-3-*o*-rut, quercetin-glu, rutin, lutein, and β-carotene.

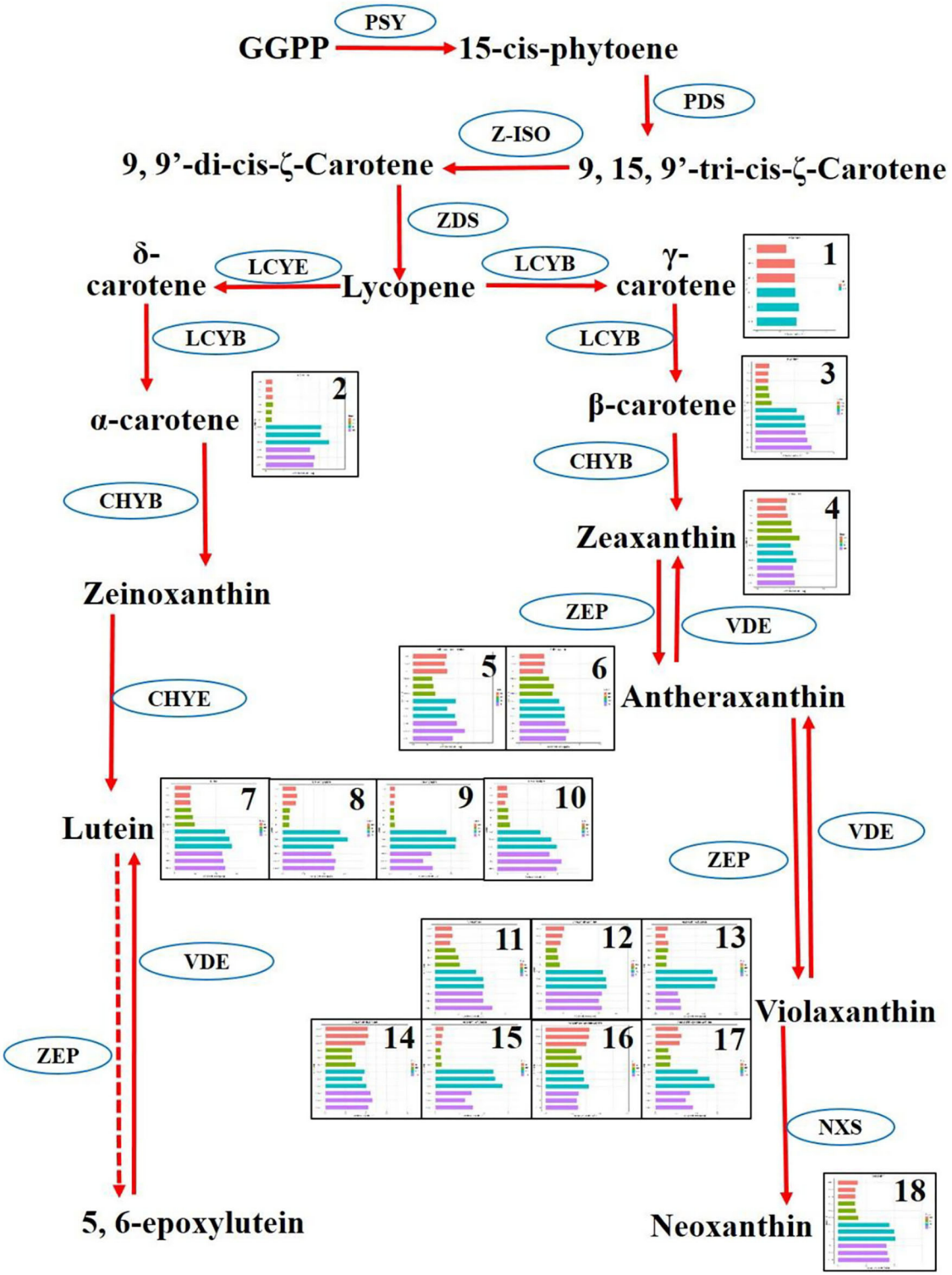


FIGURE 7 | Differentiate metabolisms of carotenoid synthesis between *Nicotiana tabacum* and *Nicotiana rustica*. Each square represented a metabolite. The replicates and four cultivars were exhibited in the squares. The red, yellow, cyan, and pink denote RFA (No-2146), RFB (Xiangyan-3), YFA (Laolaihuang), and YFB (Xiaoyelanhua), respectively. The squares were numbered and named as 1. γ-Carotene, 2. α-Carotene, 3. β-Carotene, 4. Zeaxanthin, 5. Antheraxanthin–dipalmitate, 6. Antheraxanthin, 7. Lutein, 8. Lutein–dimyristate, 9. Lutein–myristate, 10. Lutein—palmitate, 11. Violaxanthin, 12. Violaxanthin–palmitate, 13. Violaxanthin–dibutyrate, 14. Violaxanthin–dipalmitate, 15. Violaxanthin–myristate, 16. Violaxanthin-myristate-caprate, 17. Violaxanthin–myristate-palmitate, 18. Neoxanthin. The ellipses denote enzymes. PSY, Phytoene synthase; PDS, Phytoene desaturase; Z-ISO, ζ-carotene isomerase; ZDS, ζ-carotene desaturase; LCYE, Epsilon cyclase; LCYB, β-cyclase; ZEP, Zeaxanthin epoxidase; VDE, Violaxanthin de-epoxidase; NXS, Neoxanthin synthase.

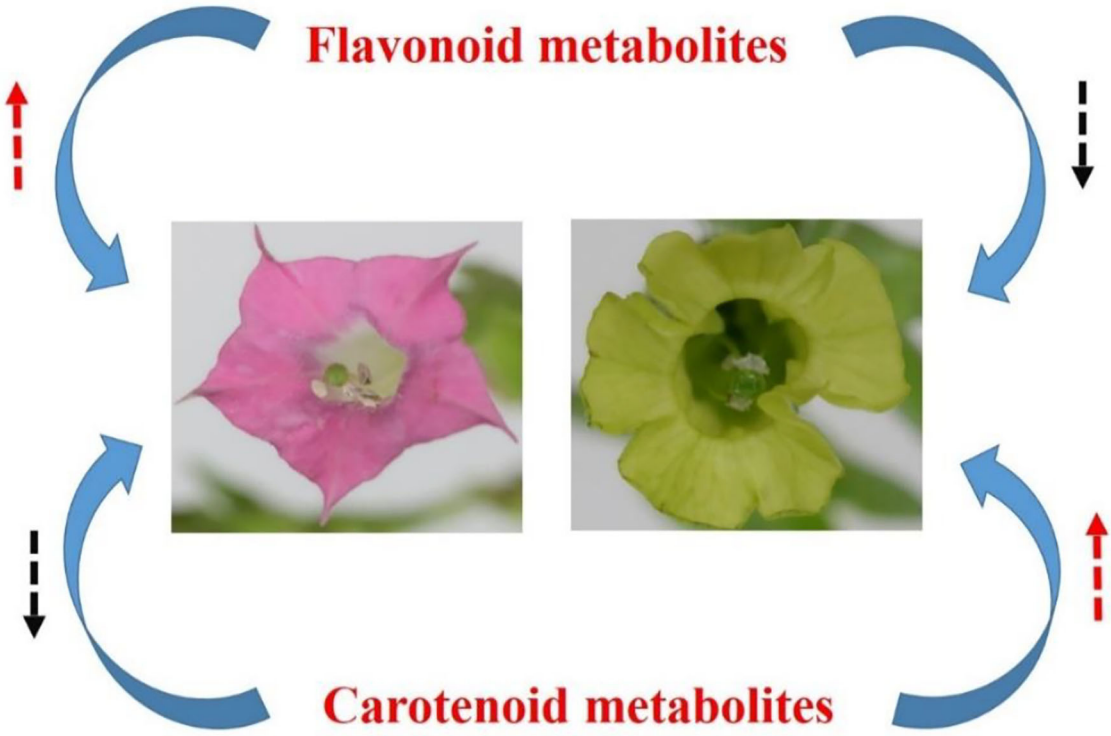


FIGURE 8 | Hypothetical metabolic model on the divergence of flower color between *N. tabacum* and *N. rustica*. The red and black dash arrows denote accumulation and decline.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**, further inquiries can be directed to the corresponding author.

AUTHOR CONTRIBUTIONS

QX and LM designed the study. QX and LL grew the plants. YZ, GC, RH, LW, and ZD performed the qualifying and quantifying on the metabolism involving anthocyanin and carotenoid. XZ, ZD, and LM collected and analyzed the data. RH, LL, and LW contributed to the illustration. All authors contributed to the manuscript writing and approved the submitted version.

FUNDING

This work was funded by the Project of Yongzhou Tobacco Monopoly Bureau with grant number *yzyc2019KJ02*, the China Postdoctoral Science Foundation with grant number *2021M690633*, and the Natural Science Foundation of Hubei Province, China, with grant number *2021CFB002*.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2022.828042/full#supplementary-material

Supplementary Figure 1-1 | Features of ion chromatogram on anthocyanins. **(A)** Extracted ion chromatogram, **(B)** overlapping spectra of TIC of QC samples in positive ion mode, and **(C)** normalized area of ion intensity. TIC, total ion chromatogram; QC, quality control.

Supplementary Figure 1-2 | Features of ion chromatogram on carotenoids. **(A)** Extracted ion chromatogram, **(B)** overlapping spectra of TIC of QC samples in positive ion mode, and **(C)** normalized area of ion intensity. TIC, total ion chromatogram; QC, quality control.

Supplementary Figure 2 | Principal component analysis on 12 samples based on all metabolites.

Supplementary Table 1 | Equation of standard curve involved in anthocyanins, where *r* represents the coefficient of correlation.

Supplementary Table 2 | Equation of standard curve involved in carotenoids, where *r* represents the coefficient of correlation.

Supplementary Table 3 | Differential metabolites among cultivars and involved Venn diagrams.

REFERENCES

Afendi, F. M., Okada, T., Yamazaki, M., Hirai-Morita, A., Nakamura, Y., Nakamura, K., et al. (2012). KNApSAcK family databases: integrated metabolite-plant species databases for multifaceted plant research. *Plant Cell Physiol.* 53, e1. doi: 10.1093/pcp/pcr165

Al-Babili, S., and Bouwmeester, H. J. (2015). Strigolactones, a novel carotenoid-derived plant hormone. *Annu. Rev. Plant Biol.* 66, 161–186. doi: 10.1146/annurev-arplant-043014-114759

Avalos, J., and Carmen Limón, M. (2015). Biological roles of fungal carotenoids. *Curr. Genet.* 61, 309–324. doi: 10.1007/s00294-014-0454-x

Bovet, L., Rossi, L., and Lugon-Moulin, N. (2006). Cadmium partitioning and gene expression studies in *Nicotiana tabacum* and *Nicotiana rustica*. *Physiologia Plantarum.* 128, 466–475. doi: 10.1111/j.1399-3054.2006.00756.x

Chen, S. M., Li, C. H., Zhu, X. R., Deng, Y. M., Sun, W., Wang, L. S., et al. (2012). The identification of flavonoids and the expression of genes of anthocyanin biosynthesis in the chrysanthemum flowers. *Biologia Plantarum.* 56, 458–464. doi: 10.1007/s10535-012-0069-3

Chiou, C. Y., Pan, H. A., Chuang, Y. N., and Yeh, K. W. (2010). Differential expression of carotenoid-related genes determines diversified carotenoid coloration in floral tissues of oncidium cultivars. *Planta.* 232, 937–948. doi: 10.1007/s00425-010-1222-x

Collard, B. C., and Mackill, D. J. (2008). Marker-assisted selection: an approach for precision plant breeding in the twenty-first century. *Philos Trans R Soc Lond B Biol Sci.* 363, 557–572. doi: 10.1098/rstb.2007.2170

Domonkos, I., Kis, M., Gombos, Z., and Ughy, B. (2013). Carotenoids, versatile components of oxygenic photosynthesis. *Prog Lipid Res.* 52, 539–561. doi: 10.1016/j.plipres.2013.07.001

Forkmann, G., and Martens, S. (2001). *Metabolic Engineering and Applications of Flavonoids*. England: Elsevier Ltd.

Fuentes, P., Pizarro, L., Moreno, J. C., Handford, M., Rodriguez-Concepcion, M., Stange, C., et al. (2012). Light-dependent changes in plastid differentiation influence carotenoid gene expression and accumulation in carrot roots. *Plant Mol Biol.* 79, 47–59. doi: 10.1007/s11103-012-9893-2

Geifman-Holtzman, O., and Ober, B. J. (2008). Prenatal diagnosis: update on invasive versus noninvasive fetal diagnostic testing from maternal blood. *Expert Rev Mol Diagn.* 8, 727–751. doi: 10.1586/14737159.8.6.727

Goodspeed, T. H. (1954). The genus nicotiana; Origins, relationships and evolution of its species in the light of their distribution, morphology and cytogenetics. Waltham, MA: Chronica Botanica Co.

Han, Y., Wang, X., Chen, W., Dong, M., Yuan, W., Liu, X., et al. (2014). Differential expression of carotenoid-related genes determines diversified carotenoid coloration in flower petal of osmanthus fragrans. *Tree Genet. Genomes.* 10, 329–338. doi: 10.1007/s11295-013-0687-8

Harbone, J. B. (1993). Phytochemistry, acad. *Press London.* 21, 2785.

Hashimoto, H., Uragami, C., and Cogdell, R. J. (2016). Carotenoids and photosynthesis. *Subcell Biochem.* 79, 111–139. doi: 10.1007/978-3-319-39126-7_4

Havaux, M. (2014). Carotenoid oxidation products as stress signals in plants. *Plant J.* 79, 597–606. doi: 10.1111/tpj.12386

He, Q., Shen, Y., Wang, M., Huang, M., Yang, R., and Zhu, S., et al. (2011). Natural variation in petal color in lycoris longituba revealed by anthocyanin components. *Plos ONE.* 6, e22098. doi: 10.1371/journal.pone.0022098

Hou, X., Rivers, J., León, P., McQuinn, R. P., and Pogson, B. J. (2016). Synthesis and function of apocarotenoid signals in plants. *Trends Plant Sci.* 21, 792–803. doi: 10.1016/j.tplants.2016.06.001

Huang, W., Zhang, D., Cao, Y., Dang, B., Jia, W., Xu, Z., et al. (2021). Differential cadmium translocation and accumulation between *Nicotiana tabacum* L. and *Nicotiana rustica* L. by transcriptome combined with chemical form analyses. *Ecotoxicol. Environm. Safety.* 208, 111412. doi: 10.1016/j.ecoenv.2020.111412

Jekni,ć, Z., Morr,é, J. T., Jekni,ć, S., Jevremovi,ć, S., Suboti,ć, A., Chen, T. H. H., et al. (2012). Cloning and functional characterization of a gene for capsanthin-capsorubin synthase from tiger lily (lilium lancifolium thunb. 'splendens'). *Plant Cell Physiol.* 53, 1899–1912. doi: 10.1093/pcp/pcs128

Jiao, F., Zhao, L., Wu, X., Song, Z., and Li, Y. (2020). Metabolome and transcriptome analyses of the molecular mechanisms of flower color mutation in tobacco. *BMC Genomics.* 21. doi: 10.1186/s12864-020-07028-5

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

Khamis, M. M., Adamko, D. J., and El-Aneed, A. (2017). Mass spectrometric based approaches in urine metabolomics and biomarker discovery. *Mass Spectrometry Rev.* 36, 115–134. doi: 10.1002/mas.21455

Kishimoto, S., and Ohmiya, A. (2006). Regulation of carotenoid biosynthesis in petals and leaves of chrysanthemum (chrysanthemum morifolium). *Physiologia Plantarum.* 128, 436–447. doi: 10.1111/j.1399-3054.2006.00761.x

Li, P., Chen, X., Sun, F., and Dong, H. (2017). Tobacco ttg2 and arf8 function concomitantly to control flower colouring by regulating anthocyanin synthesis genes. *Plant Biol (Stuttg).* 19, 525–532. doi: 10.1111/plb.12560

Liu, Y., Lv, J., Liu, Z., Wang, J., Yang, B., Chen, W., et al. (2020). Integrative analysis of metabolome and transcriptome reveals the mechanism of color formation in pepper fruit (*Capsicum annuum* l.). *Food Chemist.* 306, 125629. doi: 10.1016/j.foodchem.2019.125629

Maoka, T. (2020). Carotenoids as natural functional pigments. *J. Natural Med.* 74, 1–16. doi: 10.1007/s11418-019-01364-x

Mekapogu, M., Vasamsetti, B. M. K., Kwon, O., Ahn, M., Lim, S., and Jung, J. (2020). Anthocyanins in floral colors: biosynthesis and regulation in chrysanthemum flowers. *Int. J. Mol. Sci.* 21, 6537. doi: 10.3390/ijms21186537

Meyer, R. C., Steinfath, M., Lisec, J., Becher, M., Witucka-Wall, H., Törjék, O., et al. (2007). The metabolic signature related to high plant growth rate in arabidopsis thaliana. *Proc. Natl. Acad. Sci. U S A.* 104, 4759–4764. doi: 10.1073/pnas.0609709104

Mol, J., Cornish, E., Mason, J., and Koes, R. (1999). Novel coloured flowers. *Curr. Opini. Biotechnol.* 10, 198–201. doi: 10.1016/S0958-1669(99)80035-4

Muir, S. R., Collins, G. J., Robinson, S., Hughes, S., Bovy, A., Ric-De-Vos, C. H., et al. (2001). Overexpression of petunia chalcone isomerase in tomato results in fruit containing increased levels of flavonols. *Nat. Biotechnol.* 19, 470–474. doi: 10.1038/88150

Nambara, E., and Marion-Poll, A. (2005). Abscisic acid biosynthesis and catabolism. *Annu. Rev. Plant Biol.* 56, 165–185. doi: 10.1146/annurev.arplant.56.032604.144046

Nishihara, M., Nakatsuka, T., and Yamamura, S. (2005). Flavonoid components and flower color change in transgenic tobacco plants by suppression of chalcone isomerase gene. *Febs Lett.* 579, 6074–6078. doi: 10.1016/j.febslet.2005.09.073

Niyogi, K. K., and Truong, T. B. (2013). Evolution of flexible non-photochemical quenching mechanisms that regulate light harvesting in oxygenic photosynthesis. *Curr. Opin. Plant Biol.* 16, 307–314. doi: 10.1016/j.pbi.2013.03.011

Park, S., Kim, D., Yang, J., Lee, J., and Lim, S. (2020). Increased flavonol levels in tobacco expressing acfls affect flower color and root growth. *Int. J. Mol. Sci.* 21, 1011. doi: 10.3390/ijms21031011

Perez, D. S. L., Garbowicz, K., Brotman, Y., Tohge, T., and Fernie, A. R. (2020). The acetate pathway supports flavonoid and lipid biosynthesis in arabidopsis. *Plant Physiol.* 182, 857–869. doi: 10.1104/pp.19.00683

Quian-Ulloa, R., and Stange, C. (2021). Carotenoid biosynthesis and plastid development in plants: the role of light. *Int. J. Mol. Sci.* 22, 1184. doi: 10.3390/ijms22031184

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

Schmitzer, V. V. R. O. (2010). Color and phenolic content changes during flower development in groundcover rose. *J. Am. Soc. Hortic. Sci.* 3, 195–202.

Siva Raju, K., Sheshumadhav, M., and Murthy, T. G. K. (2008). Molecular diversity in the genus nicotiana as revealed by randomly amplified polymorphic DNA. *Physiol. Mol. Biol. Plants.* 14, 377–382. doi: 10.1007/s12298-008-0037-8

Steinfath, M., Strehmel, N., Peters, R., Schauer, N., Groth, D., Hummel, J., et al. (2010). Discovering plant metabolic biomarkers for phenotype prediction using an untargeted approach. *Plant Biotechnol. J.* 8, 900–911. doi: 10.1111/j.1467-7652.2010.00516.x

Sun, T., Yuan, H., Cao, H., Yazdani, M., Tadmor, Y., Li, L. (2018). Carotenoid metabolism in plants: the role of plastids. *Mol Plant* 11, 58–74. doi: 10.1016/j.molp.2017.09.010

Tatsuzawa, F., Saito, N., Toki, K., Shinoda, K., Honda, T. (2012). Flower colors and their anthocyanins in matthiola incana cultivars (brassicaceae). *J. Japanese Society Hort. Sci.* 81, 91–100. doi: 10.2503/jjshs1.81.91

Thevenot, E. A., Roux, A., Xu, Y., Ezan, E., Junot, C. (2015). Analysis of the human adult urinary metabolome variations with age, body mass index, and gender by implementing a comprehensive workflow for univariate and OPLS statistical analyses. *J. Proteome. Res.* 14, 3322–3335. doi: 10.1021/acs.jproteome.5b00354

Thill, J., Miosic, S., Ahmed, R., Schlangen, K., Muster, G., Stich, K., et al. (2012). 'le rouge et le noir': a decline in flavone formation correlates with the rare color of black dahlia (dahlia variabilis hort.) flowers. *BMC Plant Biol.* 12, 225. doi: 10.1186/1471-2229-12-225

Tian, L. (2015). Recent advances in understanding carotenoid-derived signaling molecules in regulating plant growth and development. *Front. Plant Sci.* 6, 790. doi: 10.3389/fpls.2015.00790

Wen, W., Alseekh, S., and Fernie, A. R. (2020). Conservation and diversification of flavonoid metabolism in the plant kingdom. *Curr. Opin. Plant Biol.* 55, 100–108. doi: 10.1016/j.pbi.2020.04.004

Yabuzaki, J. (2017). Carotenoids database: structures, chemical fingerprints and distribution among organisms. *Database.* 2017. doi: 10.1093/database/bax004

Yamagishi, M., Kishimoto, S., and Nakayama, M. (2010). Carotenoid composition and changes in expression of carotenoid biosynthetic genes in tepals of asiatic hybrid lily. *Plant Breed.* 129, 100–107. doi: 10.1111/j.1439-0523.2009.01656.x

Yonekura-Sakakibara, K., Higashi, Y., and Nakabayashi, R. (2019). The origin and evolution of plant flavonoid metabolism. *Front. Plant Sci.* 10. doi: 10.3389/fpls.2019.00943

Yuan, H., Zhang, J., Nageswaran, D., and Li, L. (2015). Carotenoid metabolism and regulation in horticultural crops. *Hort. Res.* 2. doi: 10.1038/hortres.2015.36

Zhao, D., and Tao, J. (2015). Recent advances on the development and regulation of flower color in ornamental plants. *Front. Plant Sci.* 6. doi: 10.3389/fpls.2015.00261

Zhao, D., Tao, J., Han, C., and Ge, J. (2012). Flower color diversity revealed by differential expression of flavonoid biosynthetic genes and flavonoid accumulation in herbaceous peony (paeonia lactiflora pall.). *Mol. Biol. Rep.* 39, 11263–11275. doi: 10.1007/s11033-012-2036-7

Zhu, C., Yang, Q., Ni, X., Bai, C., Sheng, Y., and Shi, L., et al. (2014). Cloning and functional analysis of the promoters that upregulate carotenogenic gene expression during flower development in gentiana lutea. *Physiol Plant.* 150, 493–504. doi: 10.1111/ppl.12129

Conflict of Interest: QX, RH, and ZD were employed by Yongzhou Tobacco Monopoly Bureau of Hunan.

