



OPEN ACCESS

EDITED BY

Milind B. Ratnaparkhe,
ICAR Indian Institute of Soybean
Research, India

REVIEWED BY

Anuj Kumar,
Dalhousie University, Canada
Rajeev Ranjan,
Purdue University, United States

*CORRESPONDENCE

Ping Chen,
✉ chenping02@caas.cn

SPECIALTY SECTION

This article was submitted to Plant
Genomics,
a section of the journal
Frontiers in Genetics

RECEIVED 26 October 2022

ACCEPTED 31 January 2023

PUBLISHED 21 February 2023

CITATION

Feng X, Abubakar AS, Chen K, Yu C, Zhu A,
Chen J, Gao G, Wang X, Mou P and
Chen P (2023), Genome-wide analysis of
R2R3-MYB transcription factors in
Boehmeria nivea (L.) gaudich revealed
potential cadmium tolerance and
anthocyanin biosynthesis genes.
Front. Genet. 14:1080909.
doi: 10.3389/fgene.2023.1080909

COPYRIGHT

© 2023 Feng, Abubakar, Chen, Yu, Zhu,
Chen, Gao, Wang, Mou and Chen. This is
an open-access article distributed under
the terms of the [Creative Commons
Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use,
distribution or reproduction in other
forums is permitted, provided the original
author(s) and the copyright owner(s) are
credited and that the original publication
in this journal is cited, in accordance with
accepted academic practice. No use,
distribution or reproduction is permitted
which does not comply with these terms.

Genome-wide analysis of R2R3-MYB transcription factors in *Boehmeria nivea* (L.) gaudich revealed potential cadmium tolerance and anthocyanin biosynthesis genes

Xinkang Feng¹, Aminu Shehu Abubakar^{1,2}, Kunmei Chen¹,
Chunming Yu¹, Aiguo Zhu¹, Jikang Chen¹, Gang Gao¹,
Xiaofei Wang¹, Pan Mou¹ and Ping Chen^{1*}

¹Institute of Bast Fiber Crops, Chinese Academy of Agricultural Sciences, Changsha, China, ²Department of Agronomy, Bayero University, Kano, Nigeria

Gene family, especially MYB as one of the largest transcription factor family in plants, the study of its subfunctional characteristics is a key step in the study of plant gene function. The sequencing of ramie genome provides a good opportunity to study the organization and evolutionary characters of the ramie MYB gene at the whole genome level. In this study, a total of 105 BnGR2R3-MYB genes were identified from ramie genome and subsequently grouped into 35 subfamilies according to phylogeny divergence and sequences similarity. Chromosomal localization, gene structure, synteny analysis, gene duplication, promoter analysis, molecular characteristics and subcellular localization were accomplished using several bioinformatics tools. Collinearity analysis showed that the segmental and tandem duplication events is the dominant form of the gene family expansion, and duplications prominent in distal telomeric regions. Highest syntenic relationship was obtained between BnGR2R3-MYB genes and that of *Apocynum venetum* (88). Furthermore, transcriptomic data and phylogenetic analysis revealed that *BnGMYB60*, *BnGMYB79/80* and *BnGMYB70* might inhibit the biosynthesis of anthocyanins, and UPLC-QTOF-MS data further supported the results. qPCR and phylogenetic analysis revealed that the six genes (*BnGMYB9*, *BnGMYB10*, *BnGMYB12*, *BnGMYB28*, *BnGMYB41*, and *BnGMYB78*) were cadmium stress responsive genes. Especially, the expression of *BnGMYB10/12/41* in roots, stems and leaves all increased more than 10-fold after cadmium stress, and in addition they may interact with key genes regulating flavonoid biosynthesis. Thus, a potential link between cadmium stress response and flavonoid synthesis was identified through protein interaction network analysis. The study thus provided significant information into MYB regulatory genes in

Abbreviations: HTH, helix–turn–helix; MEME, Multiple Em for Motif Elicitation; NJ, Neighbor–Joining; ORF, open reading frame; GSDS, Gene structure display server; PlantTFDB, Plant Transcription Factor Database; RNA-seq, RNA-sequencing; SMART, Simple Modular Architecture Research Tool; TAIR, The Arabidopsis Information Resource; TFs, Transcription Factors; qPCR, Quantification real-time PCR; MWs, Molecular weights; MEME, Multiple Expectation maximization for Motif Elicitation; pIs, Isoelectric points; GRAVY, Grand average of hydropathicity; HX_1, Hongxuan No. 1 variety; ZZ_1, Zhongzhu No. 1 variety.

ramie and may serve as a foundation for genetic enhancement and increased productivity.

KEYWORDS

ramie, MYB TFs, expression profiles, cadmium stress, anthocyanin biosynthesis, protein interaction

Background

The TFs are generally made up of four functional domains regions that include DNA binding domain, transcription regulation domain, oligomerization site, as well as the nuclear localization signal based on which as well as the number domain residues TFs were classified (Sacks et al., 2018). MYB TFs being unarguably the largest family are distinguished by its highly conserved MYB domain composed of one to four adjacent imperfect tandem repeats at the N-terminus (Riechmann and Ratcliffe, 2000). Each repeat is approximately 50–53 amino acid residues in length, containing three α -helices that together form a helix-turn-helix (HTH) secondary structure that interacts with the major DNA at the specific recognition site C/TAACG/TG during transcription (Dubos et al., 2010). On the contrary, the C-terminal region is highly variable that leads to the wide range of regulatory roles of the MYB gene family (Ogata et al., 1996). According to the number of the adjacent MYB repeats, MYB genes are categorized into four subfamilies: 1R-MYB, R2R3-MYB, R1R2R3-MYB, and 4R-MYB. The R2R3-MYB are generally the most diverse and extensively studied group (Jiang and Rao, 2020). This also led to the subfunctionalization of the large MYB family (Wani et al., 2018). For example, the typical subfamily S6, S5, S4, S7, S44, and S79 participated in the regulation of flavonoid biosynthesis. Representative S1, S2, S11, S17, S20, S22, and S38 subfamilies that contribute to biological stress response, and S9, S15, S18, S25, and S27 subfamilies that are related to growth and development (Gahlaut et al., 2016). In addition, the experimental data of S16, S24, S28, S31, S33, S36, and S37 and other subfamilies are limited and need further research (Wu et al., 2022).

The MYB TF is present in all eukaryotes and was first reported in 1987 in *Zea mays* which was functionally characterized to partake in regulating anthocyanin biosynthesis (Paz-Ares et al., 1987). Nowadays, with the advance of plant genome-wide association and molecular biology methods, an increasing number of studies have shown MYB proteins to play significant role in regulating plant primary and secondary metabolism, plant development, and in the response to various biotic and abiotic stresses (Feller et al., 2011; Baldoni et al., 2015; Ng et al., 2018; Li et al., 2019). For example, *CsTS1* regulates theanine biosynthesis (Zhang et al., 2020), *CaMYB31* regulates capsaicin biosynthesis (Arce-Rodríguez and Ochoa-Alejo, 2017), *DkMyb4* regulated proanthocyanidin biosynthesis in persimmon (Akagi et al., 2009). In addition, many MYBs associated with anthocyanin biosynthesis have been found in many species in recent years, such as *Arabidopsis* (*AtMYB75*, *AtMYB90*, *AtMYB113* and *AtMYB114*) (Bac-Molenaar et al., 2015; Jian et al., 2019; Muñoz-Gómez et al., 2021), citrus (*CsRuby1* and *CsMYB3*) (Huang et al., 2020) and tomato (*SlMYB12*) (Ballester et al., 2010). They interact with anthocyanin synthesis pathway genes through MBW complex to

form anthocyanin synthesis regulatory network (Lloyd et al., 2017). The resistance mechanism mediated by MYB has also been reported. For example, *GbMYB5* confers drought tolerance in cotton and transgenic tobacco plants (Chen et al., 2015a), another rice MYB (*OsMYB3R-2*) exhibited enhanced cold tolerance by alteration in cell cycle and ectopic expression of stress genes (Ma et al., 2009). To resist soil cadmium contamination, plants have evolved fine-tuned mechanisms to protect cells from Cd toxicity by partitioning Cd into vacuoles or trapping free Cd²⁺ in the cytosol (Clemens and Ma, 2016). Studies from several independent laboratories have demonstrated the roles of HPPs in heavy metal transport, accumulation, and detoxification (Chu et al., 2005; Tehseen et al., 2010; Zhao et al., 2013). In addition, *AtMYB59* is involved in plant growth and cadmium stress response in *Arabidopsis Thaliana* by acting as negative regulator of Ca signaling and homeostasis (Fasani et al., 2019). *AtMYB4* regulates cadmium tolerance by enhancing protection against oxidative damage and increasing expression of *phytochelatin synthase 1 (PCS1)* and *metallothionein 1C (MT1C)* (Agarwal et al., 2020). Other studies have shown that ABA is a feedback mechanism that controls Cd uptake and accumulation in plant cells (Hsu and Kao, 2003). Induced by cadmium stress, *ABI5* is upregulated and interacts with *AtMYB49*, preventing its binding to downstream gene promoters, thereby reducing Cd accumulation. On the other hand, *AtMYB49* positively regulates the expression of *bHLH38* and *bHLH101*, resulting in the activation of a metal transporter involved in Cd absorption (Zhang et al., 2019). Interestingly, these mechanisms do not exist independently. A recent study from the perspective of root development in response to abiotic stress revealed that *MYB36* balances root lignification and suberization by mediating the *MYB74-MYB36-MYB92-MYB93* subnetwork (Xu et al., 2022). The above studies indicate that the functions of MYB transcription factors are diversified, gradually forming a complex network mechanism.

Ramie is an important fiber crop in the Urticaceae family which is also widely used as feed and other industrial raw materials (Kiprioti et al., 2015; Mu et al., 2020; Wang et al., 2021). In addition, its perennial nature makes it extremely resistant, and its high cadmium tolerance makes it a soil remediation crop in mining areas in southern China. The diverse functions of R2R3-MYB can connect a variety of regulatory metabolic networks. Therefore, we proposed to study the functional differentiation of MYB family genes in secondary metabolism and abiotic stress. Despite the important role in plant specific processes, very limited reports are present to date on the functional characteristics of R2R3-MYB transcription factor in ramie. Therefore, this study comprehensively investigated and reported gene structure, gene duplication, chromosomal localization, phylogeny, and cis-acting elements of the BnGR2R3-MYB gene. Transcriptome expression profiling coupled with UPLC-QTOF-MS was used to identify BnGMYBs involved in anthocyanin synthesis. Two ramie

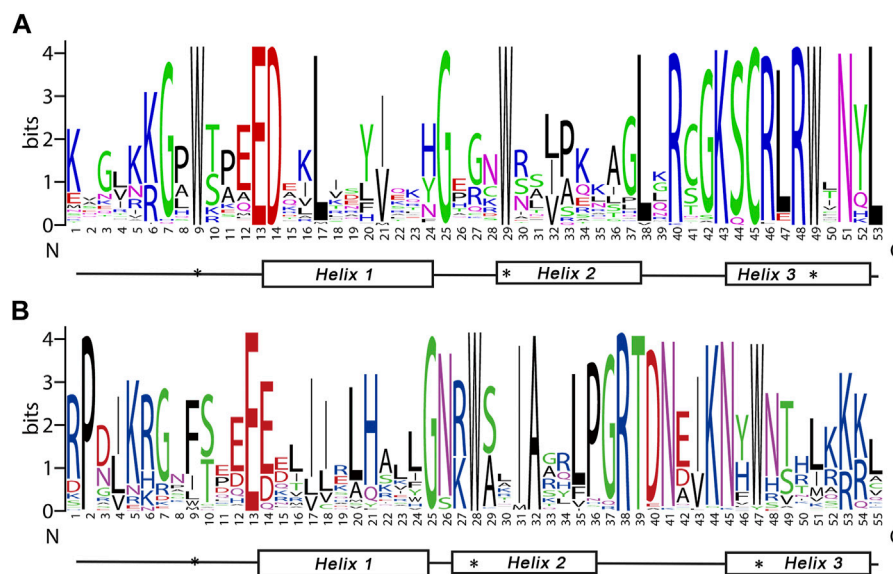


FIGURE 1

Consensus sequence and the level of conservation of R2R3-MYB domains from ramie. The sequence logos of the R2 (A) and R3 (B) MYB repeats are based on full-length alignments of all ramie R2R3-MYB domains. The bit score exhibits the information content for each position in the sequence. The position of the three α -helices were marked (Helix 1–3). The conserved tryptophan residues (Trp, W) in the MYB domain are marked with black asterisks.

varieties (HX_1 and ZZ_1) were exposed to cadmium stress and expressions of some selected MYBs explored to understand their function in stress responses in the ramie. Protein interaction network analysis in addition, revealed a potential link between cadmium stress and flavonoid synthesis.

Results and discussion

Identification and sequence feature of ramie MYB genes

To identify the MYB family genes in ramie genome, the amino acid sequence of hidden Markov model (HMM) profile of the Pfam MYB domain (PF00249) was used as a query and obtained 245 amino acid sequences that contain MYB or MYB-like repeats. Subsequent processing with FGENESH-M returned 211 sequences. After confirmation of the presence of MYB domains base on the PROSITE, Pfam and NCBI-CDD analyses, we finally established a total of 200 non-redundant MYB proteins which included 89 1R-MYB, 105 R2R3-MYB, 5 R1R2R3-MYB and 1 4R-MYB proteins. According to previous reports, R2R3-MYB family members not only represent more than half of the proportion of the total MYB proteins, but also contain most functional genes, such as specialized metabolism, including the benzenoid, phenylpropanoid, terpenoid, and glucosinolate (GSL) pathways (Dubos et al., 2010). Therefore, we selected only the R2R3-MYB for further study. Based on their order on the chromosome, the R2R3-MYB genes (BnGMYBs) were renamed *BnGMYB1* to *BnGMYB102*. Three of the 105 R2R3-MYBs (*Maker00022368*, *Maker00031449*, and *Maker00022368*)

which could not be conclusively mapped to any chromosomes, were renamed *BnGMYB103*–*BnGMYB105* respectively (Supplementary Table S1). This may be due to a large number of gaps, faulty assembly or contamination of organelle DNA and circular DNA *in vitro*. The molecular weight of the protein as obtained using ProtParam analytical tool revealed *BnGMYB52* as the smallest with 135 amino acids (aa), and *BnGMYB1* (593 aa) the largest. The MW of the proteins ranged from 15.6 to 65.0 kDa, and the pI ranged from 4.56 (*BnGMYB23*) to 10.08 (*BnGMYB96*). All the R2R3 but two are localized in the nuclear region, whereas the remaining two were in either chloroplast or cytoplasm (Supplementary Table S1).

R2, R3 repeats are significantly conserved sequences within the MYB domain regions. Sequence logos were constructed to investigate the homologous domain features and the frequency of amino acids (Figure 1). There were 108 amino acids conserved residues in the BnGR2R3-MYB. 11 residues in the R2 and 6 in the R3 were completely among all BnGMYBs. These included the most prominent series of evenly distributed and highly conserved triplet tryptophan (Trp., W) residues in each repeat at positions 9, 29 and 49 of the R2 repeat; 28 and 47 of the R3 repeat. These characteristic amino acids are known to play an important role in sequence-specific DNA binding, and are considered landmarks for plant MYB proteins (Jiang et al., 2004). R3 repeats in *BnGMYB66*, *BnGMYB56*, *BnGMYB67* have the last Trp residue replaced with Tyrosine (Tyr., Y) or phenylalanine (Phe., F) which were consistent with those of other species (Li et al., 2016; Liu et al., 2020). Insertion of leucine residue (Leu., L) between second and third helices of the R2 repeat have been reported to represents evolutionary relationship of R2R3-MYB domain (Williams and Grotewold, 1997). In this study, 83 (79.05%) BnGR2R3-MYBs were observed to have Leu-38 inserted between the two helices (Supplementary Table S2).