The remaining 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.

frontiers | Frontiers in Plant Science

ORIGINAL RESEARCH
published: 13 May 2022
doi: 10.3389/fpls.2022.814677

Metabolism of Carotenoids and β-Ionone Are Mediated by Carotenogenic Genes and *PpCCD4* Under Ultraviolet B Irradiation and During Fruit Ripening

Hongru Liu[1,2,3], Xiangmei Cao[3], Muhammad Azam[4], Chunfang Wang[1,2], Chenxia Liu[1,2], Yongjin Qiao[1,2] and Bo Zhang[3]**

[1] Crop Breeding & Cultivation Research Institute, Shanghai Academy of Agricultural Sciences, Shanghai, China, [2] Research Center for Agricultural Products Preservation and Processing, Shanghai Academy of Agricultural Sciences, Shanghai, China, [3] Laboratory of Fruit Quality Biology Zhejiang Provincial Key Laboratory of Horticultural Plant Integrative Biology, Zhejiang University, Hangzhou, China, [4] Pomology Laboratory, Institute of Horticultural Sciences, University of Agriculture Faisalabad, Faisalabad, Pakistan

OPEN ACCESS

Edited by:
Xiumin Fu,
South China Botanical Garden (CAS), China

Reviewed by:
Ilahy Riadh,
Institut National de la Recherche Agronomique de Tunisie (INRAT), Tunisia
Hongli Cui,
Shanxi Agricultural University, China
Xiongjie Zheng,
King Abdullah University of Science and Technology, Saudi Arabia

***Correspondence:**
Yongjin Qiao
yjqiao2002@126.com
Bo Zhang
bozhang@zju.edu.cn

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

Received: 14 November 2021
Accepted: 07 April 2022
Published: 13 May 2022

Citation:
Liu H, Cao X, Azam M, Wang C, Liu C, Qiao Y and Zhang B (2022) Metabolism of Carotenoids and β-Ionone Are Mediated by Carotenogenic Genes and PpCCD4 Under Ultraviolet B Irradiation and During Fruit Ripening. Front. Plant Sci. 13:814677. doi: 10.3389/fpls.2022.814677

Carotenoids are essential pigments widely distributed in tissues and organs of higher plants, contributing to color, photosynthesis, photoprotection, nutrition, and flavor in plants. White- or yellow-fleshed colors in peach were determined by expression of carotenoids cleavage dioxygenase (*PpCCD*) genes, catalyzing the degradation of carotenoids. The cracked volatile apocarotenoids are the main contributors to peach aroma and flavor with low sensory threshold concentration. However, the detailed regulatory roles of carotenoids metabolism genes remained unclear under UV-B irradiation. In our study, metabolic balance between carotenoids and apocarotenoids was regulated by the expression of phytoene synthase (*PSY*), β-cyclase (*LCY-B*), ε-cyclase (*LCY-E*), and *PpCCD4* under UV-B irradiation. The transcript levels of *PpPSY*, *PpLCY-B*, *PpLCY-E*, and *PpCHY-B* were elevated 2- to 10-fold compared with control, corresponding to a nearly 30% increase of carotenoids content after 6 h UV-B irradiation. Interestingly, the total carotenoids content decreased by nearly 60% after 48 h of storage, while UV-B delayed the decline of lutein and β-carotene. The transcript level of *PpLCY-E* increased 17.83-fold compared to control, partially slowing the decline rate of lutein under UV-B irradiation. In addition, the transcript level of *PpCCD4* decreased to 30% of control after 48 h UV-B irradiation, in accordance with the dramatic reduction of apocarotenoid volatiles and the delayed decrease of β-carotene. Besides, β-ionone content was elevated by ethylene treatment, and accumulation dramatically accelerated at full ripeness. Taken together, UV-B radiation mediated the metabolic balance of carotenoid biosynthesis and catabolism by controlling the transcript levels of *PpPSY*, *PpLCY-B*, *PpLCY-E*, and *PpCCD4* in peach, and the transcript level of *PpCCD4* showed a positive relationship with the accumulation of β-ionone during the ripening process. However, the detailed catalytic activity of PpCCD4 with various carotenoid substrates needs to be studied further, and the key transcript factors involved in the regulation of metabolism between carotenoids and apocarotenoids need to be clarified.

Keywords: carotenoids, β-ionone, metabolism, peach, UV-B irradiation

INTRODUCTION

Carotenoids are nutritional lipid-soluble tetraterpenoid pigments, which are widely distributed in a number of fruits, flowers, and vegetables, contributing to their colors, aromas, flavors, attractants for pollinators, and seed-spreaders (Nisar et al., 2015; Hou et al., 2016). Those compounds play important roles in light harvesting of photosynthesis system and photoprotection by scavenging free radicals and quenching single oxygen derived from excess light damage (Bouvier et al., 2005). In addition, carotenoids can be oxidatively cleaved by CCDs, producing norisoprenoid aroma volatiles, which are important contributors to floral aroma in fruits and flowers (Nisar et al., 2015). Their derivatives have been involved in a number of physiological processes in plants, including development and ripening of fruits, rhythmicity of flower volatile emission, metabolic of phytohormones, and small signaling molecules (Bouvier et al., 2005; Cazzonelli and Pogson, 2010; Nisar et al., 2015). Meanwhile, carotenoids have a beneficial effect on human health, being precursors of vitamin A and partial constituents, and have been used to reduce the risk of age-related macular degeneration and certain cancers (Kopsell and Kopsell, 2006; Yuan et al., 2015; Hou et al., 2016).

Carotenoids are composed of 40-carbon hydrocarbons derived from the building blocks of isopentenyl diphosphate (IPP) and dimethylallyl diphosphate (DMAPP). The basic five-carbon intermediates are produced by methylerythritol 4-phosphate (MEP) and mevalonic acid (MVA) pathways in the plastid and cytoplasm, respectively, in plants (Farré et al., 2010). Then geranylgeranyl pyrophosphate (GGPP), a common substrate of carotenoids and diterpenes, is condensed with three IPPs and one DMAPP molecule by geranylgeranyl pyrophosphate synthase (GGPPS) in a plastid. Then rate-limiting enzyme PSY converts two GGPP molecules to the first carotenoid phytoene, guiding carbon flux to the biosynthesis of carotenoids (Cazzonelli and Pogson, 2010; Fu et al., 2014). Phytoene is then transformed to lycopene by continuous desaturation and isomerization with enzyme catalyzation (Nisar et al., 2015). Subsequently, lycopene is cyclized with bifurcate pathway, either by alone *LCY-B*/chromoplast-specific lycopene β-cyclase (CYCB) to form β-carotene or by *LCY-B* and *LCY-E* to synthesize α-carotene (Alquezar et al., 2009). Moreover, α-carotene and β-carotene will be converted to lutein and zeaxanthin, respectively, *via* a hydroxylation reaction. Furthermore, the xanthophyll cycle between zeaxanthin and violaxanthin has been known to respond to the excess light stress (De Poll et al., 2011). Epoxidation of zeaxanthin is catalyzed by zeaxanthin epoxidase (ZEP) and reversed by violaxanthin de-epoxidase (VDE) (Chen and Gallie, 2012). Violaxanthin can be converted into neoxanthin by neoxanthin synthase (NSY), and 9-*cis*-violaxanthin can be cleaved at 11, 12 site and/or 9′-*cis*-neoxanthin at 11′, and 12′ double bond by 9-*cis*-epoxycarotenoid dioxygenase (NCED) to produce the precursor of abscisic acid (ABA), xanthoxin (Walter et al., 2010; Nisar et al., 2015). The catabolic of ABA is mainly catalyzed by the key enzyme of ABA8′-hydroxylase (Vallabhaneni and Wurtzel, 2010). Therefore, several rate-limiting enzymes determined the carbon flux of carotenoids metabolism, and the relationship between key enzymes and carotenoids constituencies has been intensively studied in plants (Nisar et al., 2015).

Carotenoids content are determined by the balance of biosynthesis and catabolism; unlike biosynthesis, which is determined by several key enzymes, the catabolism is mainly determined by CCD enzymes (Ibdah et al., 2006; Huang et al., 2009; Havaux, 2014; Hou et al., 2016). Various products have been produced by the oxidization of carotenoids with different cleaved sites, such as norisoprenoid volatiles, phytohormones (ABA, SLs), and some other small molecular metabolites (Havaux, 2014; Hou et al., 2016). CCD gene families have commonly been divided into CCD and NCED subclasses. Generally, CCDs consist of four members (1, 4, 7, 8), and NCED contains five members (2, 3, 5, 6, 9) in *Arabidopsis* (Gonzalez-Jorge et al., 2013; Song et al., 2016). The five NCEDs were involved in the metabolism of ABA, and VP14 from maize was the first identified member of the family (Tan et al., 2003; Messing et al., 2010). *AtCCD7/MAX3* and *AtCCD8/MAX4* members are continuously involved in the biosynthesis of strigolactone in *Arabidopsis* (Booker et al., 2004; Walter et al., 2010). Meanwhile, various norisoprenoid volatiles have been produced by CCD1 enzyme in tomato, grape, petunia, rose damascena, and *Osmanthus fragrans Lour* (Simkin et al., 2004a,b; Mathieu et al., 2005; Huang et al., 2009; Baldermann et al., 2010).

In addition, the members of CCD4 have been reported to degrade carotenoids, resulting in formation of white color (Brandi et al., 2011; Yoshioka et al., 2012; Falchi et al., 2013; Ma et al., 2013a). White-fleshed potato tubers and peach fruit showed higher expression levels of *CCD4* than the yellow-fleshed mutants, and accumulation of carotenoids was decreased in white-fleshed varieties with C_{-13} norisoprenoid volatiles content was increased (Campbell et al., 2010; Brandi et al., 2011; Falchi et al., 2013; Ma et al., 2013b; Bai et al., 2016). Besides, many studies have mentioned that protein CCD4 catalyzed carotenoids into apocarotenoid volatiles among many plants' tissues and organs. Two types of *CsCCD4* genes with different subcellular localization in crocus are involved in both the degradation of β-carotene and the release of β-ionone in stigma (Rubio-Moraga et al., 2014). BnaC3.CCD4 might use δ-and/or α-carotene as substrates to produce α-ionone and determine the white- or yellow-color of petals from *Brassica napus* (Zhang et al., 2015). The β-ionone was one of the most important flavor compounds in green and black tea, and CsCCD4 from *Camellia sinensis* could cleave β-carotene at 9,10 (9′,10′) sites to produce β-ionone during the withering process (Wang et al., 2020). Besides, β-cyclocitral, β-ionone, β-damascenone, and β-citraurin in plants, algae, fungi, and bacteria could be produced by enzymatic cleavage of CCDs, peroxidases (PODs) and lipoxygenases (LOXs) and even with non-enzymatic oxidation with various carotenoids substrates (Ma et al., 2013a; Bruno et al., 2016; Liang et al., 2021; Zheng et al., 2021).

Carotenoids and their derivatives have exhibited key roles in diverse metabolic processes during plant development, fruit ripening, and interaction with various environmental stimuli (Bouvier et al., 2005; Walter and Strack, 2011). Carotenoid

metabolites and xanthophyll cycles have been involved in the response to microclimate light and UV-B attenuation in the field during grapeberry development (Joubert et al., 2016). The cleaved metabolites, such as β-ionone and β-cyclocitral, have been reported to be important signaling molecules involved in interaction between plants and insects (Booker et al., 2004; Wei et al., 2011; Havaux, 2014; Hou et al., 2016). Plants exposed to UV-B (280–315 nm) have been widely reported to induce the accumulation of phenolic compounds in plants for their antioxidant properties (De Poll et al., 2011; Gil et al., 2012; Zeng et al., 2020). Meanwhile, carotenoids and their derivatives have also been involved in protecting against UV radiation as signal molecules, though the concentration of those compounds is usually low in plant tissues (Ramel et al., 2013; Havaux, 2014).

Previous studies have reported the metabolic mechanisms of carotenoids and apocarotenoids and identified many genes and enzymes to illustrate the molecular mechanism (Liang et al., 2021). However, the detailed regulatory roles during the development, fruit ripening, and response to environment stress were remained unclear, especially for horticultural crops (Shi et al., 2020; Koschmieder et al., 2021). Our study reported the metabolic of carotenoids and apocarotenoids were regulated by transcription changes of *PpPSY*, *PpLCY-B*, *PpLCY-E*, and *PpCCD4* under UV-B irradiation in peach. Meanwhile, the correlation between *PpCCD4* and C_{13}-norisoprenoid volatile β-ionone was analyzed during peach ripening. We clarified the key gene roles in carotenoids and apocarotenoids metabolic molecular mechanisms in peach, while the detailed enzyme activities of PpCCD4 with various carotenoid substrates, such as β-carotene, lutein, violaxanthin, and zeaxanthin, need to be detected further. The identification of transcription factors and illumination of transcriptional regulation mechanisms under the development, ripening, and interaction with environmental stimuli need to be strengthened.

MATERIALS AND METHODS

Plant Materials and Treatments

Peach (*Prunus persica* L. Batsch cv. Hujingmilu) was obtained from a well-managed orchard in Fengxian district of Shanghai, China. The fruits were harvested at three ripening stages, 110 days after bloom (110 DAB), 113 DAB, 116 DAB, and stored at 20°C for 3 days after harvest. The fruits were transported to the laboratory immediately after harvest from the orchard. Healthy fruits with uniform size, no mechanical damage, pathogen-, and disease-free were selected for our experiments. Each treatment consists of three biological replicates with five fruits each. Peel was separated from flesh, frozen in liquid nitrogen, and stored at −80°C for further analysis. UV-B irradiation treatments were processed as our previous work (Liu et al., 2017), fruits with uniform maturity were selected and irradiated with a high intensity of 1.5 W m^{-2} and a low intensity of 0.4 W m^{-2} for different times at 20°C. Ethylene and 1-methylcycloproene (1-MCP) treatments were performed as in previous work (Wu et al., 2017, 2019). Briefly, the marketable mature fruits were selected to be used, and the fruits were sealed in a bucket with a fan embedded in the lid. The fruits were suffocated for 12 h with an ethylene concentration of 1 μl L^{-1}, produced from the ethephon, and the competitive inhibitor 1-MCP treatment was performed with a concentration of 5 μl L^{-1}, and the control fruits were sealed only with the same buckets.

Determination of the Fruit Ripening Quality Index

The firmness of fruit was analyzed according to Zhang et al. (2010) using a texture analyzer (TA-XT2i Plus, Stable Micro system) fitted with a 7.9-mm diameter head. The penetration rate was 1 mm s^{-1} with a final depth of 10 mm, and the two opposite sides at the equator of the fruit were used to measure. Total soluble solids (TSS) were determined by an Atago PR-101α digital hang-held refractometer, as performed by Zhang et al. (2010). Briefly, two opposite slices at the equator of each fruit were used to evaluate TSS with three drops of juice per slice. Fruit color was measured with a reflectance spectrophotometer (Hunter Lab Mini Scan XE Plus colorimeter) at eight evenly distributed equatorial surface points of each fruit. The raw data values of L* (lightness, from black to white), a* (the degree of green-red variation), and b* (the variation between blue and yellow) were recorded, and the color index of red grapes (CIRG) was calculated using the formula CIRG = (180-H)/(L* + C), and H = arctan (b*/a*) and $C = [(a^*)^2 + (b^*)^2]^{0.5}$.

Aroma Volatiles Extraction and Determination by Head Space Gas Chromatography Mass Spectrometry

Aroma volatiles of peach peel were extracted and analyzed by head space gas chromatography mass spectrometry (HS-GC-MS) as our previous study (Wei et al., 2021). Frozen peel was ground into powder under liquid nitrogen, and 1 g of powder was immediately transferred into vials. To adjust the instrument, 30 μl of 2-Octanol (0.8 mg/ml) was added into the vials as internal standard. A solid-phase micro-extraction (SPME) needle containing a fiber coated with 65 μm of polydimethylsiloxane and divinylbenzene (PDMS-DVB) (Supeclo Co., Bellefonte, PA, United States) was used for volatile concentration. The determination was automatic injection with a SPME autosampler (Combi PAL, CTC Analytics, Agilent Technologies), connected to a 7890N gas chromatograph and a 5975C mass spectrometer from Agilent company. Volatiles were isolated with a DB-WAX capillary column (30 m × 0.25 mm i.d. × 0.25 μm film thickness; J&W Scientific, Folsom, CA, United States). Helium was used as a carrier gas with a 1.0 ml/min flow rate. The temperature program in oven started at 40°C and was increased to 100°C with a rate of 3°C/min and then to 245°C with 5°C/min. The column effluent was ionized by electron ionization (EI) at an energy of 70 eV with a transfer temperature of 250°C and a source temperature of 230°C. Volatiles were identified by comparing their electron ionization mass spectra with the NIST Mass Spectral Library (NIST-08) and the retention time of authentic standards. Quantification of volatiles was performed using the

peak area of the internal standard as a reference based on total ion chromatogram.

Isolation and Determination of Carotenoids by High Performance Liquid Chromatography (HPLC) Analysis

The extraction and determination of the carotenoids in peach peel was performed as previous reports (Zhou et al., 2007; Fu et al., 2014). Briefly, frozen peel was grounded finely under liquid nitrogen, and 100 mg of powder was used to extract with chloroform/methanol/50 mM Tris-buffer, pH 7.5. The chloroform phase was collected by 10 min of centrifugation at 10,000 × *g*, and the aqueous phases were re-extracted with chloroform. The two combined chloroform phases were finally dried under nitrogen gas. The residue was first dissolved in 20 μl of diethyl ether and then in 200 μl of 6% (W/V) KOH in methanol. After mixing fully, the tube was transferred into a heating unit for incubation at 60°C for 30 min in darkness. Tris-buffer was then added into the incubated mixture, and re-extraction was carried out as described above, and finally dried with nitrogen gas. Three replicates were used for the determination of each sample. HPLC analysis was performed using Waters Alliance 2695 system (Waters Corporation, United States) consisting of a 2695 module and a 2996 PDA detector, equipped with a reverse phase 5 μm (250 mm × 4.6 mm) C_{30} column and a 20 mm × 4.6 mm i.d. C_{30} guard (YMC Inc., Wilmington, NC, United States). Chromatography was carried out at 25°C with the elution program as previously described by Zhou et al. (2007). Carotenoids were identified on the basis of the same retention times and same spectral characteristics as standards, and peak areas were recorded at 450 nm. The quantification was performed according to the calibration curve of the corresponding standards.

Transcript Level Analysis

Total RNA extraction protocol was performed as our previous study (Liu et al., 2017), and the high-quality RNA was used to prepare the libraries for high-throughput RNA-seq as the protocols described previously (Zhang et al., 2016). RNA purity and concentration were detected with NanoDrop 2000 (Thermo Fisher Scientific, Wilmington, DE, United States). RNA integrity was measured using the RNA Nano 6000 Assay kit on the Agilent Bioanalyzer 2100 system (Agilent Technologies, CA, United States). A total of 1 μg of RNA from five fruits per sample was used for the RNA sample preparations. Sequencing libraries were made with the NEBNext Ultra™ RNA Library Prep Kit for Illumina (NEB, United States) according to the manufacturer's protocol. The RNA-seq raw data presented in the study can be found in the National Center for Biotechnology Information (NCBI) Short Red Archive database. The accession numbers PRJNA576753 and PRJNA57477 are for peach samples at different ripening processes and under ethylene and 1-MCP treatments, respectively, and the data have been published in a previous study (Cao et al., 2021), and SRP103523 represents the UV-B treatment, and the data have been published in a previous study (Liu et al., 2017).

Statistics

Figures were produced using ORIGIN 8.0 (Microcal Software, Inc., Northampton, MA, United States). The two-sample significance test was calculated using unpaired Student's *t*-test, and Duncan's test of one-way analysis of variance (ANOVA) was used to test the significance level of multiple groups ($p < 0.05$) (SPSS 21.0; SPSS Inc., Chicago, IL, United States).

RESULTS

Accumulation of Norisoprenoid Volatiles and Transcription of *PpCCD4* During Peach Ripening

The peach ripening process is accelerated after the S1 color transition stage with the fruit quality index changing markedly. Fruit firmness decreased quickly from 44 N to 25 N in 1 week, and then decreased to 5 N within 3 days stored at room temperature after harvest (**Supplementary Table 1**). In accordance with ripening, the color index of CIRG and the fruit weight increased gradually, while the soluble solids content changed slightly except at the S1 stage (**Supplementary Table 1**). In addition, six norisoprenoid volatiles were detected throughout fruit ripening and under ethylene treatment (**Figures 1A,B**). Norisoprenoid volatiles content and proportions changed obviously as peach ripening and β-ionone proportion reached nearly 40% after 3 days storage at 20°C (**Figure 1B**). β-ionone content increased from S1 to S3 gradually, while the accumulation accelerated after S4 and reached nearly 670 ng g^{-1} FW with a 20-fold increase compared to S3 (**Figure 1C**). The proportion of theaspirane gradually declined from 48 to 8.7% in contrast with β-ionone increase during peach ripening (**Figure 1B**). Interestingly, (*E, Z*)-4,6,8-Megastigmatriene was detected at S3, and the content was increased at S4 stage (**Figure 1B**). In addition, the transcript level of *PpCCD4* showed an increasing trend as peach ripened, and was dramatically elevated at S4 (**Figure 1D**). *PpCCD4* expression was positively correlated with content change of β-ionone ($R^2 = 0.85$, $p < 0.05$) (**Figure 1F**), while the transcript level of *PpCCD1* increased gradually with low change in transcript levels (**Figure 1E**).

Ethylene and 1-MCP Regulate β-Ionone Volatiles Content and *PpCCD4* Transcription

Peach fruit exhibited climacteric characteristics as ripening, and ethylene played an important role in activating the ripening process. Accumulation of norisoprenoid volatiles and expression pattern of *PpCCD4* were evaluated under ethylene and 1-MCP treatments with fruits at the turning color stage (**Figure 2**). Ripening process of peach was significantly accelerated by ethylene suffocating, and firmness decreased significantly from 60 N to 10 N after 60 h of storage at room temperature (**Figure 2A**), while 1-MCP treatment markedly delayed the fruit

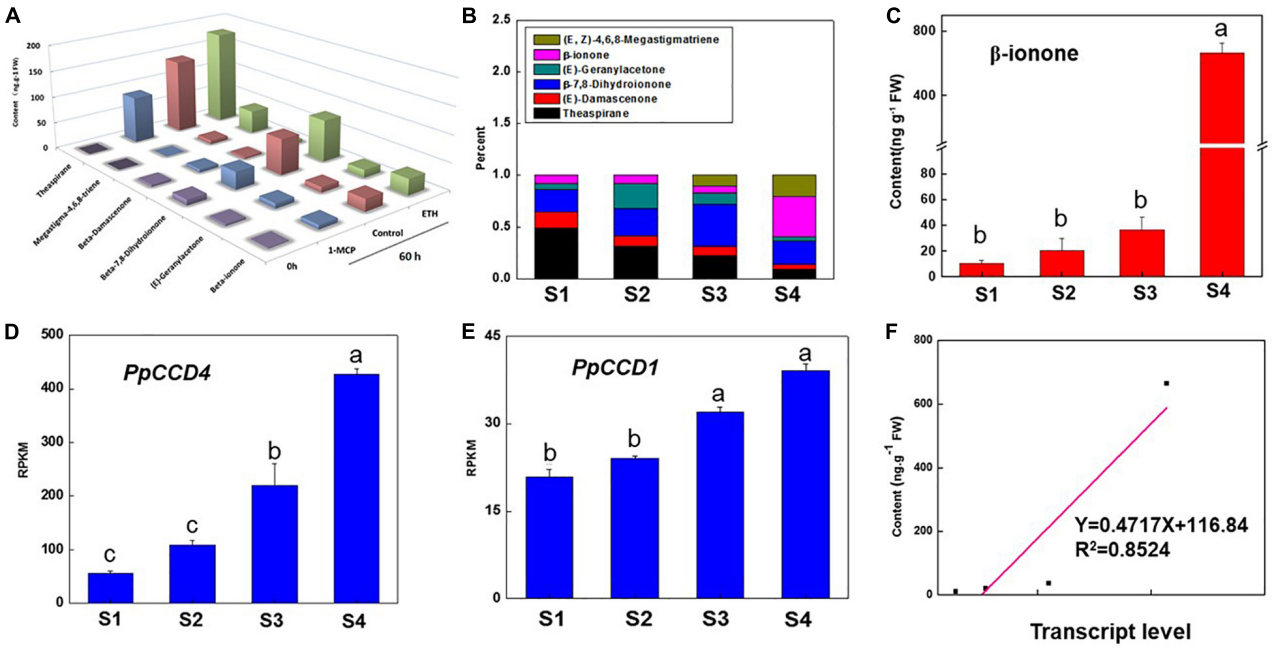


FIGURE 1 | Change of the apocarotenoid volatiles and transcript level of *PpCCD4* during peach ripening. **(A)** Change of apocarotenoid volatiles under ethylene treatment. **(B)** Percentage change of the apocarotenoid volatiles under different ripening process, S1 represents 110 days after bloom (DAB), S2 represents 113 DAB, S3 represents 116 DAB, and S4 represents fruits that were stored at room temperature for 3 days. **(C)** Change of β-ionone content. **(D)** Transcript patterns of *PpCCD4*. **(E)** Transcript patterns of *PpCCD1*. **(F)** Correlation analysis between β-ionone content and transcript levels of *PpCCD4*. Different letters indicate the significant difference with Duncan's multiple range tests at $p < 0.05$.

ripening process with firmness changed slightly (**Figure 2A**). Accumulation of all the six kinds of norisoprenoid volatiles showed an increased trend after 60 h of storage, and ethylene treatment accumulated much higher content than control, while 1-MCP treatment inhibited accumulation of all six norisoprenoid volatiles (**Figure 1A**). The concentration of β-ionone was significantly higher than control under ethylene treatment after 60 h of room temperature storage, while accumulation was inhibited dramatically by 1-MCP treatment (**Figure 2C**). Meanwhile, transcript levels of *PpCCD1* showed no sensitivity to ethylene and 1-MCP treatments (**Figure 2B**). In summary, ethylene induced the accumulation of β-ionone by activating transcription of *PpCCD4*, while opposite effects were achieved by 1-MCP treatment.

UV-B Irradiation-Mediated Carotenoids Metabolism in Peach

The color of the peach peel changed from pink to golden yellow after 48 h of UV-B irradiation (**Figure 3A**). The content of β-carotene, zeaxanthin, violaxanthin, and lutein was analyzed after 6 and 48 h of UV-B irradiation, respectively, in peach peel (**Figure 3B**). The lutein and violaxanthin accumulation occupied nearly 70% of the total four carotenoids, and β-carotene occupied 17% after 6 h of storage at room temperature (**Figure 3B**). Total content of the four carotenoids increased from 1,200 to 1,500 ng g^{-1} FW after 6 h of UV-B irradiation (**Figure 3B**). Numbers of key genes of carotenoids metabolism, such as *PSY*, *LCY-B*, *LCY-E*, and *CCDs*, have been identified in many higher plants (Fu et al., 2014; Nisar et al., 2015; Lado et al., 2016). Transcript levels of *PpPSY*, *PpLCY-B*, *PpLCY-E*, and *PpCHY-B* were significantly induced compared to control by 6 h of UV-B irradiation (**Figure 3C** and **Supplementary Table 2**). However, transcript levels of *PpPSY*, *PpLCY-B*, and *PpCHY-B* exhibited no significant difference compared with control except *PpLCY-E* after 48 h of UV-B irradiation (**Figure 3C**). After 48 h of storage at room temperature, total content of the four carotenoids reduced significantly compared with 6 h, indicating the biosynthesis of carotenoids was not the master factor in influencing carotenoids accumulation during fruit ripening (**Figure 3B**). Carotenoids proportion changed dramatically from 6 to 48 h in control group (**Figure 3B**). The proportion of β-carotene decreased from 21 to 9%, while the percentage of zeaxanthin increased from 4 to 14%, indicating the degradation of β-carotene was elevated during peach normal ripening (**Figure 3B**). However, the degradation of β-carotene was inhibited by UV-B irradiation, and β-carotene content accumulated 2-fold compared to control after 48 h of UV-B irradiation (**Figure 3B**). Meanwhile, the total content of downstream zeaxanthin and violaxanthin increased slightly by about 30% after 48 h of UV-B irradiation, prefiguring the other catabolism pathway catalyzed by CCDs that occupied the main proportion of β-carotene degradation and was inhibited after 48 h of UV-B irradiation (**Figure 3A**). In accordance with this, the transcript level of *PpCCD4* was significantly reduced after 48 h of UV-B irradiation (**Figure 4B**). On the contrary, the lutein accumulation increased significantly from 150 to 280 ng g^{-1},

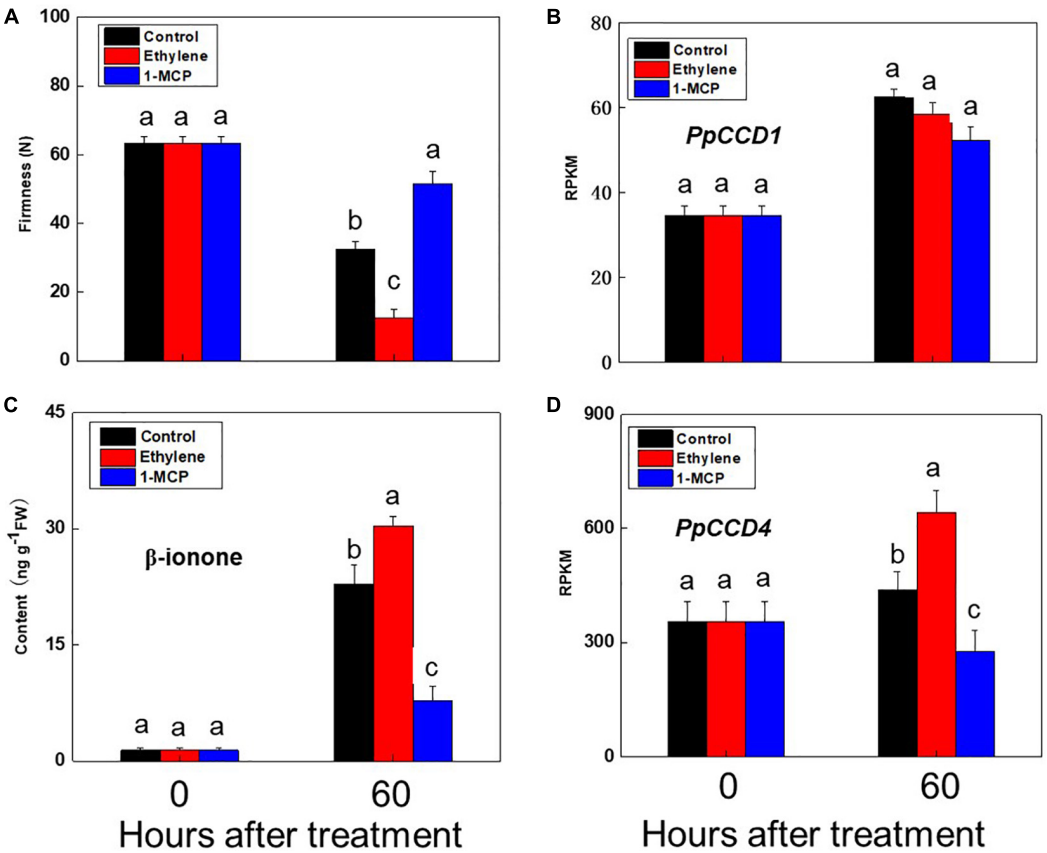


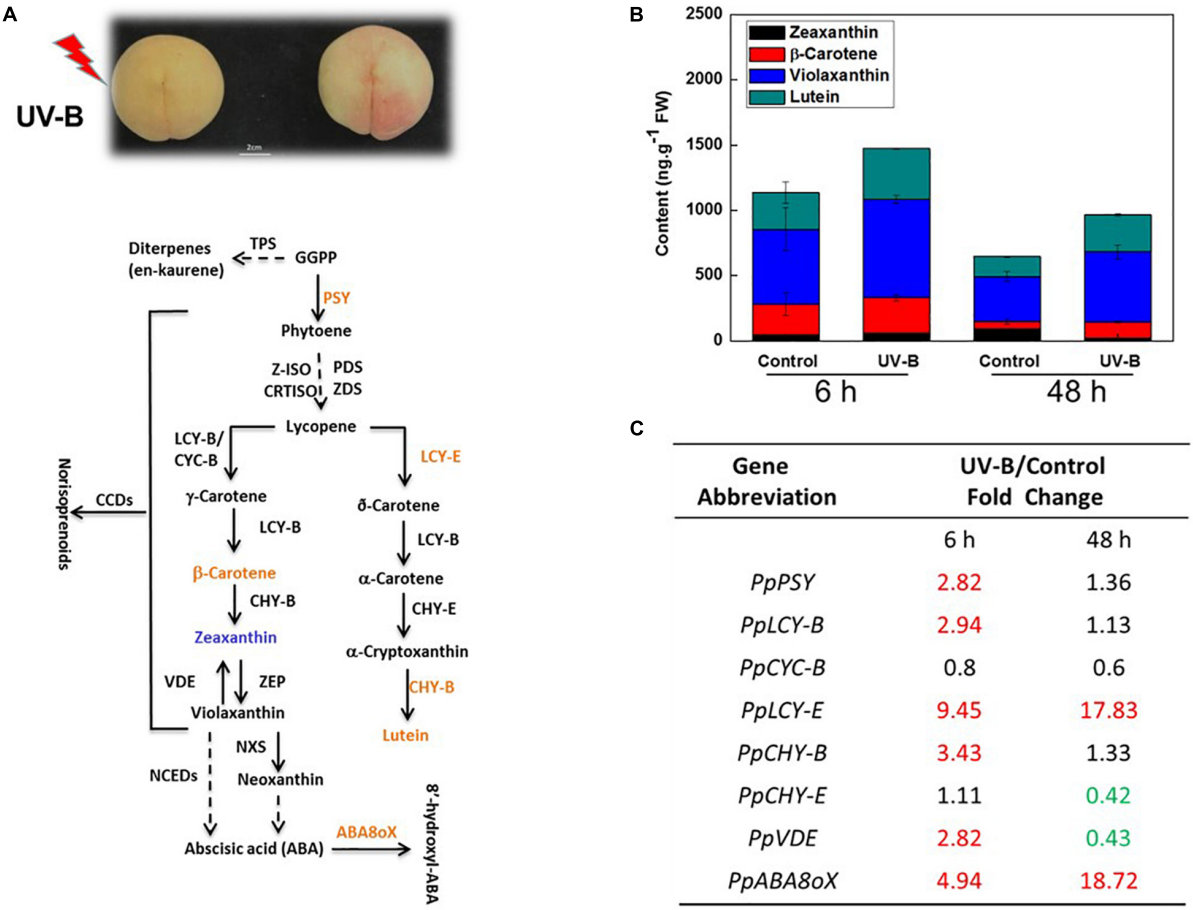
FIGURE 2 | Change of β-ionone content and *PpCCDs* transcript levels after ethylene and 1-MCP treatments. **(A)** Change of the firmness. **(B)** Transcript level of PpCCD1. **(C)** Content of β-ionone. **(D)** Transcript level of *PpCCD4*. The fruits were suffocated with 12 h and then stored at room temperature for 60 h. Different letters indicate the significant difference between the two columns ($p < 0.05$).

prefiguring that the biosynthesis of lutein contributed mostly to the accumulation at least after 48 h of UV-B irradiation for the induction of *PpLCY-E* transcript level (**Figures 3B,C**).

The biosynthesis and catabolism of carotenoids are determined by the numbers of genes labeled in the schematic map pathway, while usually several key genes in the pathway obviously influence the metabolic rates and directions (**Figure 3A**). The transcript level of the first rate-limiting gene, *PpPSY*, increased to 2.8-fold compared with control at 6 h of UV-B irradiation (**Figure 3C**). *PpLCY-B* controlled carbon metabolic flux to β-carotene access from lycopene, while *PpLCY-E* influenced the production of lutein (**Figure 3A**). Transcript levels of two vital branched genes *PpLCY-B*, and *PpLCY-E* were raised to 2.94- and 9.45-fold compared with control at 6 h of UV-B irradiation, respectively (**Figure 3C**). The downstream genes of β-β ring branched pathways *PpCHY-B, PpABA8oX* were all induced at 6 h of UV-B irradiation, indicating carotenoids catabolism and phytohormone ABA responded actively to UV-B (**Figure 3C**). However, transcript levels of the related genes were not induced as UV-B irradiation held for 48 h, except for *PpLCY-E* and *PpABA8oX* (**Figure 3C**). The continuous high transcript level of *PpLCY-E* supported the enhancement of lutein content after 48 h of UV-B irradiation (**Figure 3B**). *PpVDE* catalyzed violaxanthin to zeaxanthin and the transcript level changed from 2.82-fold induction to 57% inhibition during the UV-B irradiation duration, corresponding to a 78% reduction of zeaxanthin content after 48 h of UV-B irradiation (**Figures 3B,C**). Thus, carotenoids components and contents were influenced by transcript changes of different genes, and the metabolic sensitivity of carotenoids was increased by UV-B dose.

PpCCD4-Mediated Change of β-Carotene and Norisoprenoid Volatiles Under UV-B Irradiation

Norisoprenoid volatiles were produced by cleavage of carotenoids, and β-carotene was one of the main substrates. Five kinds of norisoprenoid volatiles were detected in peach peel after 48 h of storage at room temperature (**Supplementary Table 3**). The content of dihydro-β-ionone and β-ionone increased significantly after 48 h of storage (**Figure 4A** and **Supplementary Table 3**). However, β-ionone content was nearly undetected after 48 h of UV-B irradiation, and its hydrogenation product of dihydro-β-ionone content decreased from 250 to 50 ng g^{-1} FW (**Figure 4A**). In common with content reduction of β-ionone and dihydro-β-ionone, the transcript level of



C

Gene Abbreviation	UV-B/Control Fold Change	
	6 h	48 h
PpPSY	2.82	1.36
PpLCY-B	2.94	1.13
PpCYC-B	0.8	0.6
PpLCY-E	9.45	17.83
PpCHY-B	3.43	1.33
PpCHY-E	1.11	0.42
PpVDE	2.82	0.43
PpABA8oX	4.94	18.72

FIGURE 3 | Carotenoids content and transcript level of genes regulated under UV-B irradiation. **(A)** The schematic map of carotenoids metabolic pathways. **(B)** Change of the carotenoids content. **(C)** Fold changes of the related genes transcript levels between UV-B and the control. The red color numbers mean the significant induction, and the green color mean the significant inhibition under UV-B irradiation with $p < 0.05$.

PpCCD4 decreased to 30% compared with control after 48 h of UV-B irradiation (**Figure 4B**). To further clarify the influence of UV-B, peach fruits were irradiated with different doses and intensities. The irradiation showed no obvious inhibition effects within 0.5 h, while, β-ionone content exhibited a significant reduction at 24 h of irradiation, and inhibition effects remained till 72 h of recovery in the dark after 24 h of UV-B irradiation (**Supplementary Figure 1**). The dose experiments showed that the inhibition effects were stronger under high intensity irradiation. Besides, the transcript level of *PpCCD1* was induced by UV-B irradiation at 6 h, indicating *PpCCD1* may play redundant roles in β-carotene catabolism and the formation of β-ionone (**Figure 4B**). Thus, UV-B irradiation decreased the biosynthesis of β-ionone by inhibiting the transcript level of *PpCCD4*.

DISCUSSION

Carotenoids are compositions of various color of plants organisms, such as a class of natural pigments that attract pollinators and dispersers for plants (Lewinsohn et al., 2005; Auldridge et al., 2006; Walter et al., 2010; Yuan et al., 2015). Carotenoids are also photosynthetic accessory pigments, participating in plant photosynthesis and photoprotection (Muller et al., 2001; Cerullo et al., 2002; Holt et al., 2005). In addition, carotenoids can be cleaved by CCD enzymes to produce apocarotenoids, some of which are key contributors to flavor and floral aroma in fruits and vegetables (Lewinsohn et al., 2005; Yuan et al., 2015). The structural diversity of apocarotenoids indicates the vital roles in plant–environment interactions (Walter et al., 2010; Havaux, 2014; Brousseau et al., 2021). The key enzymes of carotenoids biosynthesis and catabolism have been identified in many plants, such as loquat (Zhou et al., 2007; Fu et al., 2014), tomato (Simkin et al., 2004a; Lado et al., 2016), and citrus (Alquezar et al., 2009; Ma et al., 2013a; Rodrigo et al., 2013). Reports of PpCCD4 have mainly been limited to the genetic comparative analysis of white- and yellow-fleshed colors in peach varieties and carotenoids accumulation in leaves (Brandi et al., 2011; Falchi et al., 2013; Ma et al., 2013b). The molecular mechanism between carotenoids and volatile apocarotenoids still exists, especially under the abiotic stress. In our study,

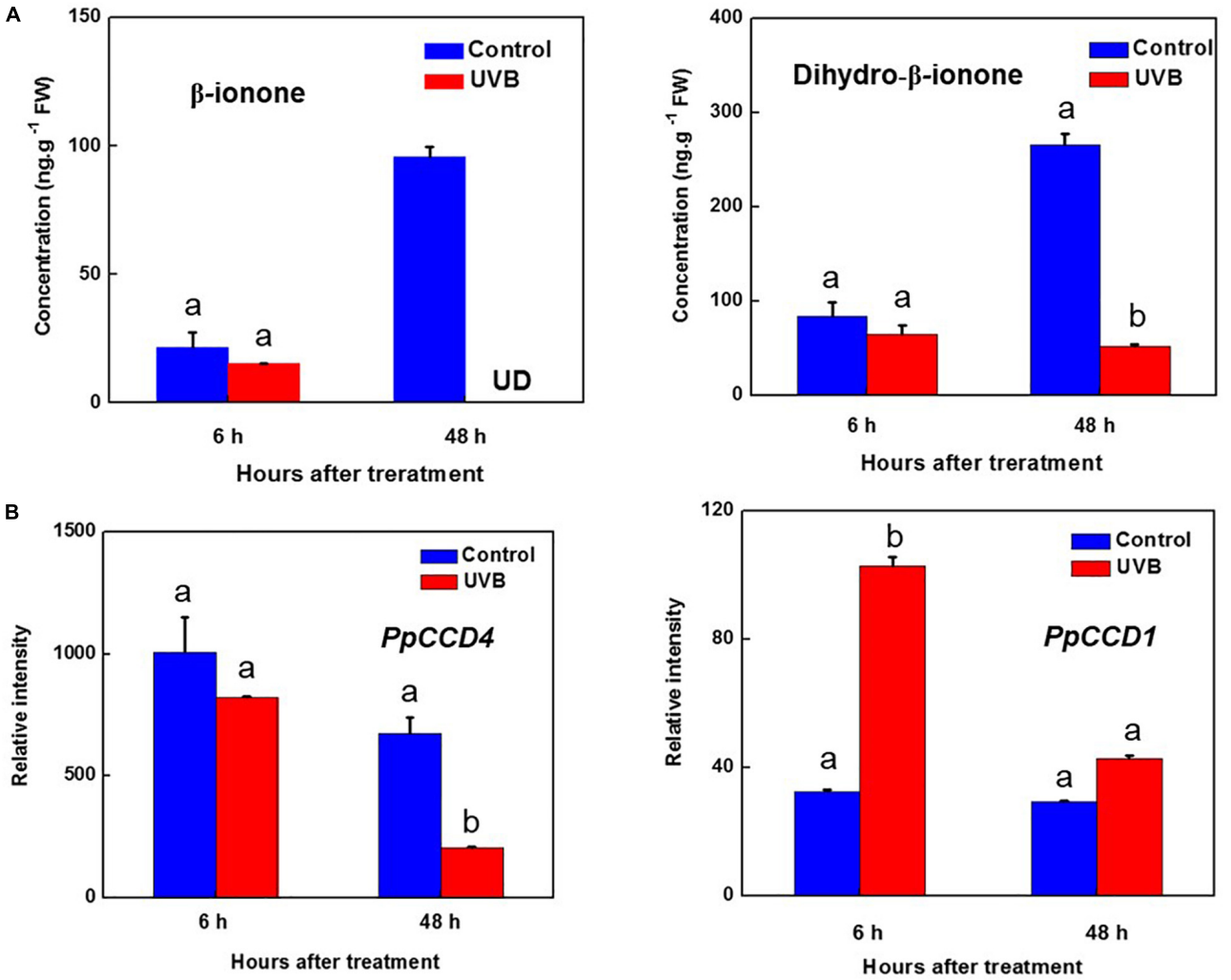


FIGURE 4 | Transcript levels of *PpCCDs* and content of norisoprenoids changed by UV-B irradiation. **(A)** Content of the β-ionone and Dihydro-β-ionone. **(B)** Transcript levels of *PpCCD4* and *PpCCD1*. Different letters indicate significant differences between UV-B treatment and control ($p < 0.05$).

change of the main carotenoids and volatile apocarotenoids were determined and mediated by the key carotenogenic genes of the pathway and *PpCCD4* under UV-B irradiation. In addition, transcription of *PpCCD4* corresponded to the increase of β-ionone through peach ripening, in accordance with the ethylene and 1-MCP treatments.

Change of Norisoprenoid Volatiles Were Correlated With Transcript Level of *PpCCD4* Through Peach Ripening and Ethylene Treatment

Norisoprenoid volatiles are widely distributed in fruits, flowers, and leaves of plants (Bouvier et al., 2005; Lewinsohn et al., 2005), exhibiting floral and fruity aroma attributes with low threshold concentration. Several reports have indicated that *PpCCD4* influenced carotenoids degradation and norisoprenoid volatiles accumulation in peach (Brandi et al., 2011; Falchi et al., 2013; Ma et al., 2013b). In our study, six main norisoprenoid volatiles were detected in peach, and their proportions changed significantly during the ripening process (**Figures 1A,B**). The proportion of theaspirane decreased gradually, while the proportion of β-ionone dramatically elevated at shelf-3-days, indicating β-ionone might contribute much to peach flavor aroma and seed dispersal (**Figures 1B,C**). Corresponding to this, the transcript level of *PpCCD4* was significantly induced as ripening (**Figure 1D**). The concentration of β-ionone and dihydro-β-ionone increased as the two apricot varieties ripened (Xi et al., 2020). To verify the metabolic mechanism of norisoprenoid volatiles through peach ripening, phytohormone ethylene and its competitive inhibitor 1-MCP were used to perform fumigation on turning color peach. Ethylene significantly promoted the peach ripening process with firmness decreasing markedly, while the fruit ripening process was obviously delayed under 1-MCP treatment with firmness decreasing less compared with the initial fruit (**Figure 2A**).

Ethylene significantly elevated β-ionone content, while 1-MCP inhibited the accumulation of β-ionone dramatically, and transcript level of *PpCCD4* exhibited similar patterns (**Figures 2C,D**). All the six norisoprenoid volatiles were induced by ethylene and inhibited by 1-MCP treatments (**Figure 1A**). However, transcript level of *PpCCD1* showed no obvious sensitivity to ethylene and 1-MCP treatment (**Figure 2B**), which was corresponding to the change pattern of transcript level during peach ripening with no obvious response to ethylene peak at shelf-3 days (**Figure 1E**). Negative correlation was observed between levels of carotenoids and norisoprenoid volatiles *via* expression of *CCD1s* increased as fruit ripened in grapes (Mathieu et al., 2005), apricots (Xi et al., 2020), and watermelon (Lewinsohn et al., 2005). The positive correlation between *PpCCD4* transcript level and β-ionone content change indicated that *PpCCD4* was the potential biosynthetic gene of β-ionone in peach (**Figure 1F**), which was consistent with the previous reports (Brandi et al., 2011; Falchi et al., 2013; Bai et al., 2016).