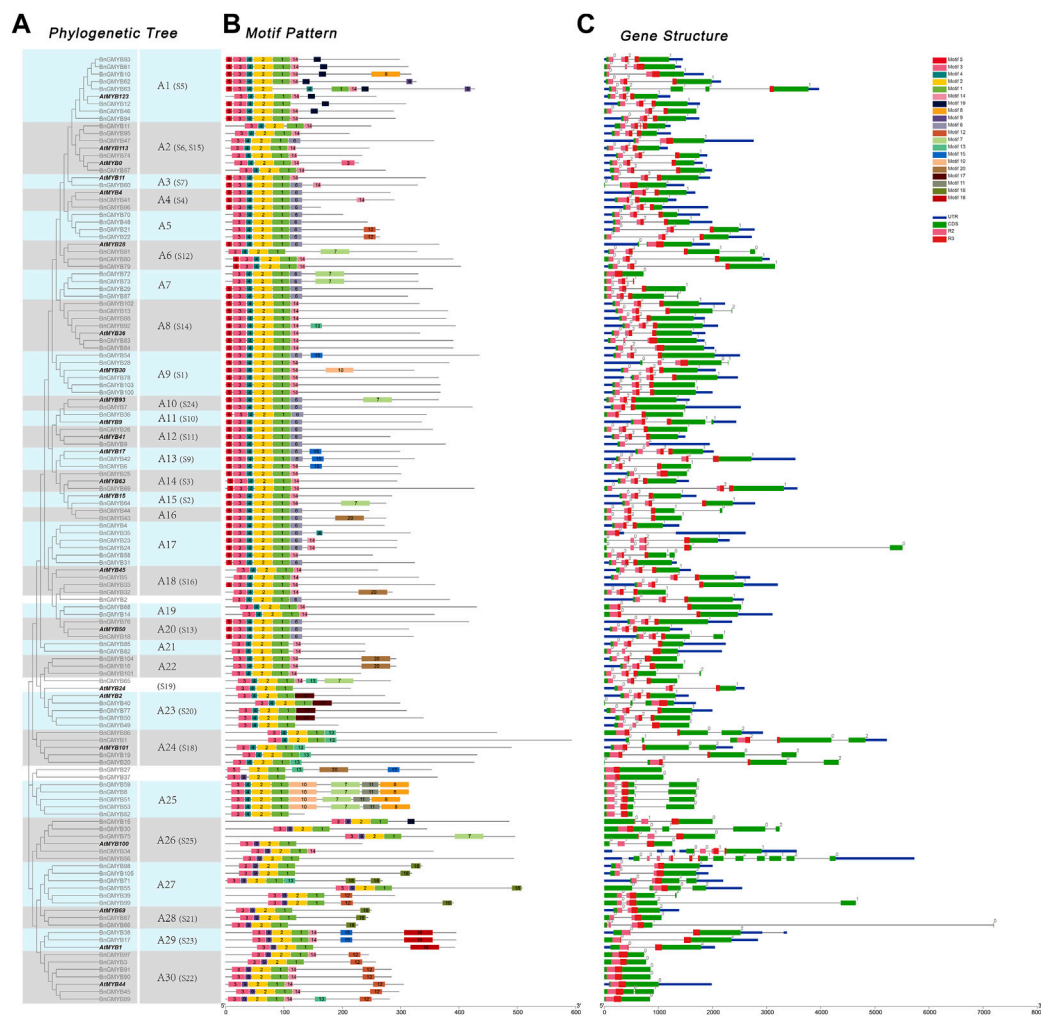


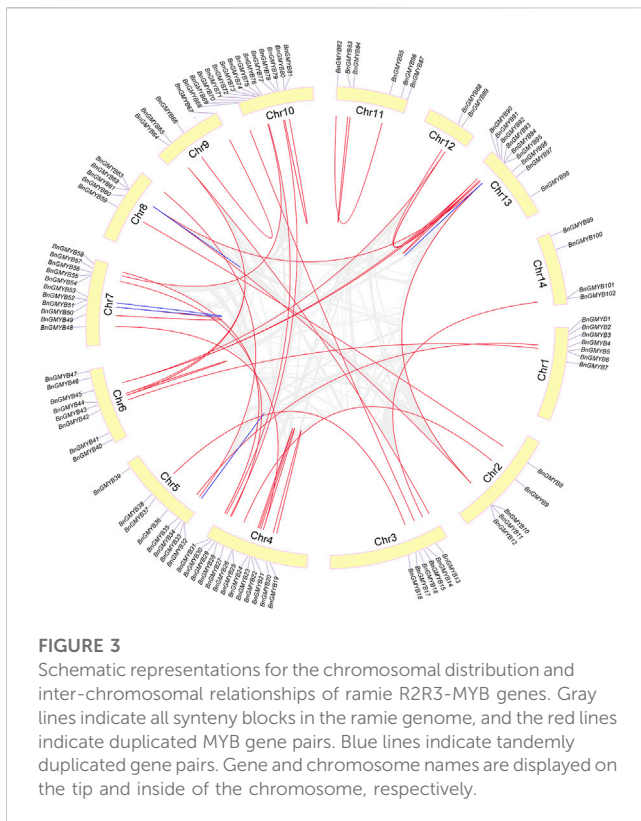
FIGURE 2 Phylogenetic relationships, gene structure and architecture of conserved protein motifs in BnGR2R3-MYB genes. **(A)** The maximum likelihood phylogenetic includes 105 R2R3-MYB proteins from ramie and 23 representatives from Arabidopsis **(B)** The motif architecture of ramie MYB proteins. 20 different motifs are displayed in different colored boxes. **(C)** Exon-intron structure of ramie MYB genes. Green boxes indicate exons; blue boxes indicate untranslated 5'- and 3'- regions; black lines indicate introns. R2, R3 MYB domain are highlighted by boxes of different shape and color.

Gene structure and motif composition of BnGMYBs

We performed a phylogenetic analysis of R2R3-MYB proteins using neighbor-joining (Guo et al., 2019) and maximum likelihood (ML) methods with 1,000 bootstrap replicates. The tree topologies obtained using the two methods were largely similar with a very few variations in the gene classification (Figure 2A and Supplementary Figure S1). Since the ML method can choose the optimal alternative model and optimize the evolutionary tree with certain topological structure and branch length, we adopted it for further characterization. The R2R3-MYB members of ramie were subdivided into 30 subgroups (designated A1-A30 in this study) using Arabidopsis MYB proteins as reference (Dubos et al., 2010). Nine clades only contained ramie R2R3-MYB members, and four genes did not fit into any subgroup.

To investigate the relationship between gene structural function and evolution, we analyzed the structural diversity of the BnGR2R3-

MYB family, which showed genes having the same genetic structure clustering together as obtained in the phylogenetic tree (Figure 2A). MYB conservative motifs as obtained using MEME online tool (Figure 2B) revealed the presence of 20 conserved motifs in the ramie MYB proteins (Supplementary Table S3). The gene structure analysis showed most of gene coding sequences disrupted by introns, except for almost all members from A30 subgroup, ungrouped *BnGMYB27/37* and *BnGMYB72* (Figure 2C). The number of introns ranges from zero to eleven, with most genes having three exons and two introns in line with previous reports (Du et al., 2012a). Interestingly, most intron insertions occur in R2, R3 conserved domains indicating the important role of these conserved domains in plants. The highly conserved gene structure and domain influence each other, which provided important evidence for the division of subgroups. Similarly, protein architecture was conserved within a specific subgroup. Phylogenetic tree, gene structure, and motif analysis suggested similar functions for MYB proteins within the same



subgroup. In addition to the conserved motif representing R2R3, other similar motifs were shared by the BnGR2R3-MYB members within the same subgroup, for example, motifs 11, 17, and 19 were unique to members in clades A25, A23, and A1 respectively.

BnGMYB genes chromosome distribution and synteny

Ramie R2R3-MYB genes were distributed unevenly in the species 14 chromosomes (Figure 3), with Chr10 having the largest number (14) followed by Chr4 (13) and the least number (2) found on Chr12. No significant correlation was obtained between the chromosome length and the number, while some of the genes were concentrated at either top or bottom of the chromosome. The closer the genes are to each end, the more likely they are to crossover (Schilling et al., 2020).

Based on BLASTP and MCScanX methods, we investigated the gene duplication events. Reference to Holub (Holub, 2001) for a description of tandem duplication, among the BnGR2R3-MYB genes, the blue lines in Figure 3 showed four pairs of tandemly duplicated genes (*BnGMYB32/33*, *BnGMYB51/52*, *BnGMYB61/62*, *BnGMYB93/94*) located in Chr5, Chr7, Chr8, and Chr13, respectively, whereas the red lines showed 39 segmental duplication pairs between BnGR2R3-MYB genes (Supplementary Table S4) including two MYB-related genes (*Maker00025472*; *Maker00000772*) and one 3R-MYB gene (*Maker00072694*). Some BnGMYBs located on Chr13, Chr6, and Chr4, such as *BnGMYB91*, *BnGMYB90*, *BnGMYB89*, *BnGMYB21*, *BnGMYB44*, and *BnGMYB45*, were involved in multiple duplication events.