Carotenoids Metabolism Were Regulated by *PpPSY, PpLCY-E, PpLCY-B,* and *PpCCD4* Under UV-B Radiation

Carotenoids are mainly divided into two subgroups, namely, xanthophylls and carotenes, and some carotenoids are accessory pigments in photosynthesis by enhancing light harvesting, while some antioxidative xanthophylls can enhance plants' tolerance to light damage by quenching reactive oxygen species induced by excessive irradiation (Muller et al., 2001; Cerullo et al., 2002; Robert et al., 2004; Holt et al., 2005). Carotenoids accumulation is associated with the expression of several key genes involved in the pathway, including *PSY*, *LCY-E*, *LCY-b*, *CHY-b*, and *VDE*, and is also regulated by related transcription factors, phytohormones, and abiotic stress (Marty et al., 2005; Welsch et al., 2007; Toledo-Ortiz et al., 2010). It is known that total carotenoids are commonly induced by UV irradiation. The field UV-B can alter β-carotene and lutein content during grapeberry development (Joubert et al., 2016). In our study, the transcript level of *PpPSY* has no significant increase after 48 h of UV-B irradiation, indicating that the total carbon flux flowing to carotenoids has no increase any more (**Figure 3C**). However, the transcript level of *PpLCY-E* continuously increased from 9.45- to 17.83-fold compared with control, resulting in the final product lutein that increased nearly 80% after 48 h of UV-B irradiation (**Figures 3B,C**). Lutein was involved in photoprotection of PSII system against the oxidative damage induced by excessed UV-B radiation in developmental grape berries (Biswas et al., 2020). Meanwhile, the production of lutein in wheat endosperm could be supplied amply with a very low expression of *LCYE-B* (Yu et al., 2021).

The content of β-carotene was regulated by expression of *PpPSY*, *PpLCY-B*, *PpCHY-B*, *PpVDE*, *PpCCD4*, and *PpCCD1* in β–βring branched pathway, and they were all induced at 6 h UV-B radiation except *PpCCD4* (**Figures 3C**, **4B**). In accordance with the transcript level change, β-carotene content increased nearly 15%. However, the degradation of β-carotene by PpCCD4 played a vital role after 48 h of UV-B irradiation, for biosynthesis had no significant elevation while β-carotene content was markedly increased to 2-fold compared to control (**Figure 3B**). Corresponding to this, the transcript level of *PpCCD4* was significantly decreased to 30% compared to control after 48 h of UV-B irradiation (**Figure 4B**). Interestingly, the content of violaxanthin increased nearly 30% at 6 h of UV-B radiation, and the increase reached 56% compared to control after 48 h of UV-B irradiation (**Figure 3B**). Violaxanthin was deoxidized to zeaxanthin by *PpVDE*, and the deoxygenation process was significantly induced at 6 h with appropriate UV-B radiation, while the process was dramatically inhibited by 48 h UV-B radiation (**Figure 3C**). Accordingly, the content of zeaxanthin was significantly increased by 30% after 6 h of UV-B radiation, while the accumulation was markedly decreased by nearly 70% after 48 h of UV-B irradiation compared with control (**Figure 3B**). Violaxanthin has been known to have more resistance to high UV-B stress compared with other xanthophylls (Joshi et al., 2007; Lidon et al., 2012; Biswas et al., 2020), and the blocking of the violaxanthin de-epoxidation process by *PpVDE* indicated the photoprotection effect of zeaxanthin was limited under high UV-B irradiation. Two *PpNCEDs* (*PpNCED2, PpNCED3*) were surveyed in peach genome, and *PpNCED2* was induced slightly by UV-B, while *PpNCED3* had no response to UV-B irradiation (**Supplementary Figure 2A**). Besides, the transcript level of *PpNCED2* was elevated as the storage duration, which was corresponding to the induction pattern under the ethylene treatment (**Supplementary Figure 2B**). The transcript level of the *PpNCED2* was significantly increased from the turning color stage (S1) to the commercial maturity (**Supplementary Figure 2C**), which might be related to the metabolism of stress phytohormone ABA during fruit ripening.

UV-B Controlled Formation of β-Ionone by Influencing Transcription of *PpCCD4*

UV-B influenced the metabolism of carotenoids by regulating key biosynthesis genes in the reaction chains. Meanwhile, catabolism of carotenoids, especially producing norisoprenoid volatiles by CCDs, was significantly influenced by UV-B irradiation, and the aroma flavor and colors of peach were changed. Transcript level of *PpCCD4* was decreased by high UV-B radiation at 48 h, while the influence was insignificant at 6 h of radiation (**Figure 4B**). In accordance with this, β-ionone content increased markedly in control, while the concentration was even under detected after 48 h of UV-B irradiation (**Figure 4A**). The content of dihydro-β-ionone occupied nearly 60% of the total norisoprenoid volatiles after 48 h of storage at room temperature (**Supplementary Table 3**), while UV-B irradiation decreased 80% of the dihydro-β-ionone content (**Figure 4B**). The inhibition of norisoprenoid volatiles content depended on UV-B irradiation intensity and doses, and the inhibition effect was obvious after 24 h of radiation and lasted till 72 h of recovery in the dark, and the high intensity (UVB-150) showed a stronger inhibition effect (**Supplementary Figure 1**). The norisoprenoid volatiles have been reported to be the substrates of glucosylation by the uridine diphosphate sugar-dependent glycosyltransferases (UGTs) in *N. benthamiana* and *M.* × *piperita* (Sun et al., 2020). The apocarotenoid modification

reactions share partial key enzymes with the detoxification induced by xenobiotics and some reactive carbonyl species (Koschmieder et al., 2021). Thus, the apocarotenoids might have been modified further in peach fruit, and the non-volatile glycosylated dihydro-β-ionone has been detected in our previous reports (Wu et al., 2017, 2019). A total of 13 *PpUGTs* were induced after 48 h of UV-B irradiation, and the transcript levels of their products were positively correlated with the content change of glycosylated dihydro-β-ionone during development (Wu et al., 2017). Thus, the candidate *PpUGTs* may transfer the free β-ionone and dihydro-β-ionone to the non-volatile glycoslated form in peach (**Supplementary Table 4**). Transcript level of *PpCCD1* was significantly induced at 6 h of UV-B irradiation, while the induction was not continuous till 48 h of irradiation (**Figure 4B**). The study indicated that the transcript level of *PpCCD4* may determine the emission of norisoprenoid volatiles by catalyzing the carotenoids in peach, especially for β-carotene, and the process was significantly inhibited by UV-B irradiation. Interestingly, 15- and 60 min-exposure of freshly cut cantaloupe melon to UV radiation induced release of β-ionone, β-cyclocitral and geranylacetone for the inhibition effects on microbial growth (Lamikanra et al., 2002). Thus, in a more general context, carotenoids and their cleavage products participated in multiple facets of plant defense mechanisms and responded to various environmental stimuli.

CONCLUSION

In our study, the metabolic and molecular mechanisms of carotenoids and apocarotenoids volatiles were clarified under UV-B irradiation and during the fruit ripening process. Concentration of carotenoids were elevated at 6 h of UV-B irradiation with the induction of *PpPSY*, *PpLCY-B, PpLCY-E*, and *PpCHY-B*, whereas accumulation pattern of carotenoids and apocarotenoid volatiles was mainly influenced by the transcript levels of *PpCCD4*, *PpLCYE*, and *PpVDE* after 48 h of UV-B irradiation. The inhibition of *PpCCD4* might cause the accumulation of β-carotene and the reduction of β-ionone and dihydro-β-ionone, and the content of lutein and violaxanthin might be caused by the induction of *PpLCYE* and inhibition of *PpVDE* after 48 h of UV-B irradiation. In addition, β-ionone content increased during peach ripening and was induced by ethylene treatment. In summary, peach can regulate constitutes of carotenoids and mediate metabolism balance between carotenoids and apocarotenoid volatiles in response to UV-B irradiation stress. Meanwhile, there are still some works needed to be studied further. Firstly, biochemical activities of the key enzymes need to be determined with wide carotenoids substrates in peach to help understand the detailed biological roles completely. Secondly, the downstream metabolism of lutein and β-ionone needs to be studied deeply and offer us more information about the metabolic mechanism of carotenoids and apocarotenoid volatiles under UV-B irradiation. Finally, the transcriptional regulatory mechanisms need to be studied further, and the related transcription factors need to be identified. Meanwhile, the ethylene and UV-B light signal transduction pathways offered us clues.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**, further inquiries can be directed to the corresponding authors.

AUTHOR CONTRIBUTIONS

BZ, HL, and YQ designed the experiments. HL and XC performed the experiments. HL, MA, CW, and CL analyzed the data and wrote the manuscript. All authors contributed the article and approved the submitted version.

FUNDING

This work was supported by Shanghai Agriculture Applied Technology Development Program China [Hu nong ke tui zi (2021) 4-1], National Natural Science Foundation of China (31972379 and 32102451), the Shanghai Agricultural Products Preservation and Processing Engineering Technology Research Center (19DZ2251600), and the Shanghai Professional Service Platform for Agriculture Products Preservation and Processing (21DZ2292200).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2022.814677/full#supplementary-material

REFERENCES

Alquezar, B., Zacarias, L., and Rodrigo, M. J. (2009). Molecular and functional characterization of a novel chromoplast-specific lycopene beta-cyclase from citrus and its relation to lycopene accumulation. *J. Exp. Bot.* 60, 1783–1797. doi: 10.1093/jxb/erp048

Auldridge, M. E., McCarty, D. R., and Klee, H. J. (2006). Plant carotenoid cleavage oxygenases and their apocarotenoid products. *Curr. Opin. Plant Biol.* 9, 315–321. doi: 10.1016/j.pbi.2006.03.005

Bai, S. L., Tuan, P. A., Tatsuki, M., Yaegaki, H., Ohmiya, A., Yamamizo, C., et al. (2016). Knockdown of carotenoid cleavage dioxygenase 4 (CCD4) *via* virus-induced gene silencing confers yellow coloration in peach fruit: evaluation of gene function related to fruit traits. *Plant Mol. Biol. Rep.* 34, 254–264.

Baldermann, S., Kato, M., Kurosawa, M., Kurobayashi, Y., Fujita, A., Fleischmann, P., et al. (2010). Functional characterization of a carotenoid cleavage dioxygenase 1 and its relation to the carotenoid accumulation and volatile emission during the floral development of *Osmanthus* fragrans Lour. *J. Exp. Bot.* 61, 2967–2977. doi: 10.1093/jxb/erq123

Biswas, D. K., Ma, B. L., Xu, H., Li, Y. G., and Jiang, G. M. (2020). Lutein-mediated photoprotection of photosynthetic machinery in Arabidopsis thaliana exposed to chronic low ultraviolet-B radiation. *J. Plant physiol.* 248:153160. doi: 10.1016/j.jplph.2020.153160

Booker, J., Auldridge, M., Wills, S., McCarty, D., Klee, H., and Leyser, O. (2004). MAX3/CCD7 is a carotenoid cleavage dioxygenase required for the synthesis of a novel plant signaling molecule. *Curr. Biol.* 14, 1232–1238. doi: 10.1016/j.cub.2004.06.061

Bouvier, F., Isner, J. C., Dogbo, O., and Camara, B. (2005). Oxidative tailoring of carotenoids: a prospect towards novel functions in plants. *Trends Plant Sci.* 10, 187–194. doi: 10.1016/j.tplants.2005.02.007

Brandi, F., Bar, E., Mourgues, F., Horvath, G., Turcsi, E., Giuliano, G., et al. (2011). Study of 'Redhaven' peach and its white-fleshed mutant suggests a key role of CCD4 carotenoid dioxygenase in carotenoid and norisoprenoid volatile metabolism. *BMC Plant Biol.* 11:24. doi: 10.1186/1471-2229-11-24

Brousseau, V. D., Wu, B. S., MacPherson, S., Morello, V., and Lefsrud, M. (2021). Cannabinoids and terpenes: how production of photo-protectants can be manipulated to enhance *Cannabis sativa* L. phytochemistry. *Front. Plant Sci.* 12:620021. doi: 10.3389/fpls.2021.620021

Bruno, M., Koschmieder, J., Wuest, F., Schaub, P., Fehling-Kaschek, M., Timmer, J., et al. (2016). Enzymatic study on AtCCD4 and AtCCD7 and their potential to form acyclic regulatory metabolites. *J. Exp. Biol.* 67, 5993–6005. doi: 10.1093/jxb/erw356

Campbell, R., Ducreux, L. J. M., Morris, W. L., Morris, J. A., Suttle, J. C., Ramsay, G., et al. (2010). The metabolic and developmental roles of carotenoid cleavage dioxygenase 4 from potato. *Plant Physiol.* 154, 656–664. doi: 10.1104/pp.110.158733

Cao, X. M., Wei, C. Y., Duan, W. Y., Gao, Y., Kuang, J. F., Liu, M. C., et al. (2021). Transcriptional and epigenetic analysis reveals that NAC transcription factors regulate fruit flavor ester biosynthesis. *Plant J.* 106, 785–800. doi: 10.1111/tpj.15200

Cazzonelli, C. I., and Pogson, B. J. (2010). Source to sink: regulation of carotenoid biosynthesis in plants. *Trends Plant Sci.* 15, 266–274. doi: 10.1016/j.tplants.2010.02.003

Cerullo, G., Polli, D., Lanzani, G., De Silvestri, S., Hashimoto, H., and Cogdell, R. J. (2002). Photosynthetic light harvesting by carotenoids: detection of an intermediate excited state. *Science* 298, 2395–2398. doi: 10.1126/science.1074685

Chen, Z., and Gallie, D. R. (2012). Violaxanthin de-epoxidase is rate-limiting for non-photochemical quenching under subsaturating light or during chilling in Arabidopsis. *Plant Physiol. Biochem.* 58, 66–82. doi: 10.1016/j.plaphy.2012.06.010

De Poll, W. H. V., Lagunas, M., de Vries, T., Visser, R. J. W., and Buma, A. G. J. (2011). Non-photochemical quenching of chlorophyll fluorescence and xanthophyll cycle responses after excess PAR and UVR in *Chaetoceros Brevis*, *Phaeocystis* Antarctica and coastal Antarctic phytoplankton. *Mar. Ecol-Prog. Ser.* 426, 119–131. doi: 10.3354/meps09000

Falchi, R., Vendramin, E., Zanon, L., Scalabrin, S., Cipriani, G., Verde, I., et al. (2013). Three distinct mutational mechanisms acting on a single gene underpin the origin of yellow flesh in peach. *Plant J.* 76, 175–187. doi: 10.1111/tpj.12283

Farré, G., Sanahuja, G., Naqvi, S., Bai, C., Capell, T., Zhu, C., et al. (2010). Travel advice on the road to carotenoids in plants. *Plant Sci.* 179, 28–48. doi: 10.1016/j.plantsci.2010.03.009

Fu, X., Feng, C., Wang, C., Yin, X., Lu, P., Grierson, D., et al. (2014). Involvement of multiple phytoene synthase genes in tissue- and cultivar-specific accumulation of carotenoids in loquat. *J. Exp. Bot.* 65, 4679–4689. doi: 10.1093/jxb/eru257

Gil, M., Pontin, M., Berli, F., Bottini, R., and Piccoli, P. (2012). Metabolism of terpenes in the response of grape (*Vitis vinifera* L.) leaf tissues to UV-B radiation. *Phytochemistry* 77, 89–98. doi: 10.1016/j.phytochem.2011.12.011

Gonzalez-Jorge, S., Ha, S. H., Magallanes-Lundback, M., Gilliland, L. U., Zhou, A., Lipka, A. E., et al. (2013). Carotenoid cleavage dioxygenase4 is a negative regulator of beta-carotene content in Arabidopsis seeds. *Plant cell* 25, 4812–4826. doi: 10.1105/tpc.113.119677

Havaux, M. (2014). Carotenoid oxidation products as stress signals in plants. *Plant J.* 79, 597–606. doi: 10.1111/tpj.12386

Holt, N. E., Zigmantas, D., Valkunas, L., Li, X. P., Niyogi, K. K., and Fleming, G. R. (2005). Carotenoid cation formation and the regulation of photosynthetic light harvesting. *Science* 307, 433–436. doi: 10.1126/science.1105833

Hou, X., Rivers, J., Leon, P., McQuinn, R. P., and Pogson, B. J. (2016). Synthesis and function of apocarotenoid signals in plants. *Trends Plant Sci.* 21, 792–803. doi: 10.1016/j.tplants.2016.06.001

Huang, F. C., Horvath, G., Molnar, P., Turcsi, E., Deli, J., Schrader, J., et al. (2009). Substrate promiscuity of RdCCD1, a carotenoid cleavage oxygenase from *Rosa Damascena*. *Phytochemistry* 70, 457–464. doi: 10.1016/j.phytochem.2009.01.020

Ibdah, M., Azulay, Y., Portnoy, V., Wasserman, B., Bar, E., Meir, A., et al. (2006). Functional characterization of CmCCD1, a carotenoid cleavage dioxygenase from melon. *Phytochemistry* 67, 1579–1589. doi: 10.1016/j.phytochem.2006.02.009

Joshi, P. N., Ramaswamy, N. K., Iyer, R. K., Nair, J. S., Pradhan, M. K., Gartia, S., et al. (2007). Partial protection of photosynthetic apparatus from UV-B-induced damage by UV-A radiation. *Environ. Exp. Bot.* 59, 166–172. doi: 10.1016/j.envexpbot.2005.11.005

Joubert, C., Young, P. R., Eyeghe-Bickong, H. A., and Vivier, M. A. (2016). Field-grown grapevine berries use carotenoids and the associated xanthophyll cycles to acclimate to UV Exposure differentially in high and low light (shade) conditions. *Front. Plant Sci.* 7:786. doi: 10.3389/fpls.2016.00786

Kopsell, D. A., and Kopsell, D. E. (2006). Accumulation and bioavailability of dietary carotenoids in vegetable crops. *Trends Plant Sci.* 11, 499–507. doi: 10.1016/j.tplants.2006.08.006

Koschmieder, J., Wust, F., Schaub, P., Alvarez, D., Trautmann, D., Krischke, M., et al. (2021). Plant apocarotenoid metabolism utilizes defense mechanisms against reactive carbonyl species and xenobiotics. *Plant Physiol.* 185, 331–351. doi: 10.1093/plphys/kiaa033

Lado, J., Zacarias, L., and Rodrigo, M. J. (2016). "Regulation of carotenoid biosynthesis during fruit development," in *Carotenoids in Nature: Biosynthesis, Regulation and Function*, ed. C. Stange (New York, NY: Springer), 161–198. doi: 10.1007/978-3-319-39126-7_6

Lamikanra, O., Richard, O. A., and Parker, A. (2002). Ultraviolet induced stress response in fresh cut cantaloupe. *Phytochemistry* 60, 27–32. doi: 10.1016/s0031-9422(02)00048-1

Lewinsohn, E., Sitrit, Y., Bar, E., Azulay, Y., Ibdah, M., Meir, A., et al. (2005). Not just colors - carotenoid degradation as a link between pigmentation and aroma in tomato and watermelon fruit. *Trends Food Sci. Technol.* 16, 407–415. doi: 10.1016/j.tifs.2005.04.004

Liang, M. H., He, Y. J., Liu, D. M., and Jiang, J. G. (2021). Regulation of carotenoid degradation and production of apocarotenoids in natural and engineered organisms. *Crit. Rev. Biotechnol.* 41, 513–534. doi: 10.1080/07388551.2021.1873242

Lidon, F. J. C., Teixeira, M., and Ramalho, J. C. (2012). Decay of the chloroplast pool of ascorbate switches on the oxidative burst in UV-B-irradiated rice J. *Agron. Crop Sci.* 198, 130–144. doi: 10.1111/j.1439-037x.2011.00489.x

Liu, H. R., Cao, X. M., Liu, X. H., Xin, R., Wang, J. J., Gao, J., et al. (2017). UV-B irradiation differentially regulates terpene synthases and terpene content of peach. *Plant Cell Environ.* 40, 2261–2275. doi: 10.1111/pce.13029

Ma, G., Zhang, L., Matsuta, A., Matsutani, K., Yamawaki, K., Yahata, M., et al. (2013a). Enzymatic formation of beta-citraurin from beta-cryptoxanthin and zeaxanthin by carotenoid cleavage dioxygenase4 in the flavedo of citrus fruit. *Plant Physiol.* 163, 682–695. doi: 10.1104/pp.113.223297

Ma, J., Li, J., Zhao, J., Zhou, H., Ren, F., Wang, L., et al. (2013b). Inactivation of a gene encoding carotenoid cleavage dioxygenase (CCD4) leads to carotenoid-based yellow coloration of fruit flesh and leaf midvein in peach. *Plant Mol. Biol. Rep.* 32, 246–257. doi: 10.1007/s11105-013-0650-8

Marty, I., Bureau, S., Sarkissian, G., Gouble, B., Audergon, J. M., and Albagnac, G. (2005). Ethylene regulation of carotenoid accumulation and carotenogenic gene expression in colour-contrasted apricot varieties (*Prunus Armeniaca*). *J. Exp. Bot.* 56, 1877–1886. doi: 10.1093/jxb/eri177

Mathieu, S., Terrier, N., Procureur, J., Bigey, F., and Gunata, Z. (2005). A carotenoid cleavage dioxygenase from *Vitis vinifera* L.: functional characterization and expression during grape berry development in relation to C13-norisoprenoid accumulation. *J. Exp. Bot.* 56, 2721–2731. doi: 10.1093/jxb/eri265

Messing, S. A. J., Gabelli, S. B., Echeverria, I., Vogel, J. T., Guan, J. C., Tan, B. C., et al. (2010). Structural insights into maize viviparous14, a key enzyme in the biosynthesis of the phytohormone abscisic acid. *Plant cell* 22, 2970–2980. doi: 10.1105/tpc.110.074815

Muller, P., Li, X. P., and Niyogi, K. K. (2001). Non-photochemical quenching. a response to excess light energy. *Plant Physiol.* 125, 1558–1566. doi: 10.1104/pp.125.4.1558

Nisar, N., Li, L., Lu, S., Khin, N. C., and Pogson, B. J. (2015). Carotenoid metabolism in plants. *Mol. Plant* 8, 68–82. doi: 10.1016/j.molp.2014.12.007

Ramel, F., Mialoundama, A. S., and Havaux, M. (2013). Nonenzymic carotenoid oxidation and photooxidative stress signalling in plants. *J. Exp. Bot.* 64, 799–805. doi: 10.1093/jxb/ers223

Robert, B., Horton, P., Pascal, A. A., and Ruban, A. V. (2004). Insights into the molecular dynamics of plant light-harvesting proteins in vivo. *Trends Plant Sci.* 9, 385–390. doi: 10.1016/j.tplants.2004.06.006

Rodrigo, M. J., Alquezar, B., Alos, E., Medina, V., Carmona, L., Bruno, M., et al. (2013). A novel carotenoid cleavage activity involved in the biosynthesis of citrus fruit-specific apocarotenoid pigments. *J. Exp. Bot.* 64, 4461–4478. doi: 10.1093/jxb/ert260

Rubio-Moraga, A., Rambla, J. L., Fernandez-de-Carmen, A., Trapero-Mozos, A., Ahrazem, O., Orzaez, D., et al. (2014). New target carotenoids for CCD4 enzymes are revealed with the characterization of a novel stress-induced carotenoid cleavage dioxygenase gene from *Crocus sativus*. *Plant Mol. Biol.* 86, 555–569. doi: 10.1007/s11103-014-0250-5

Shi, J., Cao, C., Xu, J., and Zhou, C. (2020). Research advances on biosynthesis, regulation, and biological activities of apocarotenoid aroma in horticultural plants. *J. Chem.* 2020, 1–11. doi: 10.1155/2020/2526956

Simkin, A. J., Schwartz, S. H., Auldridge, M., Taylor, M. G., and Klee, H. J. (2004a). The tomato carotenoid cleavage dioxygenase 1 genes contribute to the formation of the flavor volatiles beta-ionone, pseudoionone, and geranylacetone. *Plant J.* 40, 882–892. doi: 10.1111/j.1365-313X.2004.02263.x

Simkin, A. J., Underwood, B. A., Auldridge, M., Loucas, H. M., Shibuya, K., Schmelz, E., et al. (2004b). Circadian regulation of the PhCCD1 carotenoid cleavage dioxygenase controls emission of beta-ionone, a fragrance volatile of petunia flowers. *Plant Physiol.* 136, 3504–3514. doi: 10.1104/pp.104.049718

Song, M.-H., Lim, S.-H., Kim, J. K., Jung, E. S., John, K. M. M., You, M.-K., et al. (2016). In planta cleavage of carotenoids by Arabidopsis carotenoid cleavage dioxygenase 4 in transgenic rice plants. *Plant Biotechnol. Rep.* 10, 291–300. doi: 10.1007/s11816-016-0405-8

Sun, G. X., Natalia, P., Sina, B., Rafal, J., Tarik, F., Thomas, H., et al. (2020). Six uridine-diphosphate glycosyltransferases catalyze the glycosylation of bioactive C13-apocarotenols. *Plant Physiol.* 184, 1744–1761. doi: 10.1104/pp.20.00953

Tan, B. C., Joseph, L. M., Deng, W. T., Liu, L. J., Li, Q. B., Cline, K., et al. (2003). Molecular characterization of the Arabidopsis 9-cis epoxycarotenoid dioxygenase gene family. *Plant J* 35, 44–56. doi: 10.1046/j.1365-313x.2003.01786.x

Toledo-Ortiz, G., Huq, E., and Rodriguez-Concepcion, M. (2010). Direct regulation of phytoene synthase gene expression and carotenoid biosynthesis by phytochrome-interacting factors. *Proc. Natl. Acad. Sci. U.S.A.* 107, 11626–11631. doi: 10.1073/pnas.0914428107

Vallabhaneni, R., and Wurtzel, E. T. (2010). From epoxycarotenoids to ABA: the role of ABA 8′-hydroxylases in drought-stressed maize roots. *Arch. Biochem. Biophys.* 504, 112–117. doi: 10.1016/j.abb.2010.07.005

Walter, M. H., Floss, D. S., and Strack, D. (2010). Apocarotenoids: hormones, mycorrhizal metabolites and aroma volatiles. *Planta* 232, 1–17. doi: 10.1007/s00425-010-1156-3

Walter, M. H., and Strack, D. (2011). Carotenoids and their cleavage products: biosynthesis and functions. *Nat. Prod. Rep.* 28, 663–692. doi: 10.1039/c0np00036a

Wang, J. M., Zhang, N., Zhao, M. Y., Jing, T. T., Jin, J. Y., Wu, B., et al. (2020). Carotenoid cleavage dioxygenase 4 catalyzes the formation of carotenoid derived volatile β-ionone during tea (*Camellia Sinensis*) withering. *J. Agric. Food Chem.* 68, 1684–1690. doi: 10.1021/acs.jafc.9b07578

Wei, C., Liu, H., Cao, X., Zhang, M., Li, X., Chen, K., et al. (2021). Synthesis of flavour-related linalool is regulated by PpbHLH1 and associated with changes in DNA methylation during peach fruit ripening. *Plant Biotechnol. J.* 19, 2082–2096. doi: 10.1111/pbi.13638

Wei, S., Hannoufa, A., Soroka, J., Xu, N., Li, X., Zebarjadi, A., et al. (2011). Enhanced beta-ionone emission in Arabidopsis over-expressing AtCCD1 reduces feeding damage in vivo by the crucifer flea beetle. *Environ. Entomol.* 40, 1622–1630. doi: 10.1603/EN11088

Welsch, R., Maass, D., Voegel, T., DellaPenna, D., and Beyer, P. (2007). Transcription factor RAP2.2 and its interacting partner SINAT2: stable elements in the carotenogenesis of Arabidopsis leaves. *Plant Physiol.* 145, 1073–1085. doi: 10.1104/pp.107.104828

Wu, B. P., Cao, X. M., Liu, H. R., Zhu, C. Q., Klee, H., Zhang, B., et al. (2019). UDP-glucosyltransferase PpUGT85A2 controls volatile glycosylation in peach. *J. Exp. Bot.* 70, 925–936. doi: 10.1093/jxb/ery419

Wu, B. P., Gao, L. X., Gao, J., Xu, Y. Y., Liu, H. R., Cao, X. M., et al. (2017). Genome-wide identification, expression patterns, and functional analysis of UDP glycosyltransferase family in peach (*Prunus Persica* L. Batsch). *Front. Plant Sci.* 8:389. doi: 10.3389/fpls.2017.00389

Xi, W. P., Zhang, L. A., Liu, S. Y., and Zhao, G. H. (2020). The Genes of CYP, ZEP, and CCD1/4 play an important role in controlling carotenoid and aroma volatile apocarotenoid accumulation of apricot fruit. *Front. Plant Sci.* 11:607715. doi: 10.3389/fpls.2020.607715

Yoshioka, S., Aida, R., Yamamizo, C., Shibata, M., and Ohmiya, A. (2012). The carotenoid cleavage dioxygenase 4 (CmCCD4a) gene family encodes a key regulator of petal color mutation in *chrysanthemum*. *Euphytica* 184, 377–387. doi: 10.1007/s10681-011-0602-z

Yu, S., Li, M., Dubcovsky, J., and Tian, L. (2021). Mutant combinations of lycopene epsilon-cyclase and beta-carotene hydroxylase 2 homoeologs increased beta-carotene accumulation in endosperm of tetraploid wheat (*Triticum turgidum* L.) grains. *Plant Biotechnol. J.* 20, 564–576. doi: 10.1111/pbi.13738

Yuan, H., Zhang, J., Nageswaran, D., and Li, L. (2015). Carotenoid metabolism and regulation in horticultural crops. *Hort. Res.* 2:15036. doi: 10.1038/hortres.2015.36

Zeng, X. Q., Yuan, H. J., Dong, X. K., Peng, M., Jing, X. Y., Xu, Q. J., et al. (2020). Genome-wide dissection of co-selected UV-b responsive pathways in the UV-B adaptation of qingke. *Mol. Plant* 13, 112–127. doi: 10.1016/j.molp.2019.10.009

Zhang, B., Liu, C., Wang, Y., Xuan, Y., and Liu, K. (2015). Disruption of a carotenoid cleavage dioxygenase 4 gene converts flower colour from white to yellow in brassica species. *New Phytol.* 206, 1513–1526. doi: 10.1111/nph.13335

Zhang, B., Shen, J. Y., Wei, W. W., Xi, W. P., Xu, C. J., Ferguson, I., et al. (2010). Expression of genes associated with aroma formation derived from the fatty acid pathway during peach fruit ripening. *J. Agric. Food Chem.* 58, 6157–6165. doi: 10.1021/jf100172e

Zhang, B., Tieman, D. M., Jiao, C., Xu, Y. M., Chen, K. S., Fei, Z. J., et al. (2016). Chilling-induced tomato flavor loss is associated with altered volatile synthesis and transient changes in DNA methylation. *Proc. Natl. Acad. Sci. U.S.A.* 113, 12580–12585. doi: 10.1073/pnas.1613910113

Zheng, X., Yang, Y., and Al-Babili, S. (2021). Exploring the diversity and regulation of apocarotenoid metabolic pathways in plants. *Front. Plant Sci.* 12:787049. doi: 10.3389/fpls.2021.787049

Zhou, C. H., Xu, C. J., Sun, C. D., Li, X., and Chen, K. S. (2007). Carotenoids in white- and red-fleshed loquat fruits. *J. Agric. Food Chem.* 55, 7822–7830. doi: 10.1021/jf071273h

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.

frontiers in Plant Science

ORIGINAL RESEARCH
published: 16 February 2022
doi: 10.3389/fpls.2022.818345

Multi-Omics and miRNA Interaction Joint Analysis Highlight New Insights Into Anthocyanin Biosynthesis in Peanuts (*Arachis hypogaea* L.)

Jiawei Li, Yucong Ma, Mengdie Hu, Yulu Zhao, Bin Liu, Chunmei Wang, Min Zhang, Liping Zhang, Xinlei Yang and Guojun Mu**

State Key Laboratory of North China Crop Improvement and Regulation, North China Key Laboratory for Crop Germplasm Resources of Education Ministry, Laboratory of Hebei Provincial Crop Germplasm Resources, Hebei Agricultural University, Baoding, China

OPEN ACCESS

Edited by:
Renata Rivera-Madrid, Scientific Research Center of Yucatán (CICY), Mexico

Reviewed by:
Shaohua Zeng, South China Botanical Garden, Chinese Academy of Sciences (CAS), China
Jian Zhao, Anhui Agricultural University, China

***Correspondence:**
Xinlei Yang
yangxinlei2500@163.com
Guojun Mu
mgj99999@126.com

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

Received: *24 November 2021*
Accepted: *24 January 2022*
Published: *16 February 2022*

Citation:
Li J, Ma Y, Hu M, Zhao Y, Liu B, Wang C, Zhang M, Zhang L, Yang X and Mu G (2022) Multi-Omics and miRNA Interaction Joint Analysis Highlight New Insights Into Anthocyanin Biosynthesis in Peanuts (Arachis hypogaea L.). Front. Plant Sci. 13:818345. doi: 10.3389/fpls.2022.818345

Peanut (*Arachis hypogaea* L.) is one of the most important economic and oil crops in the world. At present, peanut varieties with rich anthocyanin in testa are rare in the market, but the selection and breeding of varieties with the related traits has always attracted the attention of breeders. In this study, two peanut varieties with the pink and purple testa, G110 (G) and Z18-40 (Z) were used to conduct interaction joint analysis of multi-omics and miRNA-target gene. The anthocyanin content of Z18-40 was 7.49–8.62-folds higher than G110 on 30 DAF (days after flowering) and 45 DAF via Ultraviolet-visible Spectrophotometer (UV-5800, Shanghai, China). And then, a total of 14 candidate genes related with the anthocyanin biosynthesis were identified for correlation in different comparison groups ($R^2 \geq 0.80$), among of a novel gene *Ah21440* related with hydroxycinnamoyl transferase (HCT) biosynthesis was identified. In addition, Cyanidin 3-O-glucoside (Kuromanin, pmb0550) was the only common differentially accumulated metabolite (DAM) identified using multi-omics joint analysis in G1_vs_G2, Z1_vs_Z2, G1_vs_Z1, and G2_vs_Z2, respectively. Correlation analysis of miRNA-target genes and DEGs in the transcriptome shows that, *AhmiR2950*, *AhmiR398*, *AhmiR50*, and *AhmiR51* regulated to HCT and chalcone biosynthesis related candidate genes (*Ah21440*, *AhCHS*, *AhCHI*). Lastly, all of 14 candidate genes and 4 differentially expressed miRNAs were validated using quantitative real-time PCR (qRT-PCR), which trends were consistent with that of the former transcriptome data. The results provide important reference for in-depth research on the anthocyanin metabolism mechanism in peanut testa.

Keywords: peanut, anthocyanin, testa, multi-omics joint analysis, miRNA interaction, qRT-PCR

INTRODUCTION

Peanut (*Arachis hypogaea* L., 2n = 4*x* = 40) is widely planted worldwide (Huang et al., 2019), with its planting area in China being only second to that of India, although the former's total yield ranks first in the world. As a secondary metabolite in plants, anthocyanins are found in 72 genera from 27 families of flowering plants (Sarma et al., 1997) and they exhibit physiological health effects

such as anti-aging, anti-mutation, prevention of cardiovascular diseases, liver protection and anti-cancer (Li et al., 2018). As greater concern is now being shown for health issues, anthocyanin biosynthesis has, therefore, become a highly studied compound for researchers (Wu, 2018). Testa colors are closely associated with anthocyanin content, with the different types of anthocyanin determined by the 3′,4′,5′-OH numbers in the tricyclic compound C6-C3-C6 (Yao et al., 2004). Based on these differences, blue-violet delphinidin, magenta pelargonidin, red cyanidin along with their derivatives (i.e., peonidin, petunidin, and malvidin) form the six main anthocyanins present in nature (Jaakola, 2013).

The pathway for anthocyanin biosynthesis has already been determined (Gao et al., 2020). Basically, delphinidin, pelargonidin and cyanidin are synthesized from phenylalanine with phenylalanine ammonia-lyase (PAL), 4-coumarate, coenzyme a ligase (4CL), chalcone synthase (CHS), chalcone isomerase (CHI), flavanone-3-hydroxylase (F3H), dioxanonol-4-reductase (DFR), and anthocyanidin synthase (ANS) (Song et al., 2019). The common substrates dihydrokaempferol (DHK), dihydroquercetin (DHQ), and dihydromyricetin (DHM) produce flavonols and proanthocyanidins (PAs) under the action of flavonol synthase (FLS), leucoanthocyanidin reductase (LAR), and anthocyanidin reductase (ANR) (Trainin et al., 2021). Proanthocyanins are plant natural products that are beneficial for human and livestock health (Lu et al., 2021). During the process, PAs are irreversibly converted to anthocyanin under high temperature conditions, which further deepens the colors of plants (Deng, 2018). Color changes in various plants have been reported to be related to anthocyanin biosynthesis (e.g., white clover (*Trifolium repens*) (Duan et al., 2020), alfalfa (*Medicago sativa*) (Xue et al., 2019), white primrose (*Primula vulgaris*) (Li et al., 2019), strawberry (*Fragaria* × *ananassa*) (Zhang et al., 2018), and other plants.

The mRNA Interfering Complementary RNA (miRNA) or antisense RNA has been demonstrated to regulate different types of genes (Stephenson and Zamecnik, 1978; Izant and Weintraub, 1984). In this respect, for the rice genome Rhoades et al. (2002) identified 34 miRNA regulatory loci that were associated with cell growth and differentiation. AtmiR156 has a positive regulatory effect on anthocyanin and upregulate the gene DFR by inducing and explaining the expression of SPL9 in Arabidopsis (Gou et al., 2011). miR828 in tomato can negatively regulate the biosynthesis of anthocyanin by acting on the target gene S1Myb7-lik (Jia et al., 2015). Therefore, in-depth studies on miRNA-target gene interactions in the regulation of peanut testa colors could be useful.

Nowadays, high-throughput omics methods, especially transcriptomics and metabolomics, have been widely used by botanists and deepen their comprehension of different biological pathways (Raza, 2020). Study investigated the differentially expressed genes (DEGs) of anthocyanin synthesis using transcriptome-metabolome joint analysis for the testa of both parents, and analyzed the miRNA-target gene interactions. The results of this study are expected to provide helpful insights into the cultivation of peanut varieties with concentrated anthocyanin. Additionally, this work is expected to be a reference for in-depth research on the metabolic mechanism of peanut testa anthocyanin.

MATERIALS AND METHODS

Sample Preparation

In this study, the purple testa peanut variety with high oleic acid (O: 79.52%; L: 5.48%; O/L: 14.51%), Z18-40, obtained from the cross of "G110 × Zizhenzhu." G110 with pink testa and high-oleic acid (O: 74.62%, L: 5.48%; O/L: 13.62) was used as maternal parent and Zizhenzhu with purple testa and normal oleic acid (O: 44.83%, L: 35.17%; O/L: 1.27) was used as paternal parent. Two kompetitive allele-specific PCR (KASP) markers, A004807 and A004808 were used to identify F_2 offspring and a total of 66 high-oleic acid progenies with genotype *aabb* were detested and continuously self-pollination up to F_7, resulting in superior lines namely Z18-40 (**Supplementary Figure 1**). Z18-40 was one of the new varieties both enrich in anthocyanin and high oleic acid. During the breeding process, different testa color existed between G110 and Z18-40. This phenomenon attracted us to detect the molecular mechanism of anthocyanin. G110 and Z18-40 were planted separately at Yixian station (Baoding, China) on May 2020, and were labeled with a tag after flowering between July and October. Samples consisting of 1–3 g of testa were taken on both 30 DAF and 45 DAF for G110 (G1, G2) and Z18-40 (Z1, Z2) before being wrapped in tin foil. They were then flash-frozen in liquid nitrogen for 5 s and stored at −20°C. Three biological replicates for each sample were sent to Gene Denovo Biotechnology Co. (Guangzhou, China) for sequencing.

Measurement of Differences in the Values of Testa Color

The a, b, and L values of testa at 30 DAF, 35 DAF, 40 DAF, 45 DAF, 50 DAF, and 55 DAF for the two varieties were measured with a colorimeter (CR-10Plus, Japan) using three biological replicates. One-way ANOVA was performed using SPSS26 and the size of the difference in color (ΔE) was calculated with 30 DAF as the reference according to the following equation:

$$\Delta E = \sqrt[2]{(\Delta L)^2 + (\Delta a)^2 + (\Delta b)^2}$$

ΔE values 0.00–0.25, 0.25–4.00, and >4.00 were considered to be an ideal, an acceptable and substantial differences, respectively. Peanuts were sectioned and each section was stained with 0.5% vanillin solution for 20 min. Changes in the testa color at 30 DAF and 45 DAF were then measured by stereomicroscopy (XDL-7000, China). In this case, the background was processed using Adobe Photoshop 2021.

Transcriptome Analysis

The testa of G110 and Z18-40 at 30 DAF and 45 DAF were collected. Testa samples were flash-frozen using liquid nitrogen and stored at -80°C until RNA extraction. RNA was extracted by Trizol precipitation with three biological replicates (Kingston et al., 2001). The purity and integrity of RNA were analyzed using NanoDrop ND-1000 UV/Vis spectrophotometer (Thermo

Fisher Scientific, Wilmington, DE, United States) and Agilent 2100 Bioanalyzer (Agilent, United States), respectively. Specific samples (RIN $\approx$ 10; 28S/18S $\geq$ 1.5; 1.7 < OD260/OD280 < 2.0) were selected for transcriptome sequencing and 12 libraries of complementary DNA (2 varieties × 2 periods × 3 replicates) were constructed using a five-step process prior to sequencing on the Illumina platform.[1] High-quality sequences were then compared with the tetraploid cultivar peanut reference genome[2] using the HISAT2 system (Kim et al., 2015).[3] The fragments per kilobase per million (FPKM) values of the two sets of samples were also recorded while the ratio of FPKM (Fold Change, FC) as well as the false discovery rate (FDR) for comparing differential expression were calculated using DESeq2 (Love et al., 2014), with thresholds of $|\log2FC| \geq 1$ and FDR < 0.05 being applied in order to screen for the differentially expressed genes (DEGs).

Metabolome Analysis

To evaluate the testa anthocyanin content of G110 and Z18-40, the method was used for content measurement and improvement as Wrolstad et al. (2001). Firstly, testa samples at 30 DAF and 45 DAF, frozen immediately in liquid nitrogen and stored at -70°C until use. Total anthocyanin content was measured using a spectrophotometric (UV-5800, Shanghai, China) method. In short, total anthocyanin was extracted using 1% HCl/methanol in the dark at 4°C overnight with occasional shaking. The extracts were centrifuged at 10,000 r for 10 min, and the supernatant was centrifuged at 12,000 r for 15 min and measured at 530 and 657 nm for absorbance determination. The equation A530 - 0.25 × A657 was used to compensate for the absorption of chlorophyll and its degradation products at 530 nm. Total anthocyanin content was calculated using the subtracted absorbance/fresh weight. The anthocyanin analysis of each sample was repeated three times using three independent biological replicates.

Liquid chromatography tandem mass spectrometry (LC-MS/MS) was used for the qualitative and quantitative determination of the different testa. After meeting the liquid phase and mass spectrometry conditions, methanol was used as a solvent and stored at -20°C. The samples were diluted with 70% methanol to yield different gradients before mass spectrometry, and were then extracted with methanol. The reagents were chromatographically pure ethanol and acetonitrile from Merck (United States) as well as standards (analytical purity) from BioBioPha or Sigma-Aldrich. The experiment was conducted with three biological replicates. The structural analysis of metabolites was carried out based on existing mass spectrometry public databases such as HMDB,[4] MoToDB,[5] and METLIN.[6] The variable importance in the projection (VIP) > 1 and $p < 0.05$ was the threshold in order to screen DAMs.

[1]http://www.illumina.com

[2]https://v1.legumefederation.org/data/public

[3]https://v1.legumefederation.org/data/public/Arachis_hypogaea/Tifrunner.gnm1.ann1.CCJH/

[4]http://www.hmdb.ca/

[5]http://appliedbioinformatics.wur.nl

[6]http://metlin.scripps.edu/index.php

Interaction of miRNAs-Target Gene

After Illumina sequencing, discard any low-quality reads, adaptors, contaminated sequences, and sequences shorter than 18 nt. Only the remaining high-quality sequences between 18 and 30 nt were further analyzed. All unique sequences were compared with the GeneBank tabase[7] to identify miRNAs. Known miRNAs were identified using a BLAST search against the miRNA database miRBase release 20.[8] Reads that were not annotated for any category were used to predict novel miRNAs using the miRNA prediction program MIREAP.[9] Calculate the expression amount of miRNA in each sample, and use the TPM (Transcripts Per Million) algorithm to normalize the expression amount. Differential expression analysis of miRNAs was performed using edgeR software (Robinson et al., 2010), with $|\log2FC| \geq 1$ and FDR < 0.05 selected as screening criteria to identify DEGs. Prediction and enrichment analysis of differentially expressed target genes by clusterProfiler among sample groups were analyzed, additionally, the enrichment degree of the pathways were analyzed by enrichment factor, and significance of enrichment was calculated by Fisher's exact test.

Quantitative Real-Time PCR Analysis

Analysis of Differentially Expressed Genes by Quantitative Real-Time PCR

Quantitative real-time PCR (qRT-PCR) was applied to verify the transcription levels of candidate genes. The same RNA samples in high-throughput sequencing were used for qRT-PCR. Each 10-μl qRT-PCR reaction mixture contained 1 μl of 10-fold diluted first-strand complementary DNA, 0.3 μl of each primer (10 μM), and 5-μl 2 × PowerUP^{TM} SYBR^{TM} Green Master Mix (Applied Biosystems, Carlsbad, CA, United States). A Bio-Rad CFX96 real-time PCR system was used under the following conditions: 50°C for 2 min, 95°C for 2 min, followed by 40 cycles of a denaturing step at 95°C for 10 s, an annealing step at 60°C for 20 s, and an extension step at 72°C for 45 s. For this purpose, primers were designed using Premier 5.0 software and the experiment was carried out with three biological as well as three technical replicates (**Supplementary Table 1**). The *ACT7* was used as an internal reference gene, to analyze the qRT-PCR results by $2^{-\Delta\Delta Ct}$ analysis (Livak and Schmittgen, 2001). *T*-tests were used to compare differences in gene expression.

Analysis of miRNAs by Quantitative Real-Time PCR

Quantitative real-time PCR was applied to verify the sequencing levels of candidate miRNAs. The same RNA samples in high-throughput sequencing were used for qRT-PCR. Firstly, the reverse transcription system was configured according to Tiangen kit instructions (miRcute Plus miRNA First-Strand cDNA Synthesis, Beijing, China) for synthesizing the first strand. The 20-μl reverse transcription system contained 10-μl 2 × miRNA RT Reaction Buffe, 2-μl miRNA RT Enzyme Mix, 1-μl Total RNA and 7-μl RNase-Free ddH_2O. The reverse transcription program was used under the following conditions: 42°C for 60 min and

[7]https://cipotato.org/genebankcip/

[8]http://mirbase.org/

[9]http://sourceforge.net/projects/mireap/

95°C for 2 min. The reverse primer used universal primer and the design and improvement of the forward primer followed the methods such as Chen et al. (2005) by using SnapGene 4.1.8. The 5 s rRNAs was used as an internal reference (**Supplementary Table 2**). Each 20-μl qRT-PCR reaction mixture contained 10-μl 2 × miRcute Plus miRNA and Premix (with SYBR&ROX), 0.3-μl Forward Primer (10 μM) and Reverse Primer (10 μM), 1-μl the first strand cDNA of miRNA and ddH_2O Until 20-μl. The detection was carried out according to the instructions of the fluorescence quantification kit (miRcute Plus miRNA qPCR Detection Kit, Beijing, China). The reaction program of qRT-PCR was designed under the following conditions: 40 cycles of qRT-PCR used a two step reaction which consisted of a pre-denaturation step at 95°C for 15 min, PCR cycle step at 94°C for 20 s, and an annealing and extension step at 60°C for 34 s. The analysis method of qRT-PCR results was the same as the above methods.