BnGMYB89 (Chr12) was however interrelated with *BnGMYB90/91* (Chr13) and *BnGMYB45* (Chr6). Interestingly, all segmentally duplicated genes fall into A30 subgroups together with *A. thaliana AtMYB44* (Figure 2A) making us speculate that this specific population of R2R3-MYB genes duplication in ramie might be closely related to stress response (Kranz et al., 1998; Shim and Choi, 2013; Nguyen and Cheong, 2018), and might play a vital role in the growth and development of the plant. The segmental and tandem duplication events are major drivers of the gene family expansion (Cannon et al., 2004). The positive pressure criteria were selected based on M. Lynch (Lynch and Conery, 2000): $Ka/Ks < 1$ stands for purifying selection, $Ka/Ks = 1$ means neutral selection, while $Ka/Ks > 1$ signifies positive selection. We found a high sequence divergence value among the four segmental gene pairs. These sequences diverged greatly, and evolution distance was long, owing to the occurrence of a large number of synonymous mutations. The remaining segmental and tandem duplicated BnGMYB gene pairs were $Ka/Ks < 1$ except *BnGMYB66/67* ($Ka/Ks = 1.04$). These results suggest that the ramie R2R3-BnGMYB gene family had evolved under the effect of purifying selection (Supplementary Table S4).

Syntenic maps of ramie with five other plants, comprising dicots [*A. thaliana*, *Cannabis sativa* (*C. sativa* L.), *A. venetum* (*Apocynum venetum* L.)] and monocots (rice and maize) conducted to further explore potential evolutionary mechanisms of BnGR2R3-MYB genes, revealed 88, 58, 53, 8, and 4 numbers of orthologous gene pairs (Figure 4). Collinearity analysis showed highest collinearity segments of the ramie with *A. venetum* and *C. sativa* in consistent with the phylogenetic divergence time estimation which supported the placement of ramie within the same clade as *C. sativa* (Luan et al., 2018). Some of the BnGR2R3-MYBs were found to be associated with multiple syntenic gene pairs, particularly *BnGMYB89* and four *A. thaliana* genes, suggestion that these genes might have played an important role in MYB gene family during evolution. Remarkably, the number of collinear gene pair between ramie and *A. venetum*, *C. sativa*, and *A. thaliana* (dicotyledonous) was far greater than between ramie and *Oryza sativa/Z. mays* (monocotyledonous) (Liu et al., 2018; Luan et al., 2018; Huang et al., 2019). 19 BnGR2R3-MYBs identified in dicotyledonous plant were absent in monocotyledonous plant, suggestion that these orthologous pairs (19 BnGR2R3-MYBs) formed after the divergence of dicotyledonous and monocotyledonous plants. Additionally, one collinear pair (*BnGMYB64*) was identified in all the six detected plants, indicating that these orthologous pairs might have already existed before the ancestral divergence (Supplementary Table S5).

Comparative phylogenetic analysis of BnGMYB and R2R3-MYB family from six different plant species

A Neighbor-joining (Guo et al., 2019) phylogenetic tree was constructed using the 105 full-length R2R3-MYB protein sequences of ramie and those from *Arabidopsis* (126), rice (81), maize (156), tomato (121) and pineapple (93) and the result shown in Figure 5 and S2. The phylogenetic analysis generated 35 clades (C1-C35) among six species and the numbers of the MYB members in each

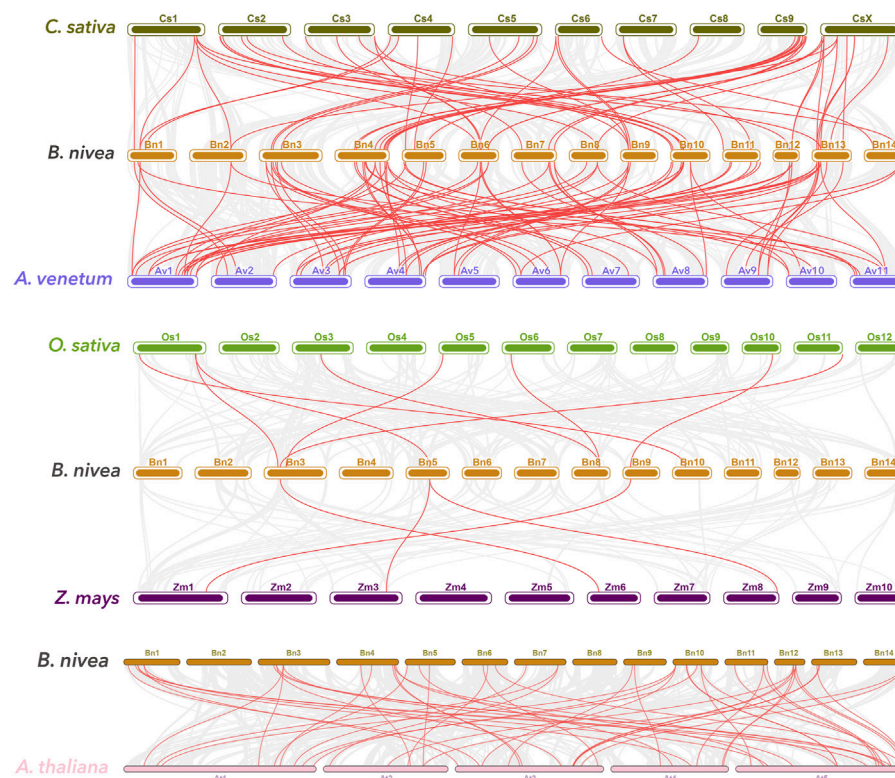


FIGURE 4

Synteny analyses of R2R3-MYB genes between ramie and five representative plants: *Boehmeria nivea*, *Cannabis sativa*, *A. venetum*, *O. sativa*, *Zea mays* and *A. thaliana*. Gray lines in the background indicate the collinear blocks within ramie and other plant genomes, while the red lines highlight the syntenic R2R3-MYB gene pairs.

species were listed (Figure 5). The fact that this tree was in good agreement with the classification results of *A. thaliana* (Dubos et al., 2010) and other plants (Li et al., 2016; Liu et al., 2017; Liu et al., 2020) demonstrated the reliability of the data. Moreover, the comprehensive phylogenetic analysis of the R2R3-MYB proteins in these six species might provide more evolutionary historical clues. For example, monocotyledon MYB proteins in the same group were often found clustered together, while MYB proteins from dicotyledonous plants were clustered side by side in lateral branches. Meanwhile, several clades were found only in some particular species. For example, none of rice, maize, or pineapple MYB members fall in the C5, C11, and C16. Similarly, C31 had only monocotyledons. This indicated a great evolutionary distance between monocotyledons and dicotyledons, and that these specific MYB genes might have evolved after divergence of the two classes. Apart from these, all other clades contained at least a member from each of the six species, suggesting common ancestry before the species diverged. The varying number of MYBs from the species in a particular clade is an indicative that MYB genes expansion may be more active in some plant species, and this expansion inequality may be related to the well conserved chromosome karyotype (Ming et al., 2015). Furthermore, MYBs in the same clade may have similar function with to the model *Arabidopsis*, thus provide direction for functional identification of ramie MYB proteins (Kranz et al., 1998; Stracke et al., 2001; Dubos et al., 2010). reference Polyploidy and subsequent genes loss exist in

most species is an important driving force of species evolution. Based on the comparative genomic analysis, it could be inferred (Supplementary Figure S3) that ramie has undergone at least one round of genome-wide duplication events. Phylogenetic analysis of MYB gene family showed that the number of BnGMYB members in some clades was consisted with rice and maize. Ramie also undergone the genome-wide duplication event common to rice (Thiel et al., 2009) and maize (Si et al., 2019).