RESULTS

Analysis of Testa Color Differences

The color of Z18-40 plants was obviously darker than that of G110 plants (**Figure 1A**). The stems and pegs of G110 were of green color, while those of Z18-40 were purple-red. In the case of flowers, those of G110 were faint yellow in color, while those of Z18-40 were orange-red, with the red color being prominent in the flag and wing petals (**Figure 1B**). In terms of testa color, a visual change from light to dark was observed during all six periods. This was particularly obvious from microscopic observations which showed that G110 was pinkish white at 30 DAF before gradually changing to pinkish red at 45 DAF. Similarly, Z18-40 was purple-black at 30 DAF and turned purple-red at 45 DAF. The results of color measurements showed that the "a" values of G110 were 11.2–15.2 ($p < 0.05$), the "b" values were 12.1–16.6 ($p < 0.05$) and the "L" values were 31.2–41.7 ($p < 0.05$). The ΔE value changed from 2.34 to 10.89 during the whole pod growth period, and ΔE > 4 after 45 DAF. The "a" values of Z18-40 were 2.1–7.8 ($p < 0.05$), "b" values were 2.2–5.9 ($p < 0.05$) and "L" values were 0.1–6.7 ($p < 0.05$). The ΔE value changed from 0.47 to 20.54 during the whole pod growth period, and ΔE > 4 after 45 DAF (**Figures 1C,D**).

Transcriptome Analysis of Anthocyanin Between Two Peanut Varieties

To further explore the molecular mechanism of peanut testa pigment differences, 12 cDNA libraries were constructed for the testa at 30 DAF and 45 DAF. This generated 79.87 Gb of clean data, of which 5.91 Gb was obtained for each sample (92.4% of reads ≥ Q30). Sequence comparison between the reads of each variety and the reference genome indicated a similarity level of 84.1–96.1%. Furthermore, a total of 26,840 DEGs were identified in the G1_vs_G2, Z1_vs_Z2, G1_vs_Z1 and G2_vs_Z2 comparison groups, with 3,387, 4,643, 3,111, and 1,883 upregulated genes and 2,581, 3,277, 5,104, and 2,854 downregulated genes, respectively. These included the anthocyanin biosynthesis-related genes *AhPAL*, *AhCHS*, *AhCHI*, *Ah4CL*, *AhF3H*, *AhFLS*, *AhDFR*, *AhANR*, *AhLAR*, *AhLDOX* as well as the regulatory genes *AhbHLH* and *AhbZIP* genes (**Supplementary Figure 2**). Venn analysis further revealed that 455, 1,824, and 6,994 genes overlapped when during four-, three-, and two-group comparisons, respectively, while 1,093, 1,710, 2078, and 669 genes were unique to the above comparison groups (**Figure 2A**). In addition, results of KEGG analysis indicated that 119, 125, 124, and 122 pathways were enriched in the four comparison groups, and plant hormone signal transduction (ko04075), flavonoid biosynthesis (ko00941), isoflavonoid biosynthesis (ko00943), phenylalanine metabolism (ko00360) and phenylalanine, tyrosine and tryptophan biosynthesis (ko00400) were related to anthocyanin biosynthesis. The results of four comparisons indicated that there were 100, 66,123, and 126 DEGs enriched in plant hormone signal transduction and 20, 13, 34, and 32 DEGs enriched by flavonoid biosynthesis, which were the two key differential enrichment pathways in anthocyanin biosynthesis (**Figure 2B** and **Supplementary Table 3**). GO enrichment analysis resulted in two GO Classify1 (Biological process and Molecular function) and seven GO Classify2 [flavonoid biosynthetic process (GO:0009813), anthocyanin-containing compound biosynthetic process (GO:0009718), response to UV-B (GO:0010224), oxidation-reduction process (GO:0055114), proanthocyanidin biosynthetic process (GO:0010023), Iron ion binding (GO:0005506), and oxidoreductase activity (GO:0016706)] related to anthocyanin biosynthesis. oxidation-reduction processes were enriched with 359, 460, 429, and 244 DEGs for the four comparison groups. The enrichment factors for the proanthocyanidin biosynthetic process in the four groups (except for G2_vs_Z2) were 5.98, 4.41, and 5.75, which was the highest in GO Classify2 (**Figure 2C** and **Supplementary Table 4**).

Metabolome Analysis of Anthocyanin

The metabolome analysis of anthocyanin resulted in the detection of 12 DAMs in a total of 16 metabolites. Two, nine, nine and seven DAMs, with log2FC values of 0.7–1.0, -1.8 to 1.6, -1.5 to 15, and -0.9 to 7.7 were identified in G1_vs_G2, Z1_vs_Z2, G1_vs_Z1, and G2_vs_Z2, respectively. The procyanidin B3, B2, A2, and A1, were all upregulated in Z1_vs_Z2 but downregulated in G1_vs_Z1. DAMs were involved in regulating expression obviously in two comparison groups as follows: (1) In G1_vs_Z1, log2FC > 5 for peonidin 3-sophoroside-5-glucoside, cyanidin 3-O-galactoside, cyanidin 3-O- glucoside, and kuromanin were significantly upregulated; (2) In G2_vs_Z2, log2FC > 5 for cyanidin 3-O-galactoside, cyanidin 3-O-glucoside, and kuromanin were significantly upregulated (**Figure 3A**, **Supplementary Figure 3**, and **Supplementary Table 5**).

Two pathways for anthocyanin biosynthesis (ko00942) and flavonoid biosynthesis (ko00941) were obtained based on KEGG enrichment analysis. The total anthocyanin contents at 30 DAF and 45 DAF showed that the anthocyanin content of Z18-40 was 7.49–8.62-folds higher than that of G110 ($p = 1.843 \times 10^{-8}$), with highly significant differences. For the same variety, the differences were not significant ($p = 0.799$) between G1 and G2, while were highly significant ($p = 1.802 \times 10^{-5}$) between Z1 and Z2 (**Figure 3B** and **Supplementary Table 6**).

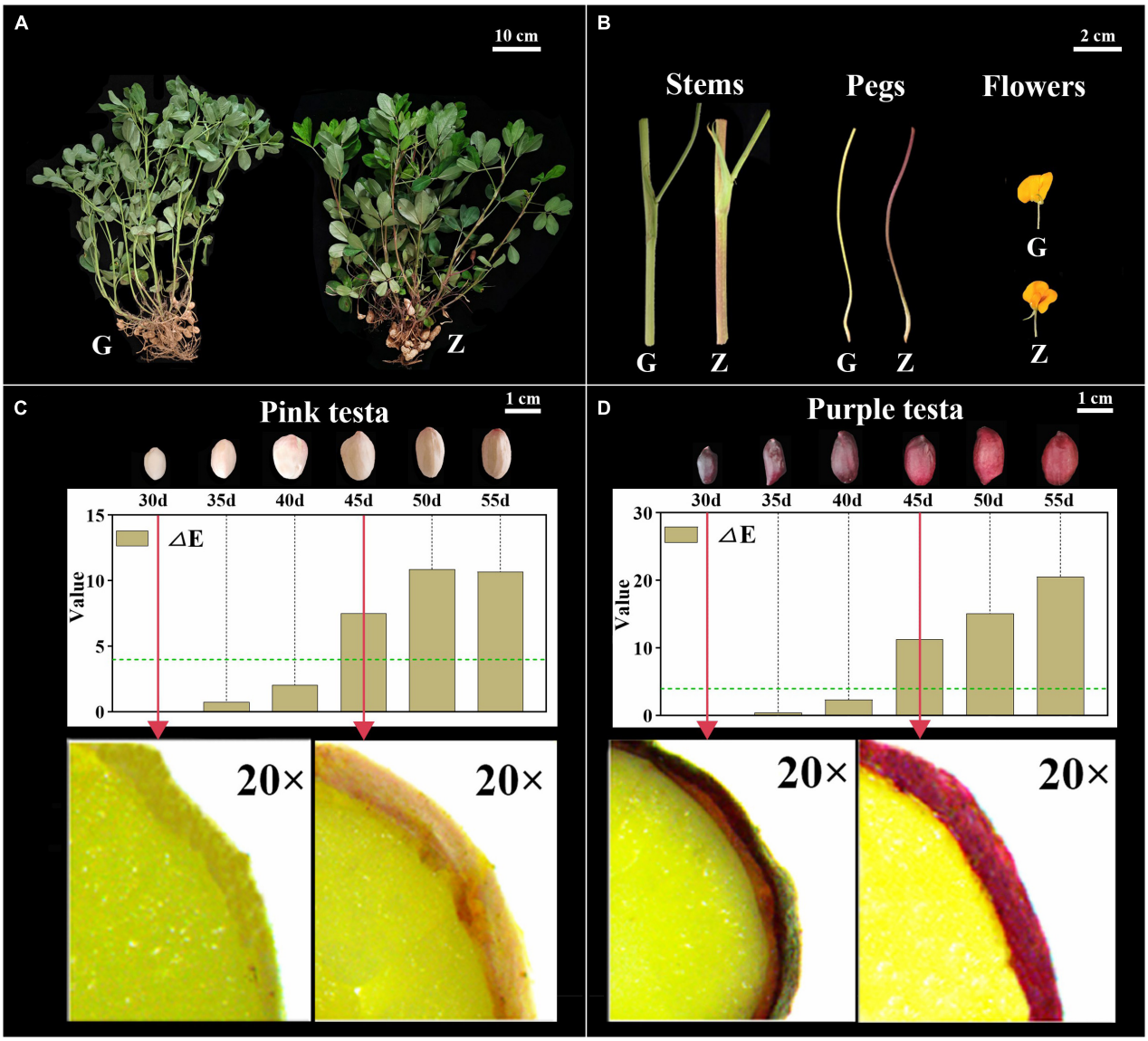


FIGURE 1 | Observation and Identification of botany phenotypes between G110 and Z18-40. Panel **(A)** shows the color difference comparison of the whole plants between G110 (G) and Z18-40 (Z). G110 plant is on the left, Z18-40 plant is on the right, and the scale bar is 10 cm. Panel **(B)** shows the color difference comparison of the stems, pegs and flowers between G and Z, and the scale bar is 2 cm. Panels **(C,D)** show the testa color changing trend between G and Z in six growth periods, and the scale bar is 1 cm. In addition, "ΔE" value of G and Z testa is determined in six periods and the green dotted line means ΔE = 4. The red arrows point to the testa of 30 DAF and 45 DAF in G and Z, respectively, which is observed with a stereo microscope on 20-folds magnification.

Transcriptome and Metabolome Joint Analyses of Anthocyanin

Based on the DEGs and DAMs of the four comparison groups obtained by transcriptome and metabolome analysis, correlation analysis was carried out in the enrichment pathway of flavonoid biosynthesis (ko00941). In. G1_vs_G2, differential gene expression was correlated with peonidin 3-O-glucoside chloride (pmf0203) as well as cyanidin 3-O-glucoside (Kuromanin, pmb0550), with 25 DEGs, 12 positive and 13 negative correlations. For the Z1_vs_Z2 comparison, differential gene expression was associated with cyanidin 3-O-glucoside (Kuromanin, pmb0550) and delphinidin (pme0442), and involved 33 DEGs, of which 15 were positively and 18 were negatively associated. In G1_vs_Z1, differential gene expression was associated with cyanidin 3-O-glucoside (Kuromanin, pmb0550), procyanidin A1 (pme0430), procyanidin B2 (pme0434), and delphinidin (pme0442). For this group, 24 DEGs were associated, with 6 being positively and 18 being negatively associated. Finally, in G2_vs_Z2, differential gene expression was associated with cyanidin 3-O-glucoside (Kuromanin, pmb0550) and was associated with 29 DEGs, with 10 showing positive and 19 showing negative associations (**Supplementary Table 7**). Lastly, we filtered 14 key candidate genes in different

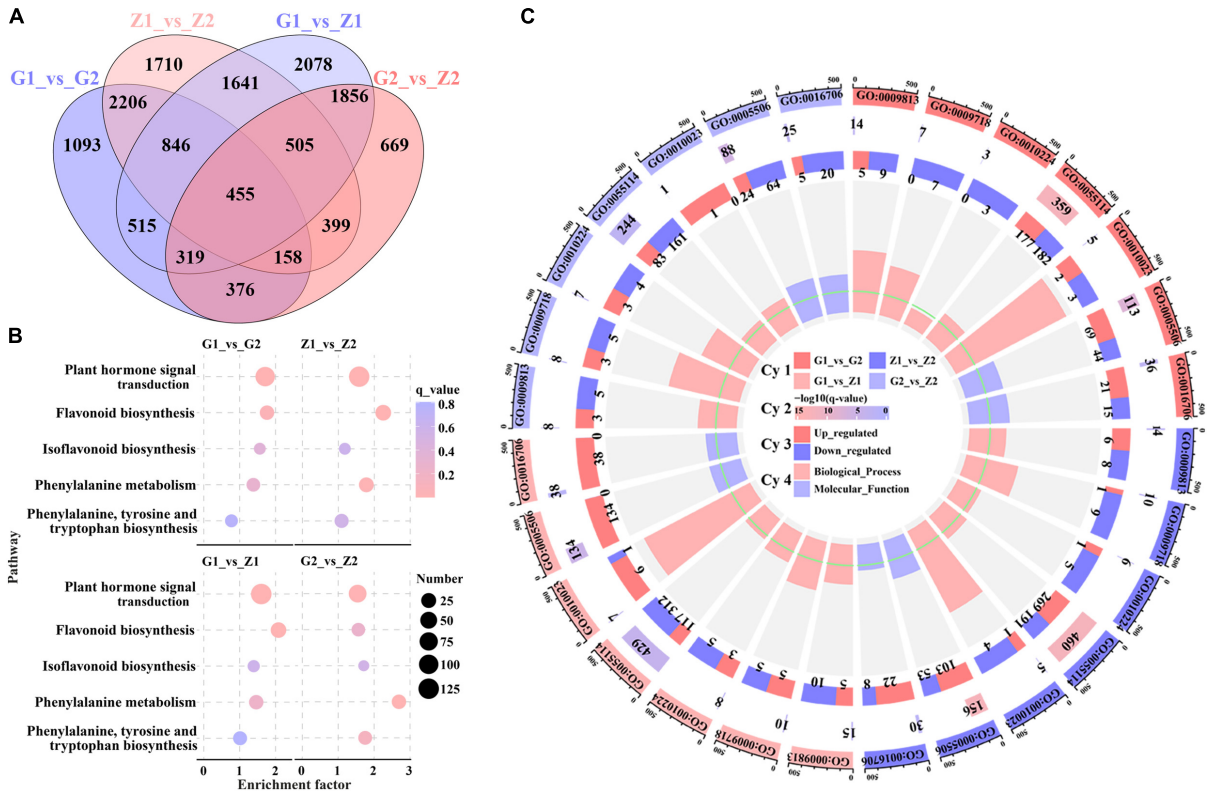


FIGURE 2 | Screening and enrichment analysis of anthocyanin DEGs. Venn diagrams **(A)** displaying overlap between DEGs identified in peanut testa in G1_vs_G2, Z1_vs_Z2, G1_vs_Z1 and G2_vs_Z2. The number on the part without overlapping colors represents the DEGs in a single comparison group. The number on the overlapping color part indicates the DEGs shared among different comparison groups. Picture **(B)** shows KEGG pathway enrichment analysis among the above four comparison groups, in which the abscissa means the enrichment factor, the ordinate means the enrichment pathways, the size of the circles means the number of DEGs, and the color of each circle means q_value. Picture **(C)** is made up of four cycles. Cycle 1 (Cy1, the outermost cycle) indicates 7 GO Classify2 localized to the four comparison groups respectively. Cycle 2 (Cy2, the secondary outer cycle) means the number of DEGs corresponding to the GO Classify2. The color of the rectangles below the number indicates the q value (-log10). Cycle 3 (Cy3, the sub-internal cycle) means the number of upregulated and downregulated genes corresponding to GO Classify2. Cycle 4 (Cy4, the innermost cycle) means the enrichment scope of GO Classify2, which include biological process and molecular function. The green circular line means the enrichment factor = 1.

comparisons and determined their chromosomal (Chr) position, *Ah21440* (novel Gene), *AhCHS* (arahy. UDJX6I, 8633464-8632413), *AhCHI* (arahy. TJ3PHW, 152096417-152097158), *AhPAL* (arahy. EEZ4Y8, 391227-389497), *Ah4CL* (arahy. X2F5F9, 2943040-2944560), *AhC-CoA* (arahy. JB63H4, 15250522-15251181), *AhF3H* (arahy. 79B99S, 11007614-11007083), *AhFLS* (arahy. 4WXU8P, 19358597-19357631), *AhDFR* (arahy. 7JZ58T, 154671103-154670569), *AhLAR* (arahy. T1J2UZ, 110006213-110006838), *AhANR* (arahy. W8TDEC, 126128334-126128550), *AhLDOX* (arahy. UQ0Z3E, 13645491-13647176), *AhbHLH* (arahy. JV9T2X, 4730198-4731619) and *AhbZIP* (arahy. R4ID1P, 74478269-74477851), associated with anthocyanin biosynthesis ($R^2 \geq 0.80$) (**Figure 4**).

Interaction of miRNAs and Target Genes and Screening of Key Candidate Genes

A total of 131.09 Gb clean tags were obtained after miRNA sequencing of the four peanut testa samples and through comparison with the miRBase database, 28 (0.54%) existing miRNAs, 430 (9.35%) known miRNAs and 407 (0.43%) novel miRNAs were identified. A total of 531 differentially expressed miRNAs were also screened for the four comparison groups, with 20, 101, 93, and 63 upregulated and 22, 115, 59, and 58 downregulated expression levels, respectively. Moreover, 4,174, 10,047, 5,037, and 5,791 target genes were predicted for the four groups. Additionally, the results of KEGG enrichment analysis indicated that 99, 129, 110, and 126 metabolic pathways could be identified for the four comparison groups. Among these, five pathways were related to anthocyanin biosynthesis and they included 30, 25, 24, 19, and 32 differentially expressed target genes (**Figure 5A**). The analysis of differentially expressed target genes and miRNAs interactions further showed that 17 miRNAs were localized within the comparison groups, with 1, 11, 4, and 6 upregulated gene expression and 2, 1, 3, and 2 downregulated gene expression, respectively (**Figure 5B**). Correlation analysis of target genes and DEGs in the transcriptome shows that the *Ah21440* (novel Gene), *AhCHS* (arahy. UDJX6I), and *AhCHI* (arahy. TJ3PHW), which were related to anthocyanin biosynthesis, were

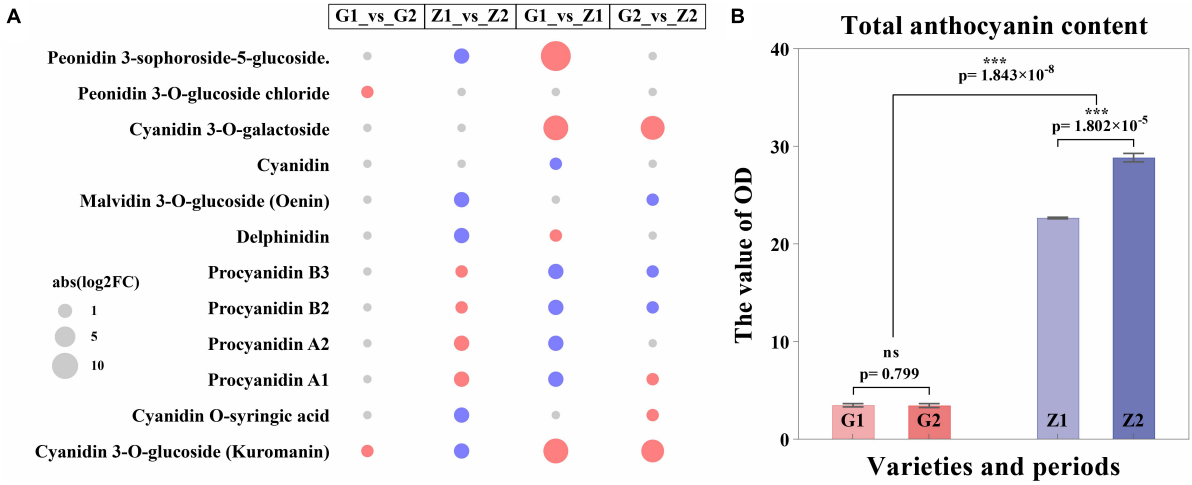


FIGURE 3 | Expression regulation of DAMs and differential analysis of anthocyanin content. Panel **(A)** shows the expression regulation of DAMs, in which the abscissa means the comparison groups, the ordinate means DAMs, the size of the circle means abs (log2FC), the red means upregulated expression, and the blue means downregulated expression. Panel **(B)** shows differential analysis of anthocyanin content, where the abscissa means varieties and periods. G1 and G2 means 30 DAF and 45 DAF of G110, respectively, Z1 and Z2 means 30 DAF and 45 DAF of Z18-40, respectively. The ordinate means the OD value (OD/g). "ns" means no significant difference ($p > 0.05$) and "****" means significant difference ($p < 0.001$).

regulated by *AhmiR2950*, *AhmiR398*, *AhmiR50*, and *AhmiR51* respectively, and were identified as the major miRNAs involved in anthocyanin regulation (**Figure 5C** and **Supplementary Table 8**).

Verification of Candidate Genes and miRNAs by Quantitative Real-Time PCR

Fourteen DEGs related to anthocyanin biosynthesis were screened through multi-omics and miRNA joint analysis results and validated by qRT-PCR. In G1_vs_G2 comparisons, nine DEGs were validated, including four upregulated genes and five downregulated genes. Two of the four upregulated genes and three of the five downregulated genes showed significant differences in expression. In Z1_vs_Z2, ten DEGs, including four upregulated and six downregulated genes, were validated. Three of the four upregulated genes and four of the six downregulated genes showed significant differences in expression. In G1_vs_Z1, nine DEGs, including four were upregulated genes and five were downregulated genes. Four upregulated genes and one of the five downregulated genes showed significant differences in expression. Finally, in G2_vs_Z2, 10 DEGs, including three upregulated genes and three in seven downregulated genes, were validated. Two of the three upregulated genes and three of the seven downregulated genes showed significant differences in expression. The key candidate genes, *Ah21440* (novel Gene), *AhCHS* (arahy. UDJX6I), *AhCHI* (arahy. TJ3PHW), *AhPAL* (arahy. EEZ4Y8), *Ah4CL* (arahy. X2F5F9), *AhF3H* (arahy. 79B99S), *AhFLS* (arahy. 4WXU8P), *AhDFR* (arahy. 7JZ58T), *AhLAR* (arahy. T1J2UZ), *AhANR* (arahy. W8TDEC), *AhLDOX* (arahy. UQ0Z3E) related to the anthocyanin biosynthesis process, were all validated and significantly different for the four comparison groups and these trends were consistent with the transcriptome data (**Figure 6A** and **Supplementary Table 9**).

Four miRNAs related to target genes interaction were screened through interaction of miRNA and target genes and validated by qRT-PCR. In G1_vs_G2 comparisons, *AhmiR2950* was downregulated. In Z1_vs_Z2, *AhmiR2950* was upregulated. In G1_vs_Z1, *AhmiR50*, *AhmiR51*, and *AhmiR398* were upregulated, and *AhmiR2950* was downregulated. In G2_vs_Z2, *AhmiR50*, *AhmiR51*, and *AhmiR398* were all upregulated. These trends were consistent with the miRNA sequencing data (**Figure 6B** and **Supplementary Table 10**).

DISCUSSION

Advantages of Research Materials

Two peanut varieties with different level of anthocyanin were selected, and pink testa (3.62 OD/g) was used as the maternal parent while purple testa (25.99 OD/g) was used as the paternal parent. These characteristics were selected so that purple testa could be distinguished as a true hybrid based on testa color in F_2 populations, thus improving the accuracy and efficiency of hybrid selection. Similar cross combinations were used in the study of Chen et al. (2021). At the same time, stable high oleic acid homozygous individual plants could be selected from F_2 by way of KASP. This method is more practical than SSR (Zhang et al., 2014) and AS-PCR (Li, 2015). The pink testa of the female parent G110 and the superior Z18-40 purple testa of F_7 were analyzed. Simultaneously, the testa color of the two varieties was different and were adopted to explore the molecular mode of anthocyanins in this research.

Testa Color-Based Identification

The botany phenotypic results showed that the testa color gradually darkened with the developmental period. Purple testa

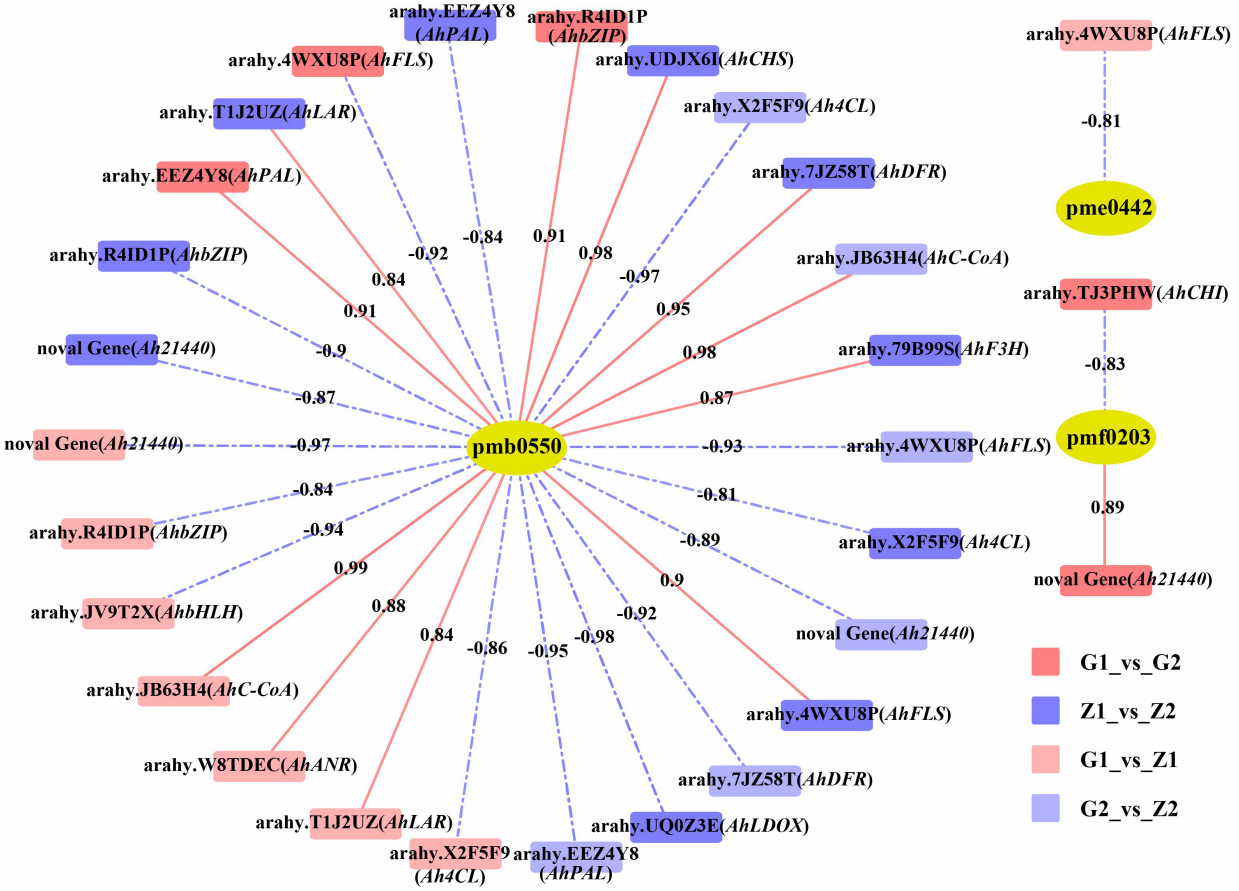


FIGURE 4 | Correlation analysis between DAMs and candidate DEGs. The pmb0550, pme0442, and pmf0203 above the yellow ellipses indicate cyanidin 3-O-glucoside (Kuromanin), delphinidin and peonidin 3-O-glucoside chloride, respectively. DEGs are labeled above the rectangles. Rectangles with red, blue, light red and light blue indicate G1_vs_G2, Z1_vs_Z2, G1_vs_Z1, and G2_vs_Z2, respectively. Red solid lines and blue dashed lines indicate positive correlation and negative correlation between DEGs and DAMs, on which are the correlation value.

showed purple-red color expressed on 45 DAF, and this color change was observed, in whole plants, stems, fruit needles, and flowers. Obviously, "L" value was inversely related to anthocyanin content (Zhang et al., 2020). The results hint that the metabolic mechanism of anthocyanin in varieties with high anthocyanin content may impact all parts of the plant. The $\Delta E > 4$ after 45 DAF showed significant color differences, thereby suggesting that 45 DAF was the key period for testa color change. Li et al. (2017) explicitly suggested that the period of peak color as a result of gene expression occurred between 40 DAF and 45 DAF, which is consistent with our results.

Regulation of Structural Candidate Genes and Metabolites

PAL is the first structural gene in anthocyanin biosynthesis, with Bar-Akiva et al. (2010) demonstrating that PAL activity can promote the formation of anthocyanins. In this study, although G 110 was higher than Z18-40 in *AhPAL* at both 30 DAF and 45 DAF, no significant differences were observed in the later developmental period, thus implying that *AhPAL* may be the first key DEG for color changes in pink testa. In a study involving pink testa peanuts, Wan et al. (2018) demonstrated that the expression of *CHS* increased as the reproductive period progressed. In this study, *AhCHS* continued to be highly expressed in G110, but was only observed for the first time in Z1 at 30DAF. It was suggested that *AhCHS* could be the first key DEG for early color differentiation in peanut testa, with these results being consistent with previous studies (Hu et al., 2021a). It was also found that the expression level of *CHI* was positively correlated with anthocyanin content (Liu et al., 2019), and that *CHI* activity could increase the flavonoid content (Li et al., 2006). Additionally, even though *AhCHI* was consistently highly expressed in Z1, it was first highly expressed in Z2. The result hints that the stage from *AhCHS* to *AhCHI* is a critical period for purple testa pigment accumulation.

F3H is a relay gene in anthocyanin biosynthesis, whose high expression can result in orange-red petals in dahlias (Enrique et al., 2020). In this study, *AhF3H* was highly expressed in Z1, and it led to the accumulation of the substrates of DHK, DHM, and DHQ from the *AhDFR* and *AhFLS* genes, thus forming a large amount of anthocyanin while deepening testa

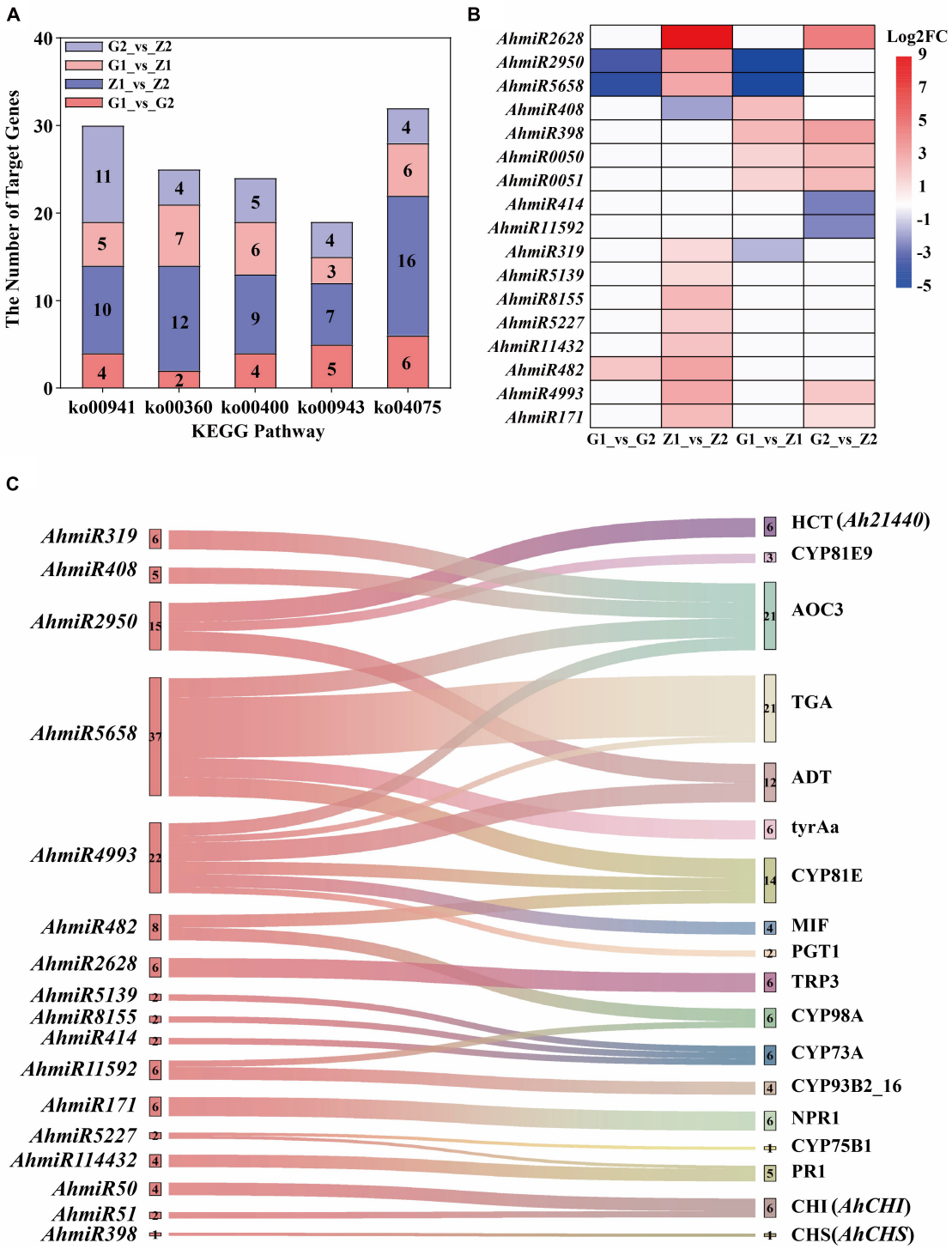


FIGURE 5 | Interaction analysis of miRNAs and target genes. The five bar graphs in panel **(A)** are composed of red, blue, light red, and light blue components, representing G1_vs_G2, Z1_vs_Z2, G1_vs_Z1, and G2_vs_Z2, on which is the number of the target genes, respectively. The abscissa means the KEGG pathways, named flavonoid biosynthesis (ko00941), phenylalanine metabolism (ko00360), tryptophan biosynthesis (ko00400), isoflavonoid biosynthesis (ko00943), and plant hormone signal transduction (ko04075). The ordinate means the number of target genes. Panel **(B)** shows that the regulation of differentially expressed miRNAs in the above four comparison groups. The abscissa means the comparison groups, and the ordinate means the differentially expressed miRNAs. The red means upregulated expression, the blue means downregulated expression, and the white means that there is no differential expression. Panel **(C)** shows the interaction of miRNA-target genes. The left is the differentially expressed miRNAs and the number beside miRNAs means the number of corresponding target genes. The right is the ko_names of the target genes, and the number beside ko_names means the number of miRNAs.

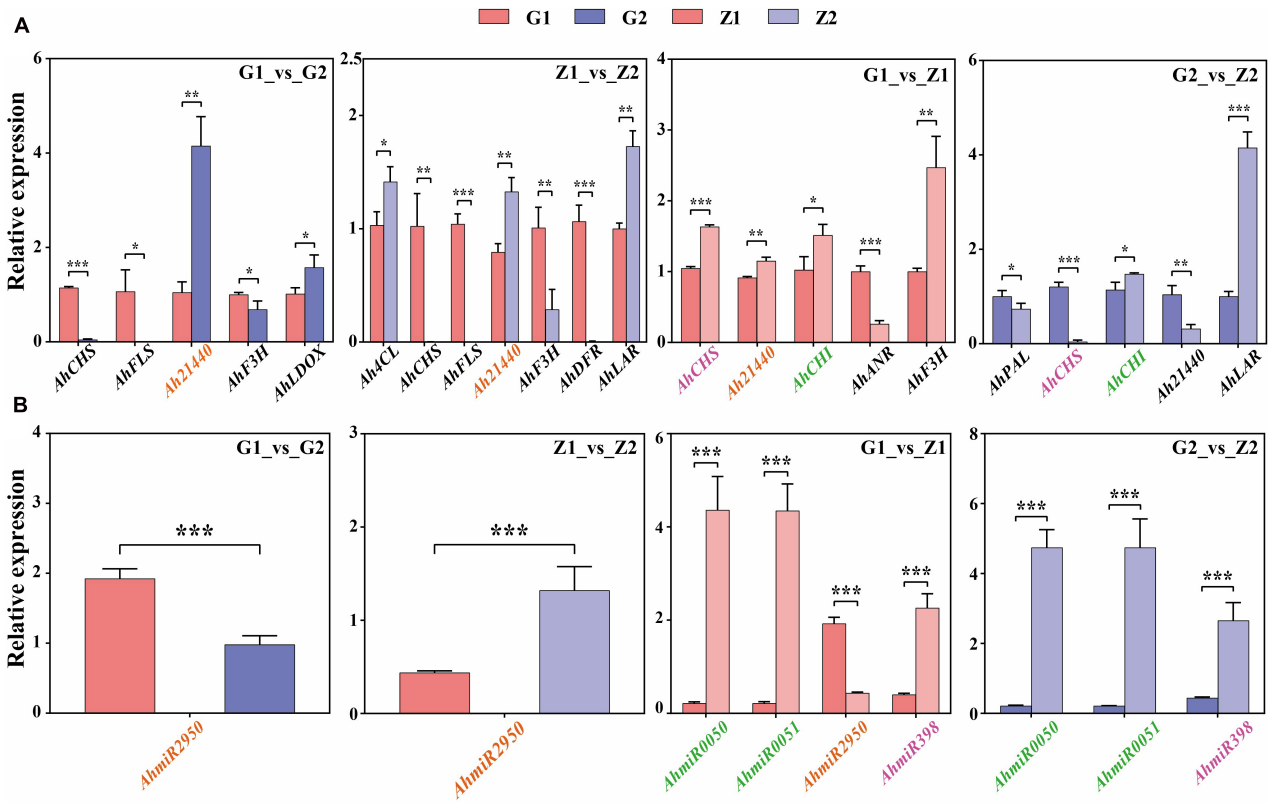


FIGURE 6 | Verification of candidate DEGs and miRNAs by qRT-PCR. Panel **(A)** shows that the verification of candidate DEGs by qRT-PCR in G1_vs_G2, Z1_vs_Z2, G1_vs_Z1, and G2_vs_Z2, respectively. The abscissa means the candidate DEGs and the ordinate means the relative expression. "*" indicates $p < 0.05$, "**" indicates $p < 0.01$, and "****" indicates $p < 0.001$. Panel **(B)** shows that the verification of differentially expressed miRNAs by qRT-PCR. The abscissa indicates the candidate miRNAs, the ordinate indicates the relative expression. "*" indicates $p < 0.05$, "**" indicates $p < 0.01$, and "****" indicates $p < 0.001$. The target genes in panel **(A)** that have an interaction relationship with the miRNAs in panel **(B)** are marked with the same color.

color. This was consistent with the results on the phenotypes of Z1. In G1_vs_G2, *AhF3H* was downregulated, and it suggested that the accumulation of substances for pink testa occurred at an early stage. Regarding *FLS* and *DFR*, they represent two competing genes which lead to the flavonol and anthocyanin pathways, respectively (Lim et al., 2016). In this study, *AhFLS* was downregulated in G1_vs_G2, while *AhDFR* was not differentially expressed. These results clearly indicated that, in the downstream pathway, *AhFLS* catalyzed large amounts of substrates within the flavonol pathway, while only few substrates were converted by *AhDFR*, hence resulting in less anthocyanin accumulation. Similarly, *AhFLS* was downregulated in Z1_vs_Z2, but in this case, *AhDFR* was upregulated, causing substantial substrate shifts to the flavonol pathway in Z1. As a result of the high substrate accumulation of *AhF3H*, Z1 could accumulate anthocyanin but the combined higher expression of both *AhDFR* and *AhF3H* allowed more anthocyanin to be accumulated, with these results being consistent with those for total anthocyanin content and metabolome analysis.

Cui et al. (2016) reported that the upregulation of *LDOX* could increase the anthocyanin content in purple podzolic lentils. In this study, *AhLDOX* was highly expressed in G2 and Z1, and it further catalyzed the formation of anthocyanin from the substrates in G2 and Z1, resulting in the deepening of testa color. This indicated that *AhLODX* was a key gene determining the anthocyanin content as indicated by the phenotypic results. *ANR* and *LAR* are also key genes for the formation of proanthocyanin as they negatively regulate proanthocyanin synthesis (Albert et al., 1997). Additionally, the high expression of *AhLAR* in Z2 increased its proanthocyanin content as reflected in the results of metabolome analysis. In this study, *AhANR* was downregulated in G1_vs_Z1 as well as in G2_vs_Z2, resulting in a decrease of proanthocyanin content in Z1. Combined with the metabolome analysis where both proanthocyanin B2 and B3 were downregulated and proanthocyanin A1 was upregulated in G2_vs_Z2, it can be suggested that *AhANR* could be a key gene that regulates proanthocyanin B2 and B3. Based on the metabolic data, the purple testa at 45 DAF, identified as purple-red by the phenotype identification, could be ascribed to the regulation of proanthocyanin. Proanthocyanin is reddish brown in color, but it can darken the testa color (Li, 2014). On the other hand, the low abundance of blue delphinidin-like substances in Z2 further deepens the red color of testa but these differences need to be further investigated.

In this study, *Ah21440* was a functional novel gene, which catalyzed the transfer of hydroxycinnamoyl transferase (HCT)

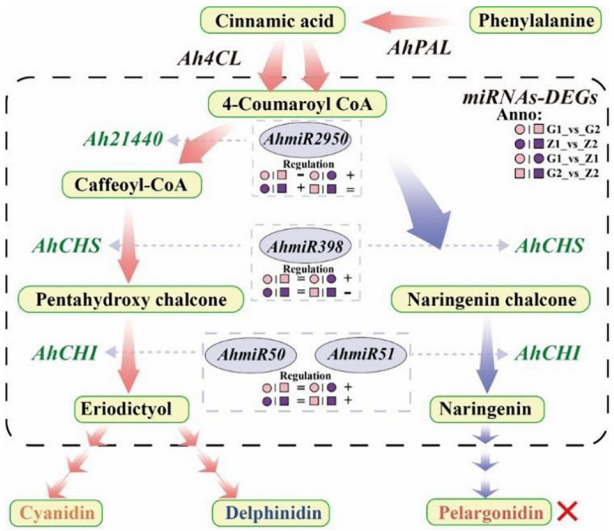


FIGURE 7 | Flow-process diagram of anthocyanin biosynthesis and miRNA interaction with target genes. The picture shows that rectangles with light yellow shading mean metabolites, light red arrows between metabolites mean preferred pathways, and light blue arrows indicate non-preferred pathways yet. The dashed rectangle indicates miRNAs-DEGs interaction. The green labels beside the arrows mean DEGs, the miRNAs corresponding to DEGs are marked in a gray ellipse, and both of them are connected by light gray arrows, and the regulation is indicated below the miRNAs. "∘" with pink and purple means 30DAF of G110 and Z18-40, respectively. "□" with pink and purple means 45DAF of G110 and Z18-40, respectively. "+," "-," and "=" indicate that there is a positive, negative and none regulatory relationship between DEGs and miRNAs in the comparison on the left. The red cross indicates that there are no regulation expression in the metabolites or genes.

to form anthocyanin (Sun et al., 2018). Transcriptome and qRT-PCR results showed that *Ah21440* was upregulated in G1_vs_G2 and Z1_vs_Z2, whilst being downregulated in G1_vs_Z1 and G2_vs_Z2. These results indicated that *Ah21440* continued to be highly expressed as the reproductive period progressed, and that both pink and purple testa favored the formation of caffeoyl coenzyme A before the formation of delphinium and cyanidin. However, no pelargonidin was identified and this was consistent with the metabolome data. This observation may be similar to the substrate specificity of *DFR* and the pathway preference of *HCT* (Yoshida et al., 2010), which needs to be further investigated (**Figure 7**).

Interaction Between miRNA and Target Genes

Anthocyanin contribute to resistance to abiotic stress (Valerio et al., 2021). In particular, *ACO3* and *ADT* are involved in responses to various abiotic stresses, while *NPR1* can promote the binding of *TAG* and *PR1*, thus improving plant stress resistance (Johnson et al., 2003). In this study, all nine miRNAs interacted with the target genes to varying degrees and therefore, the results indicated that miRNAs were capable of regulating anthocyanin-based resistance to abiotic stress. The miRNAs involved in the regulation of purple testa were significantly richer than those of pink testa. It was suggested that a potentially positive correlation between testa color and resistance to abiotic stresses. Furthermore, cytochrome P450 could catalyze the synthesis and metabolic reactions of terpenoids, fatty acids, flavonoids and isoflavones (Schuler and Werck-Reichhart, 2003). Additionally, 34 *AhCYP* genes were identified as being involved for catalyzing the hydroxylase and synthase of anthocyanin biosynthesis, and their corresponding nine miRNAs were involved in the regulation of the process in varying degrees. It is likely that miRNAs could regulate the enzymatic functions of anthocyanin biosynthesis. In fact, *AhmiR2950*, *AhmiR5658*, *AhmiR4993*, and *AhmiR5227* regulate both the anti-biotic stress function and the biosynthetic enzyme function of anthocyanin, indicating that miRNAs also have a multi-causal function.

It was shown that *miR398* was involved in regulation through *AP2* and *SPL3* target genes (Li et al., 2013), while miR398_x was involved in the regulation of *F3′H* target genes (Hu et al., 2021b). In this study, *AhmiR2950* regulated *Ah21440*, *AhmiR398* regulated *AhCHS* while the *AhmiR50* and *AhmiR51* regulated *AhCHI*. Combined with the transcriptomes data, it was found that *AhmiR2950* is negatively regulated in peanuts with low anthocyanin content, and positively regulated in peanuts with high anthocyanin content and the regulation period tends to be early in development. Additionally, *AhmiR398* was positively regulated at 30 DAF, but negatively regulated at 45 DAF. Both *AhmiR50* and *AhmiR51* were positively regulated on 30 DAF and 45 DAF. These results indicated that the regulatory pattern of miRNAs may vary depending on the variety and the progression of the reproductive period (**Figure 7**).

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: https://www.ncbi.nlm.nih.gov/, PRJNA773958.