Analysis of the cis-acting elements of the BnGR2R3-MYB genes and gene ontology annotation

The upstream promoter regions (1000bp) of the BnGR2R3-MYB genes were predicted for better understanding of the R2R3-MYB genes functions. There were total of 41 responsive cis-acting elements (Figure 6 and Supplementary Table S6), which made up of 22 light responsive elements, 8 phytohormone responses, 6 plant growth development-related responses and 5 abiotic stresses (adversity). Light response elements being the most abundant are present in all genes. Box-4, G-box (light responsiveness), ARE (anaerobic induction) and ABRE (abscisic acid responsiveness) were the most prominent elements, and presence in more than 100 of the 105 BnGR2R3-MYB genes, suggesting that these elements might play an important role in regulating gene expression. A total

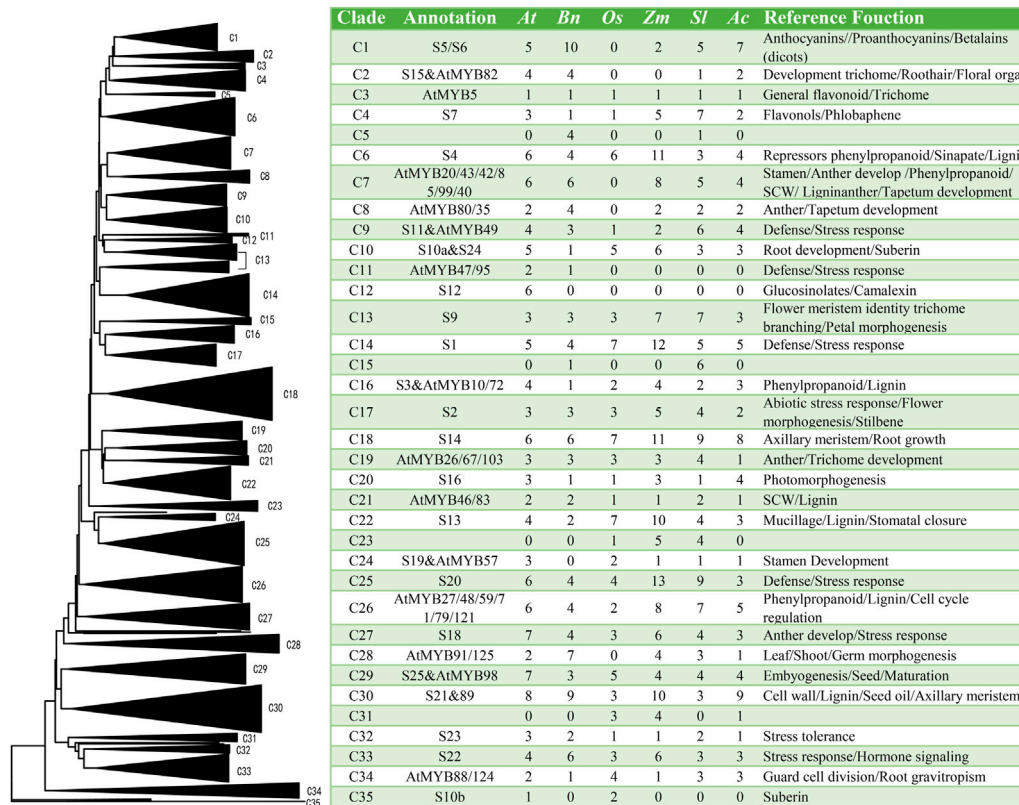


FIGURE 5 Neighbor-joining (Guo et al., 2019) tree showing relationships among R2R3-MYB proteins from *Boehmeria nivea* (Chen et al., 2015b), *A. thaliana* (At), *O. sativa* (Os), *Zea mays* (Pierrugues et al., 2001), *S. lycopersicum* (Sl) and *A. comosus* (Ac). MYB proteins from the six species were designated as C1 to C35. The table on the right indicated the number of the subgroup members and function of clades in each species.

of 65 (61.9%) genes were found to contain ABRE. Genes containing anaerobic induction, wound response, low temperature response, and drought stress related elements accounted for 55.2%, 26.7%, 24.8% and 17.1%, respectively. In addition, there are 118 elements in all genes involved in the MeJA response. These results indicate the potential function of BnGR2R3-MYB genes in response to various stresses.

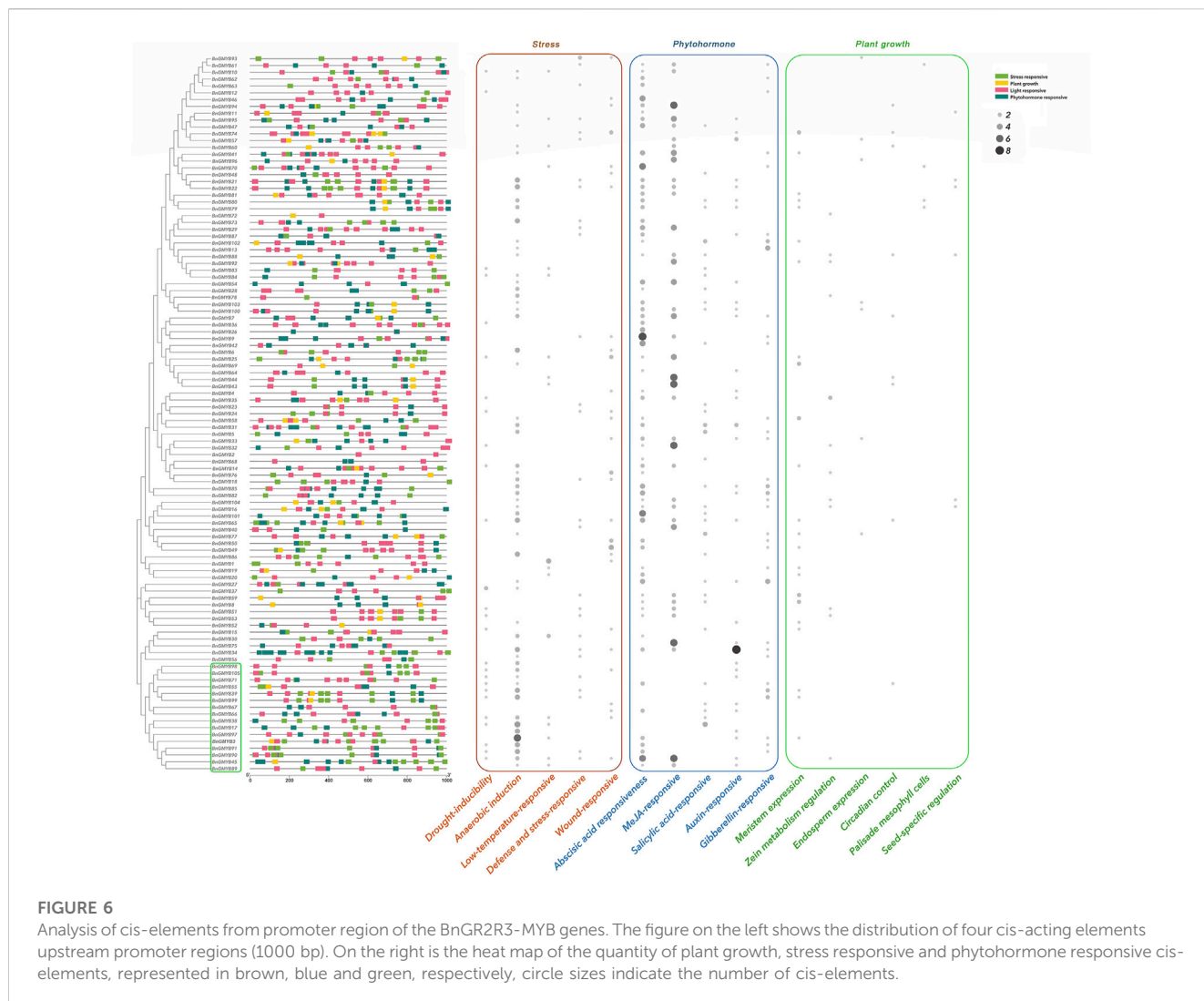
Almost all BnGR2R3-MYB genes as obtained in this study contained several cis-acting elements associated with abiotic stresses such as anaerobic induction, drought, low temperature, defense and stress, which is in line with literature (Yamaguchi-Shinozaki and Shinozaki, 2006; Chen et al., 2021). The abiotic stress responsive elements were mostly clustered together, for example, A29 and A30 phylogenetic subgroups in Figure 6 (corresponding to the C32 and C33 subgroups in Figure 5) which were consistent with their functional clustering and might be related to the function of gene groups. This further reflected the accuracy of the results.

Gene ontology (GO) (Ashburner et al., 2000) annotations of 105 BnGR2R3-MYB proteins is displayed in Supplementary Figure S4 and indicated the proteins involvement in biological process (BP), cellular component (CC) and molecular function (MF). The GO term “binding” (GO: 0005488) best described the greatest number of genes (96, 92.38%), and “single-organism developmental process” (GO: 0044767), “developmental process” (GO:0032502) were significantly enriched in top 20 of biological process. These GO

annotations of BnGR2R3-MYB proteins were in agreement with the experimental findings in other species (Katiyar et al., 2012; Smita et al., 2015; Chen et al., 2017).