AUTHOR CONTRIBUTIONS

GM and XY developed the concept, planned, coordinated, and executed the research. JL wrote the manuscript. JL and YM performed most of the experiments and analyzed the data. YZ, BL, CW, MZ, and LZ prepared the **Figures 1–7**. GM, XY, and MH revised the manuscript. All authors read and approved the final manuscript.

FUNDING

This research was funded by Key Project of Science and Technology Research in Colleges and Universities of the Department of Education in Hebei Province (ZD2019051), Key Project of Science and Technology Research of Modern Seed Industry of the Department of S&T in Hebei Province (19226363D), Project of Agricultural Science and Technology Park Construction in Baoding City, Hebei Province (2111N004),

the National Natural Science Foundation of China (31701459), and the Support Program for the Top Young Talents of Hebei Province (0602015).

ACKNOWLEDGMENTS

We are grateful to all members of Peanut Breeding Laboratory of Hebei Agricultural University, Baoding, China, for their efforts in anthocyanin detection and qRT-PCR. Special thanks to the staff of Baoding Yiyuan Ecological Agriculture Technology Co., Ltd., Baoding, China, for the labor in planting management of experimental materials throughout the growth period.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2022.818345/full#supplementary-material

REFERENCES

Albert, S., Delseny, M., and Devic, M. (1997). BANYULS, a novel negative regulator of flavonoid biosynthesis in the Arabidopsis seed coat. *Plant J.* 11, 289–299. doi: 10.1046/j.1365-313x.1997.11020289.x

Bar-Akiva, A., Ovadia, R., Rogachev, I., Bar-Or, C., Bar, E., Freiman, Z., et al. (2010). Metabolic networking in *Brunfelsia calycina* petals after flower opening. *J. Exp. Bot.* 61, 1393–1403. doi: 10.1093/jxb/erq008

Chen, C. F., Ridzon, D. A., Broomer, A. J., Zhou, Z. H., Lee, D. H., Nguyen, J. T., et al. (2005). Real-time quantification of microRNAs by stem-loop RT-PCR. *Nucl. Acids Res.* 33:e179. doi: 10.1093/nar/gni178

Chen, H., Chen, X., Xu, R., Liu, W., Liu, N., Huang, L., et al. (2021). Fine-mapping and gene candidate analysis for AhRt1, a major dominant locus responsible for testa color in cultivated peanut. *Theor. Appl. Genet.* 134, 1–10. doi: 10.1007/s00122-021-03924-w

Cui, B. L., Hu, Z. L., Zhang, Y. J., Hu, J. T., Yin, W. C., Feng, Y., et al. (2016). Anthocyanins and flavonols are responsible for purple color of *Lablab purpureus* (L.) Sweet pods. *Plant Physiol. Biochem.* 103, 183–190. doi: 10.1016/j.plaphy.2016.03.011

Deng, S. F. (2018). *Studies on the Molecular Mechanism of Peel Color Diference Formation Based on RNA-Seq in the Wild Banana(Musa itinerans)*. Master's thesis. China: Fujian Agriculture and Forestry University.

Duan, H. R., Wang, L. R., Cui, G. X., Zhou, X. H., Duan, X. R., and Yang, H. S. (2020). Identification of the regulatory networks and hub genes controlling alfalfa floral pigmentation variation using RNA-sequencing analysis. *BMC Plant Biol.* 20:1205–1361. doi: 10.1186/s12870-020-2322-9

Enrique, G. A., Ana, G., Antonia, G., Ana, G. V., Javier, G. M. F., and Beatriz, R. S. (2020). Elicitation with Bacillus QV15 reveals a pivotal role of *F3H* on flavonoid metabolism improving adaptation to biotic stress in blackberry. *PLoS One* 15:e0232626. doi: 10.1371/journal.pone.0232626

Gao, G. Y., Wu, X. F., Zhang, D. W., Zhou, D. G., Zhang, K. X., and Yan, M. L. (2020). Research progress on the MBW complexes in plant anthocyanin biosynthesis pathway. *Biotechnol. Bull.* 36, 126–134.

Gou, J., Felippes, F., Liu, C., Weigel, D., and Wang, J. (2011). Negative regulation of anthocyanin biosynthesis in *Arabidopsis* by a miR156-targeted SPL transcription factor. *Plant Cell* 23, 1512–1522. doi: 10.1105/tpc.111.084525

Hu, M. D., Li, J. W., Cui, S. L., Hou, M. Y., Yang, X. L., Liu, L. F., et al. (2021a). Transcriptomic-metabolomics joint analysis of anthocyanin synthesis for variegated testa in peanut (*Arachis hypogaea* L.). *J. Plant Genet. Resour.* 22, 1732–1745.

Hu, M. D., Li, J. W., Wang, M., Liu, B., Ma, Y. C., Zhao, Y. L., et al. (2021b). Identification and analysis of miRNA and its target genes related to anthocyanin biosynthesis in variegated testa of peanut. *J. Plant Genet. Resour.* 23, 226–239.

Huang, J. Y., Xing, M. H., Li, Y., Cheng, F., Gu, H. H., Yue, C. P., et al. (2019). Comparative transcriptome analysis of the skin-specific accumulation of anthocyanins in black peanut (*Arachis hypogaea* L.). *J. Agricult. Food Chem.* 67, 1312–1324. doi: 10.1021/acs.jafc.8b05915

Izant, J. G., and Weintraub, H. (1984). Inhibition of thymidine kinase gene expression by anti-sense RNA: a molecular approach to genetic analysis. *Cell* 36, 1007–1015. doi: 10.1016/0092-8674(84)90050-3

Jaakola, L. (2013). New insights into the regulation of anthocyanin biosynthesis in fruits. *Trends Plant Sci.* 18, 477–483. doi: 10.1016/j.tplants.2013.06.003

Jia, X. Y., Liu, H., Shen, J., Li, F., Ding, N., Sun, Y., et al. (2015). Micro RNA828 negatively regulates tomato anthocyanin biosynthesis induced by phosphorus deficiency stress. *Sci. Agricult. Sin.* 48, 2911–2924.

Johnson, C., Boden, E., and Arias, J. (2003). Salicylic Acid and NPR1 induce the recruitment of trans-activating TGA factors to a defense gene promoter in Arabidopsis. *Plant Cell* 15, 1846–1858. doi: 10.1105/tpc.012211

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

Kingston, R. E., Chomczynski, P., and Sacchi, N. (2001). Guanidine methods for total RNA preparation. *Curr. Protoc. Mol. Biol.* 14, 4.2.1–4.2.9. doi: 10.1002/J.1934-3647.1991.TB00205.X

Li, F., Jin, Z., Qu, W., Zhao, D., and Ma, F. (2006). Cloning of a cDNA encoding the Saussurea medusa chalcone isomerase and its expression in transgenic tobacco. *Plant Physiol. Biochem.* 44, 455–461. doi: 10.1016/j.plaphy.2006.08.006

Li, H. F., Qiu, J. M., Chen, X. P., Hong, Y. B., and Liang, X. Q. (2017). Differential expression of genes associated with anthocyanin synthesis in peanut cultivars with different testa color. *Chin. J. Oil Crops* 39, 600–605.

Li, H. F., Qiu, J. M., Chen, X. P., and Liang, X. Q. (2018). Clone and sequence analysis of genes associated with anthocyanin synthesis in peanut cultivars with different testa color. *J. Trop. Crops* 39, 93–99.

Li, L. (2015). *Creation of Peanut Germplasm with High Oleic Acid and Genotype Effect Analysis of ah FAD2A*. Master's thesis. China: Agricultural University of Hebei.

Li, L., Zhai, Y. H., Luo, X. N., Zhang, Y., and Shi, Q. Q. (2019). Comparative transcriptome analyses reveal genes related to pigmentation in the petals of red and white *Primula vulgaris* cultivars. *Physiol. Mol. Biol. Plants* 25, 1029–1041. doi: 10.1007/s12298-019-00664-6

Li, S. B., Liu, L., Zhuang, X. H., Yu, Y., Liu, X. G., Cui, X., et al. (2013). MicroRNAs inhibit the translation of target mRNAs on the endoplasmic reticulum in Arabidopsis. *Cell* 153, 562–574. doi: 10.1016/j.cell.2013.04.005

Li, W. P. (2014). *Research on the Tea Excellent Germplasm Resources of High Content of PC, Folate and Free Amino Acid*. Master's thesis. China: Fujian Agriculture and Forestry University.

Lim, S.-H., You, M.-K., Kim, D.-H., Kim, J. K., Lee, J.-Y., and Ha, S.-H. (2016). RNAi-mediated suppression of dihydroflavonol 4-reductase in tobacco allows fine-tuning of flower color and flux through the flavonoid biosynthetic pathway. *Plant Physiol. Biochem.* 109, 482–490. doi: 10.1016/j.plaphy.2016.10.028

Liu, X., Ahmad, N., Yang, L., Fu, T., Kong, J., Yao, N., et al. (2019). Molecular cloning and functional characterization of chalcone isomerase from *Carthamus tinctorius*. *AMB Express* 9:132. doi: 10.1186/s13568-019-0854-x

Livak, K. J., and Schmittgen, T. D. (2001). Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) method. *Methods* 25, 402–408. doi: 10.1006/meth.2001.1262

Love, M. I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biotechnol.* 15:550. doi: 10.1186/s13059-014-0550-8

Lu, N., Rao, X. L., Li, Y., Jun, J. H., and Dixon, R. A. (2021). Dissecting the transcriptional regulation of proanthocyanidin and anthocyanin biosynthesis in soybean (*Glycine max*). *Plant Biotechnol. J.* 19, 1429–1442. doi: 10.1111/pbi.13562

Raza, A. (2020). Metabolomics: a systems biology approach for enhancing heat stress tolerance in plants. *Plant Cell Rep.* Online ahead of print. doi: 10.1007/s00299-020-02635-8

Rhoades, M. W., Reinhart, B. J., Lim, L. P., Burge, C. B., Bartel, B., and Bartel, D. P. (2002). Prediction of plant microRNA targets. *Cell* 110, 513–520.

Robinson, M. D., Mccarthy, D. J., and Smyth, G. K. (2010). edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* 26, 139–140. doi: 10.1093/bioinformatics/btp616

Sarma, A. D., Sreelakshmi, Y., and Sharma, R. (1997). Antioxidant ability of anthocyanins against ascorbic acid oxidation. *Phytochemistry* 45, 671–674. doi: 10.1016/s0031-9422(97)00057-5

Schuler, M. A., and Werck-Reichhart, D. (2003). Functional Genomics of P450S. *Annu. Rev. Plant Biol.* 54, 629–667. doi: 10.1146/annurev.arplant.54.031902.134840

Song, X. W., Wei, X. B., Di, S. K., and Pang, Y. Z. (2019). Advances in The Regulation of anthocyanin transcription factors and metabolic Engineering. *Acta Bot. Sin.* 54, 133–156.

Stephenson, M. L., and Zamecnik, P. C. (1978). Inhibition of Rous sarcoma Viral RNA Translation by a Specific Oligodeoxyribonucleotide. *Proc. Natl Acad. Sci. U. S. Am.* 75, 285–288 doi: 10.1073/pnas.75.1.285

Sun, C. H., Yang, C. Y., and Jason, T. C. (2018). Molecular identification and characterization of hydroxycinnamoyl transferase in tea plants (*Camellia sinensis* L.). *Int. J. Mol. Sci.* 19:3938. doi: 10.3390/ijms19123938

Trainin, T., Harel-Beja, R., Bar-Ya'akov, I., Ben-Simhon, Z., Yahalomi, R., Borochov-Neori, H., et al. (2021). Fine Mapping of the "black" Peel Color in Pomegranate (*Punica granatum* L.) Strongly Suggests That a Mutation in the Anthocyanidin Reductase (*ANR*) Gene Is Responsible for the Trait. *Front. Plant Sci.* 12:642019. doi: 10.3389/FPLS.2021.642019

Valerio, C., Vincenzo, D. A., Marco, E., Chiara, A., Petronia, C., Domenico, C., et al. (2021). Anthocyanins are key regulators of drought stress tolerance in tobacco. *Biology* 10:139. doi: 10.3390/biology10020139

Wan, L. Y., Li, B., Lei, Y., Yan, L. Y., Huai, D. X., Kang, Y. P., et al. (2018). Transcriptomic profiling reveals pigment regulation during peanut testa development. *Plant Physiol. Biochem.* 125, 116–125. doi: 10.1016/j.plaphy.2018.01.029

Wrolstad, R. E., Acree, T. E., Decker, E. A., Penner, M. H., Reid, D. S., Schwartz, S. J., et al. (2001). *Characterization and Measurement of Anthocyanins by UV-Visible Spectroscopy*. New Jersey: John Wiley & Sons, Inc.

Wu, W. C. (2018). Research progress of plant Anthocyanins. *J. Contemp. Chem. Ind.* 9, 183–185.

Xue, L., Wang, J., Zhao, J., Zheng, Y., Wang, H. F., Wu, X., et al. (2019). Study on cyanidin metabolism in petals of pink-flowered strawberry based on transcriptome sequencing and metabolite analysis. *BMC Plant Biol.* 19:423. doi: 10.1186/s12870-019-2048-8

Yao, L. H., Jiang, Y. M., Shi, J., TomÁs-BarberÁn, F. A., Datta, N., Singanusong, R., et al. (2004). Flavonoids in food and their health benefits. *Plant Foods Hum. Nutr.* 59, 113–122.

Yoshida, K., Iwasaka, R., Shimada, N., Ayabe, S.-I., Aoki, T., and Sakuta, M. (2010). Transcriptional control of the dihydroflavonol 4-reductase multigene family in *Lotus japonicus*. *J. Plant Res.* 123, 801–805. doi: 10.1007/s10265-010-0325-6

Zhang, H., Wu, G. R., Yang, W. M., Du, W. J., Zhao, J. Z., Yue, A. Q., et al. (2020). Analysis of pigment content in different color soybean seed coats. *J. Shanxi Agric. Sci.* 48, 492–496.

Zhang, H. S., Tian, H., Chen, M. X., Xiong, J. B., Cai, H., and Liu, Y. (2018). Transcriptome analysis reveals potential genes involved in flower pigmentation in a red-flowered mutant of white clover (*Trifolium repens* L.). *Genomics* 110, 191–200. doi: 10.1016/j.ygeno.2017.09.011

Zhang, Y. Y., Ma, R. J., Yu, M. L., and Song, H. F. (2014). Screening on SSR markers linked to flesh colour geneand nectarine gene of *Prunus persica*(L.)Batsch. *J. South. Agric.* 45, 1160–1165.

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.

frontiers | Frontiers in Plant Science

CORRECTION
published: 16 May 2022
doi: 10.3389/fpls.2022.929085

Corrigendum: Multi-Omics and miRNA Interaction Joint Analysis Highlight New Insights Into Anthocyanin Biosynthesis in Peanuts (*Arachis hypogaea* L.)

Jiawei Li, Yucong Ma, Mengdie Hu, Yulu Zhao, Bin Liu, Chunmei Wang, Min Zhang, Liping Zhang, Xinlei Yang and Guojun Mu**

State Key Laboratory of North China Crop Improvement and Regulation, North China Key Laboratory for Crop Germplasm Resources of Education Ministry, Laboratory of Hebei Provincial Crop Germplasm Resources, Hebei Agricultural University, Baoding, China

OPEN ACCESS

Edited and reviewed by:
Renata Rivera-Madrid, Scientific Research Center of Yucatán (CICY), Mexico

***Correspondence:**
Xinlei Yang
yangxinlei2500@163.com
Guojun Mu
mgj99999@126.com

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

Received: *26 April 2022*
Accepted: *03 May 2022*
Published: *16 May 2022*

Citation:
Li J, Ma Y, Hu M, Zhao Y, Liu B, Wang C, Zhang M, Zhang L, Yang X and Mu G (2022) Corrigendum: Multi-Omics and miRNA Interaction Joint Analysis Highlight New Insights Into Anthocyanin Biosynthesis in Peanuts (Arachis hypogaea L.). Front. Plant Sci. 13:929085. doi: 10.3389/fpls.2022.929085

Keywords: peanut, anthocyanin, testa, multi-omics joint analysis, miRNA interaction, qRT-PCR

A Corrigendum on

Multi-Omics and miRNA Interaction Joint Analysis Highlight New Insights Into Anthocyanin Biosynthesis in Peanuts (*Arachis hypogaea* L.)

by Li, J., Ma, Y., Hu, M., Zhao, Y., Liu, B., Wang, C., Zhang, M., Zhang, L., Yang, X., and Mu, G. (2022). Front. Plant Sci. 13:818345. doi: 10.3389/fpls.2022.818345

In the original article, there were two mistakes in **Figure 7** and **Supplementary Figure 1** as published. The author mistakenly wrote "AhmiRNA398" as "AhmiRNA50" in **Figure 7** and the paternal parent "Zizhenzhu" had been wrongly written as "Z18-40" in **Supplementary Figure 1**, which caused a discrepancy with the original text and affected readers' understanding. The corrected **Figure 7** and **Supplementary Figure 1** appears below.

The authors apologize for this error and state that this does not change the scientific conclusions of the article in any way. The original article has been updated.

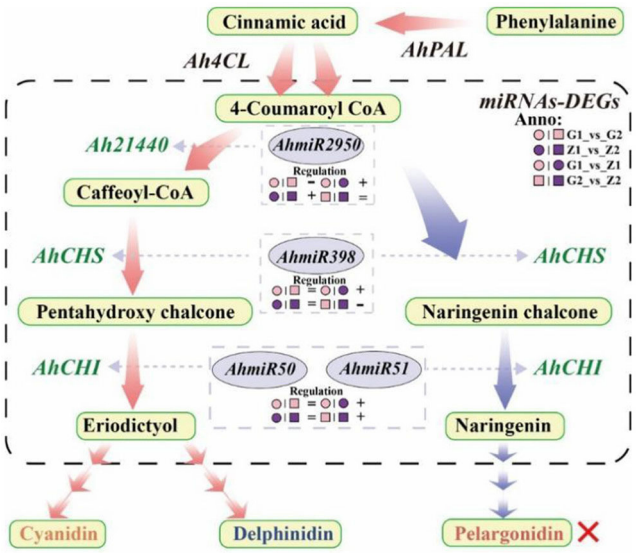
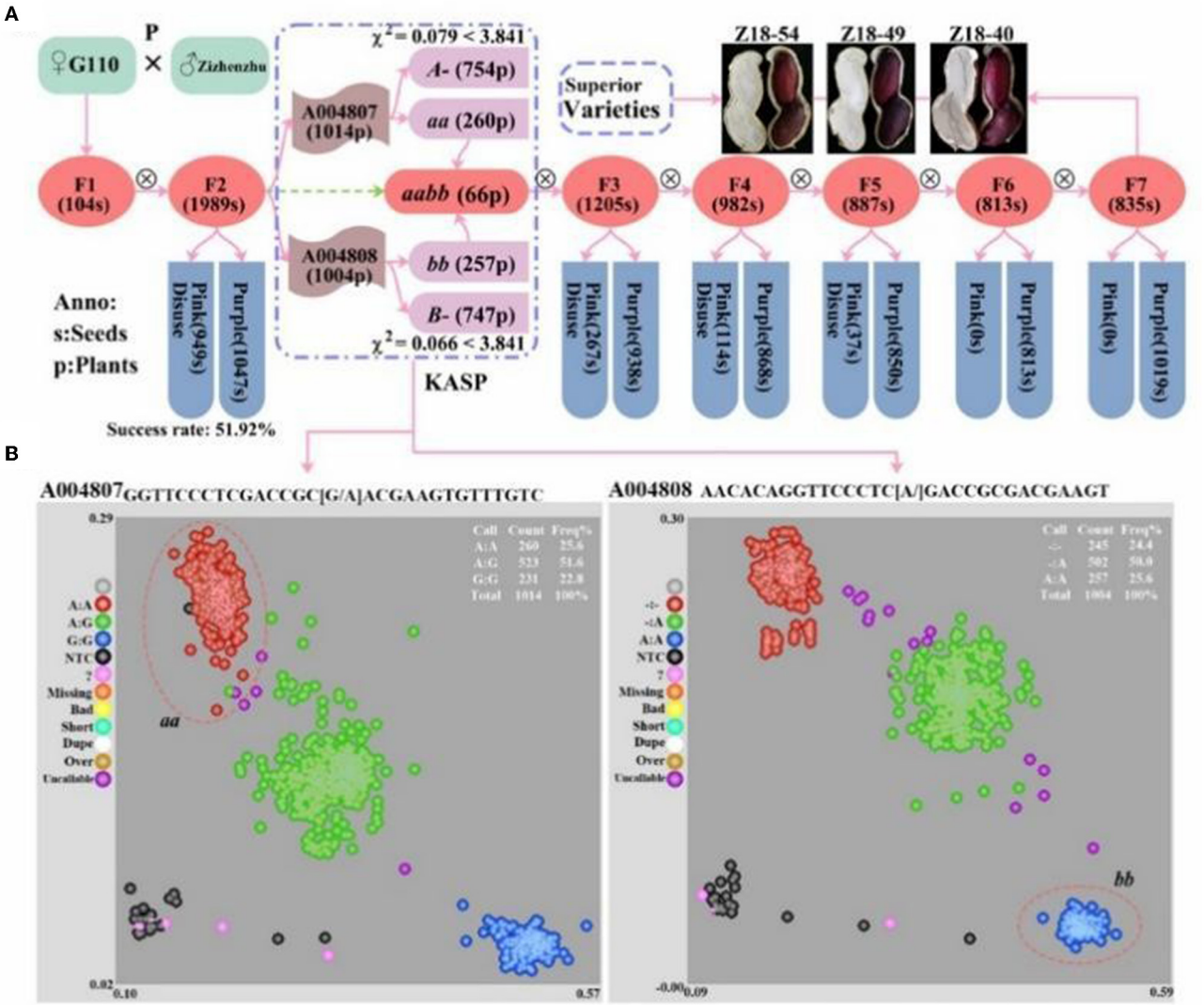


FIGURE 7 | Flow-process diagram of anthocyanin biosynthesis and miRNA interaction with target genes. The picture shows that rectangles with light yellow shading mean metabolites, light red arrows between metabolites mean preferred pathways, and light blue arrows indicate non-preferred pathways yet. The dashed rectangle indicates miRNAs-DEGs interaction. The green labels beside the arrows mean DEGs, the miRNAs corresponding to DEGs are marked in a gray ellipse, and both of them are connected by light gray arrows, and the regulation is indicated below the miRNAs. "○" with pink and purple means 30DAF of G110 and Z18-40, respectively. "□" with pink and purple means 45DAF of G110 and Z18-40, respectively. "+," "-," and "=" indicate that there is a positive, negative and none regulatory relationship between DEGs and miRNAs in the comparison on the left. The red cross indicates that there are no regulation expression in the metabolites or genes.



SUPPLEMENTARY FIGURE 1 |

Advantages of publishing in Frontiers

OPEN ACCESS
Articles are free to read for greatest visibility and readership

FAST PUBLICATION
Around 90 days from submission to decision

HIGH QUALITY PEER-REVIEW
Rigorous, collaborative, and constructive peer-review

TRANSPARENT PEER-REVIEW
Editors and reviewers acknowledged by name on published articles

Frontiers
Avenue du Tribunal-Fédéral 34
1005 Lausanne | Switzerland

Visit us: www.frontiersin.org
Contact us: frontiersin.org/about/contact

REPRODUCIBILITY OF RESEARCH
Support open data and methods to enhance research reproducibility

DIGITAL PUBLISHING
Articles designed for optimal readership across devices

FOLLOW US
@frontiersin

IMPACT METRICS
Advanced article metrics track visibility across digital media

EXTENSIVE PROMOTION
Marketing and promotion of impactful research

LOOP RESEARCH NETWORK
Our network increases your article's readership