Expression profiling of BnGR2R3-MYB genes with RNA-seq in different tissues

We analyzed transcript levels obtained from transcriptome data of 9 different samples comprising tissues (phloem, root, leaf) and developmental stages of ramie to profile expression of BnGR2R3-MYB. The results (Figure 7 and Supplementary Table S7A) showed most BnGR2R3-MYBs to have exhibited different expression patterns in line with what was obtained in other plants (Chen et al., 2021). Three BnGR2R3-MYBs (BnGR2R3MYB15; BnGR2R3MYB75; BnGR2R3MYB101) were not detected in all nine samples, suggesting that they might be pseudogenes, or only expressed in other tissues. Hierarchical cluster analysis of the BnGR2R3-MYBs from 9 the different samples enabled grouping them into three: high expression, preferential expression, and relatively lower expression. The expression grouping in most cases indicated that the expression patterns of genes were significantly different. However, some closely related MYB genes showed highly similar transcript levels, for example, three members of A30 (*BnGR2R3MYB89/90/91*) all belonged to high expression group. A total of 57 genes were expressed in all tissues,

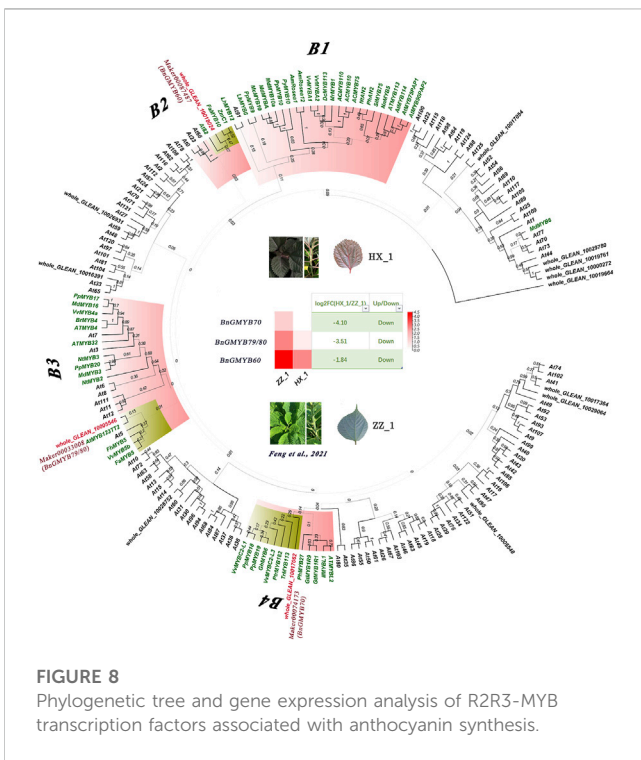
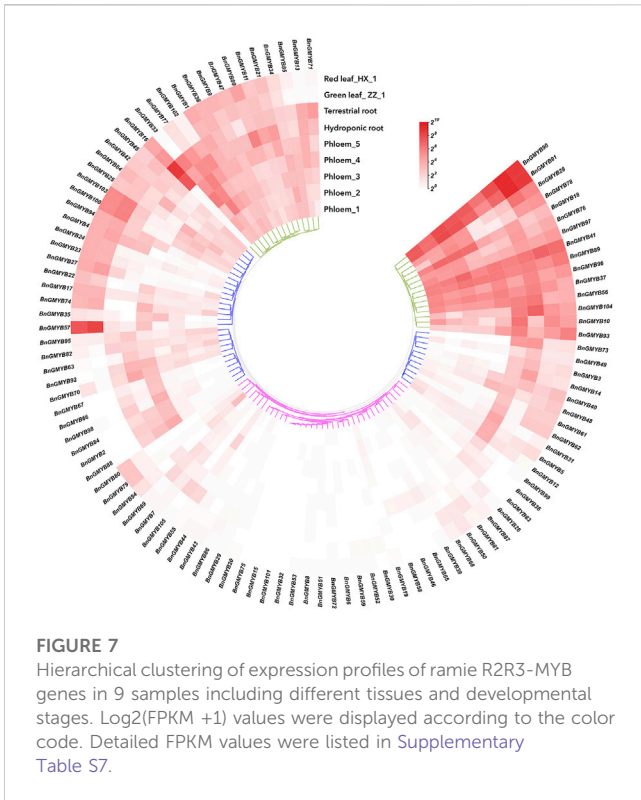


and 16 from which showed constitutive expression (FPKM>2). In order to better understand the preferential expression of genes, we filtered genes based on meeting two set rules: FPKM>2 (at least in one tissue); tissues with the highest expression levels should have twice as much expression in at least one of the other tissues. Based on this, 11 genes in phloem_third period, 19 genes in leaf of variety ZZ_1, 22 genes in terrestrial root were found to have exhibited preferential expression over the others (Supplementary Table S7C).

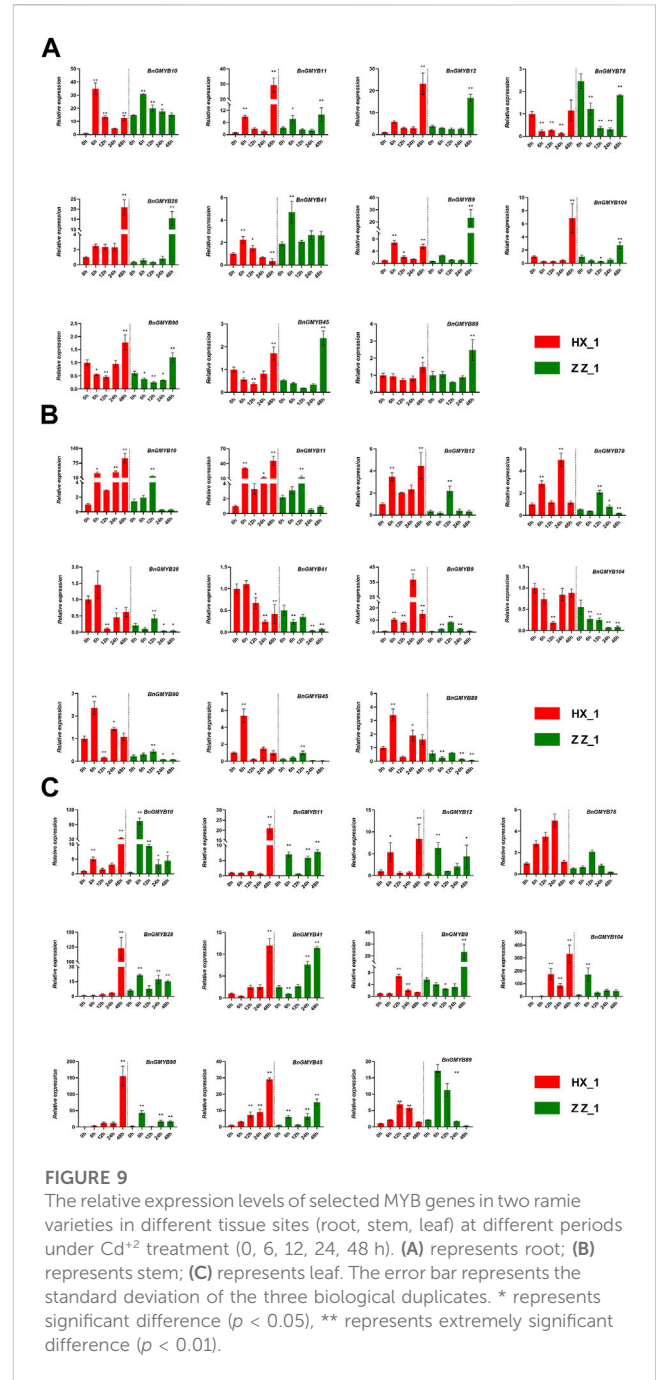
In our previous study, we conducted transcriptomic and non-targeted metabolomic studies of two ramie varieties which have distinct differences in leaf color, HX_1 (red-leaf variety) and ZZ_1 (green-leaf variety), and revealed that the color differences between the duo was due to differences in expression of anthocyanin synthesis pathway related gene, with no clarity in the absolute content of ramie anthocyanin and the function of regulatory gene (e.g., transcription factors) (Feng et al., 2021). In this study, we examined the composition and content of anthocyanin in the two varieties. The results showed that the anthocyanin in ramie was mainly Cyanidin-3-rhamnoside (Supplementary Figure S5). The Cyanidin-3-rhamnoside content in HX_1 was as computed based on the peak value was 3.147 $\mu\text{g}/\text{mL}$, whereas the anthocyanin was

almost absent in ZZ_1. Thus, the differences in the leaf color were attributed to this.

Several transcription factors including MYB have been reported to interact with key genes involved in anthocyanin synthesis (Lloyd et al., 2017). Fourteen differentially expressed MYB transcription factors searched in the transcriptome data showed 5 to have been upregulated with the remaining 9 downregulated. The highest differential expression ploidy reached up to 17-fold, implying that MYB was very active in two varieties of different colors (Supplementary Table S7B). Therefore, we selected these MYBs for the construction of phylogenetic tree with 50 anthocyanin synthesis-related MYB genes identified in other plants and R2R3-MYBs of the model plant Arabidopsis. As shown in Figure 10, the 50 anthocyanin synthesis related MYBs were grouped into five subgroups, and three of the 14 differentially expressed MYB genes were clustered into these subgroups: *BnGMYB60* (B2), *BnGMYB79/80* (B3), and *BnGMYB70* (B4). All three genes which are homolog of *PtrMYB182* (Yoshida et al., 2015)/*VvMYBC2L3* (Cavallini et al., 2015), *AtMYB123* (TT2) (Zimmermann et al., 2004), and *AtMYB82* (Zhou et al., 2020) involve anthocyanin synthesis repression were downregulated in HX_1 (Figure 8). No



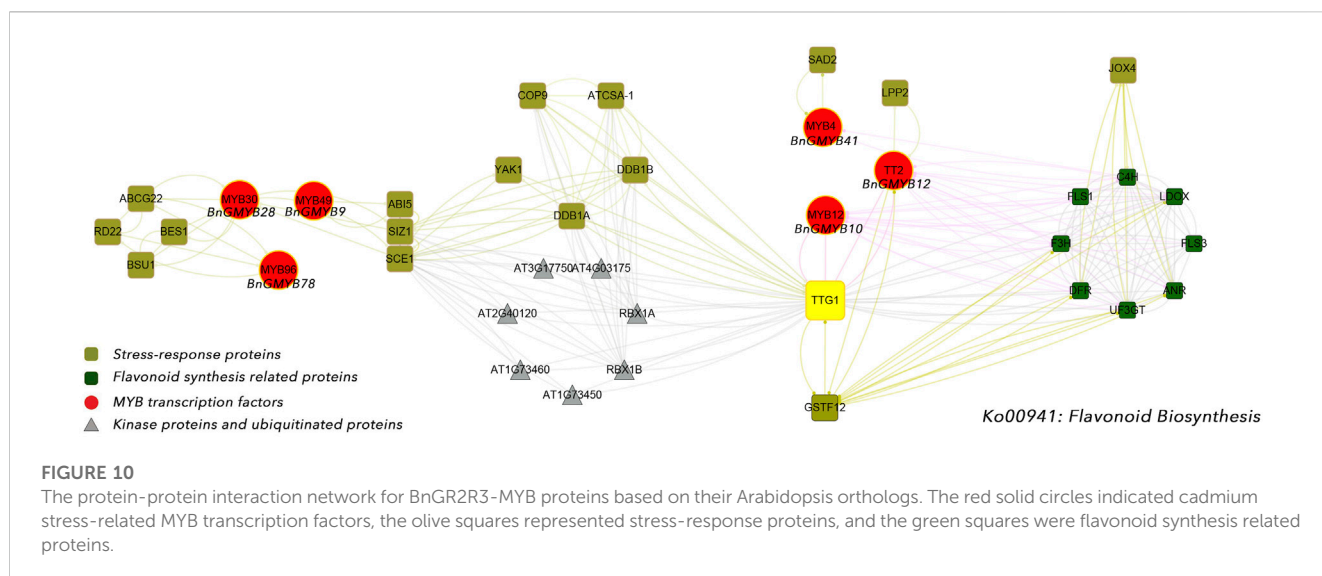
ramie MYB was found in the B1 subfamily that positively regulates anthocyanin, and therefore anthocyanin synthesis in ramie is not regulated by the B1 subgroup MYB transcription factors. It will be envisaged that, upregulation of *BnGMYB60* (B2), *BnGMYB79/80*



(B3), and *BnGMYB70* (B4) caused the inhibition of ZZ_1 anthocyanin synthesis, which ultimately affected color formation.

Expression profiles of BnGMYBs under Cd²⁺ stress

Ramie is a promising plant for remediation of heavy metal contaminated land and has high cadmium tolerance and accumulation capacity (She et al., 2015). An increasing number of studies have linked R2R3-MYB genes to various stress responses and regulation. However, there was not much information on the



involvement of MYB genes in cadmium tolerance response in ramie. We thus screened 11 MYBs candidates by combining RNA-seq expression profiles, homologous gene relationships, as well as their expression patterns under cadmium stress.

We analyzed them in three dimensions: treatment time, species, and tissue site. The results (Figure 9) showed varied expression patterns of the BnGMYB genes under different Cd²⁺ treatment times with four different trends of continuously rising, rise and fall, continuously falling, and rising after falling. Majority of the genes showed characteristics fall upon treatment before rising. The expression tendency of 11 genes were consistent in the two ramie cultivars, while some genes responded more strongly in HX_1 than in ZZ_1, such as *BnGMYB28*, *BnGMYB9*, *BnGMYB78*, indicating that the response mechanisms to Cd²⁺ stress is similar in the two cultivars. The degree of response in different parts of the plant under the same stress was also different, however, all the genes were upregulated in the leaves on cadmium treatment. Combining the results of different treatment times and tissue sites, we found that genes such as *BnGMYB9/11/41/89* showed a bottom-up response pattern in the order (root-stem-leaf) which may relate to cadmium transport and stress signaling.

The expression of *BnGMYB10/12/41* in roots, stems and leaves all increased more than 10-fold after cadmium stress, and phylogenetic analysis placed these three genes into the flavonoid synthesis MYB subgroup (C1), suggesting a potential association between flavonoid synthesis and cadmium tolerance (Dubos et al., 2010). For further understanding the association of flavonoid synthesis and cadmium stress, we selected *AtMYB4* (Agarwal et al., 2020), *AtMYB30* (Fichman et al., 2020), *AtMYB49* (Zhang et al., 2019), and *AtMYB96* (Seo et al., 2011) which are respectively homologs of *BnGMYB41*, *BnGMYB28*, *BnGMYB9*, *BnGMYB78* with established role in defense responses, and *AtMYB4* (Agarwal et al., 2020), *AtMYB12* (Mehrtens et al., 2005), *TT2* (Baudry et al., 2004) (homologs of *BnGMYB10/BnGMYB12/BnGMYB41*, respectively) involved in regulating flavonoid synthesis pathway in *Arabidopsis thaliana* for protein interaction network analysis (Figure 10). These MYB TFs were found to interact with flavonoid biosynthetic genes, in addition to being tightly linked to some

resistance proteins. *BnGMYB12* and *BnGMYB41* interacted with *LPP2* (lipid phosphatase 2) (Pierrugues et al., 2001), *SAD2* (Zhao et al., 2007) (sensitive protein to ABA and drought 2), respectively. A number of proteins in the flavonoid metabolic pathway have been reported to interact with jasmonic acid-induced oxygenase 4 (*JOX4*) (Smirnova et al., 2017). We also found that *BnGMYB9/28/78* interacts with many important stress resistance proteins such as *SIZ1* (Miura et al., 2020), *RD22* (Ma et al., 2018), *YAK1* (Kim et al., 2016), *ABI5* (Ma et al., 2018), etc. Interestingly, we found an interface between the flavonoid metabolic pathway and cadmium stress, the *TGG1* (thioglucoside glucohydrolase 1) node connecting two possible stress coping mechanisms. This finding provides a basis for understanding the association of flavonoid synthesis and stress regulation in ramie.

Conclusion

A total of 105 BnGR2R3-MYBs unevenly distributed among 14 chromosomes, phylogenetically divided into 35 distinct subfamilies, were identified in ramie. The segmental duplication events especially prominent in distal telomeric regions, played a crucial role in the expansion of BnGR2R3-MYB gene family. The result showed highest orthologous gene pairs between ramie and *Apocynum venetum* (58) than other plants including *Arabidopsis*. Expression analysis led to the identification of tissue preferential and cadmium stress responsive expression patterns of the BnGR2R3-MYB genes. Additionally, putative functions of ramie MYB genes were assigned based on the phylogenomic results and gene expression data. We performed comparative transcriptome analysis between HX_1 and ZZ_1 coupled with phylogenetic and UPLC-QTOF-MS analysis and identified *BnGMYB60*, *BnGMYB79/80* and *BnGMYB70* as potential anthocyanin biosynthesis repressor genes. Expression profiling and analysis of some selected BnGMYB revealed *BnGMYB9/10/12/28/41/78* as potential cadmium stress regulators and *BnGMYB10/12/41* in addition, as regulatory genes of

flavonoid biosynthesis which is in accordance their *Arabidopsis* annotated homologues. Thus, this study provided significant information into MYB regulatory genes in ramie and may serve as a foundation for genetic enhancement and increased productivity.

Materials and methods

Identification of MYB protein in *Boehmeria nivea* (L.) gaudich

For the comprehensive identification of the ramie MYB TFs, 168 MYB family sequences and 97 MYB-related family sequences (*A. thaliana*), 208 MYB family sequences and 85 MYB-related family sequences (*G. arboreum*), 130 MYB family sequences and 106 MYB-related family sequences (*Oryza sativa*) were retrieved from the PlantTFDB (<http://planttfdb.gao-lab.org/>). The protein sequences of these species were then used as reference to filter the possible sequence of target specie by Blast Wrapper, with expectation cut-off value (E-value) of e^{-5} set as threshold significance. The sequences obtained were aligned to Swissport Database (<https://www.expasy.org/resources/uniprotkb-swiss-prot/>)/Reference Species Whole Protein Sequence Library (<https://www.ncbi.nlm.nih.gov/refseq/>) and screened by annotation information. The resulting datasets of MYB sequences were confirmed based on the completeness of MYB domains using Pfam, CDD and SMART with an E-value set at 0.01. To correct for deletion of some conserved sites, FGENESH-M (<http://www.softberry.com/berry.phtml>) was used to predict multiple variants potential genes in genomic DNA. All the resulting ramie R2R3-MYB proteins were manually inspected to ensure that it contained two complete MYB domains and ultimately identified and classified as members of the MYB family in *Boehmeria nivea*. Length of sequences, molecular weights, GRAVY and pI of the MYB proteins were obtained using the ExPASy online tool (Wilkins et al., 1999) and subcellular localization predicted using the pLoc-mPlant tools for Batch Prediction.

Sequence analysis and structural characterization of BnGR2R3-MYB genes

Multiple sequence alignments of the MYB domains sequences were performed using MEGA-X with default parameters. The deduced amino acid sequences of the MYB motifs were adjusted manually in Jalview software (Waterhouse et al., 2009) and WEBLOGO (Crooks et al., 2004) used to show up the sequence logos of R2 and R3 MYB domain repeats. The exon-intron organizations of the BnGR2R3-MYB genes, including intron distribution patterns, phases, intro-exon boundaries and highlighted region of the MYB domains were graphically displayed with the aid of TBtools (Chen et al., 2020). The Conserved motifs of the ramie MYB proteins were predicted by using the MEME (Bailey et al., 2009) version 5.1.0 with

optimized parameters: zero or one per sequence; maximum number of motifs set at 20, and visualized in TBtools.

Analyses of chromosome distribution, gene duplication and synteny for BnGR2R3-MYB genes

The detailed chromosome distribution information of each BnGR2R3-MYB genes was obtained from the ramie genome annotation documents. Circos (Krzywinski et al., 2009) was used to locate all the BnGR2R3-MYB genes on the ramie chromosome. Tandem, segmental duplication and collinearity within species were obtained by using Multiple Collinearity Scantoolkit (MCscanX) (Wang et al., 2012). Synteny blocks of the orthologous R2R3-MYB genes between ramie and other species were obtained also using MCscanX and both results visualized in TBtools (Chen et al., 2020). Non-synonymous and synonymous substitution of each duplicated BnGR2R3-MYB genes were calculated using KaKs_Calculator 2.0 (Wang et al., 2010).

Phylogenetic analysis and classification of ramie BnGR2R3-MYB proteins and WGD analysis

Multiple sequence alignments of the R2R3-MYB proteins from ramie and *Arabidopsis thaliana* were performed using MEGA 7.0 with default parameters (Kumar et al., 2016) and neighborhood linkage (Guo et al., 2019) phylogenetic tree constructed with the following parameters: Poisson model; pairwise deletion; and 1,000 bootstrap replications. R2R3-MYB proteins from other species [*Arabidopsis* (Holub, 2001), rice (Katiyar et al., 2012), maize (Du et al., 2012b), tomato (Li et al., 2016), pineapple (Liu et al., 2017)] were obtained based on the description in corresponding literatures, and adopted same.

To establish WGD events, paralogous genes of ramie were detected using all-vs-all homology searches in BLASTP with an E-value threshold of $1e^{-5}$. Syntenic blocks within a genome were identified based on the detected homologous gene pairs using MCscan.

Identification of the cis-elements of BnGR2R3-MYB genes and GO annotation

The promoter regions of the BnGR2R3-MYB genes, 1,000 bp genomic DNA upstream sequences of each of the 105 BnGR2R3-MYB were selected, and the cis-elements predicted using PlantCare (http://bioinformatics.psb.ugent.be/web_tools/plantcare/html/). Response class elements such as light-responsive, plant growth, stress-responsive and phytohormone-responsive were filtered and the cis-acting elements visualized in TBtools, and displayed with a heatmap. Following blasting in BLASTP against Swissport database,

functional annotation of BnGMYB proteins was performed in Blast2GO Tool and subsequent mapping Gene Ontology (GO) terms, and visualization with R ggplot2.

Expression patterns of BnGR2R3-MYB genes in the representative tissues of the ramie

Transcriptome data of 9 samples comprising phloem, root and leaf tissues as well as, developmental periods and growth condition (hydroponics or terrestrial) obtained from Genome Sequence Archive database were used. Transcript abundance of the BnGR2R3-MYB genes was calculated as fragments per kilobase of exon model per million mapped reads (FPKM). $\text{Log}_2(\text{FPKM} + 1)$ values were displayed according to the color code. Detailed FPKM values were listed in [Supplementary Table S7](#).

Determination of anthocyanin content by UPLC-QTOF-MS

Anthocyanin content was determined following the method of [Wang et al. \(2020\)](#). Leaves samples of HX_1 and ZZ_1 (0.4 g each) was separately extracted at 4°C for 12 h using methanol/formic acid (4 mL) mixture (9:1, v/v). The extracts were then centrifuged for 20 min at 3,900 rpm and the supernatant filtered using a 0.22 µm nylon membrane into brown glass tube. The UPLC-QTOF-MS analysis was performed on Waters ACQUITY UPLC I-Class-Xevo G2-XS QTOF/PDA eLambda Detector equipped with an ACQUITY UPLC CSH C18 (2.1 mm × 100 mm i.d., 1.7 µm) column. Acetonitrile with 5% formic acid and water with 5% formic acid were respectively used as mobile phase A and B at 190 nm–800 nm wavelength photo-diode array (PDA). It was operated at positive ion mode for mass and the injection volume set at 1 µL for all samples. ESI source parameters were according to literature ([Wang et al., 2020](#)) and MS/MS for fragments ranged from 50 m/z to 1,000 m/z in a continuum mode. Ramp was in the range of 6 V–80 V. Peonidin-3-glucoside chloride (15 µg) was used as reference standard.

Stress treatment (Cd^{2+}) of ramie under hydroponic conditions

Two ramie varieties (ZZ_1; HX_1) were used for cadmium stress treatment. The two species have conspicuous leaf color attributed by different anthocyanin content ([Feng et al., 2021](#)) Shoots from each of the varieties at similar growth stage during the same period were selected for hydroponic cutting according to [Gao et al. \(2018\)](#). These were monitored for 15 days and seedlings with inconsistent growth were removed. CdCl_2 to a final concentration of 50 mg/L was added and three separate biological replicates were sampled at 0, 12, 24, and 48 h after the treatments. All samples were frozen quickly in liquid nitrogen and stored at –80°C until used.

BnGMYB protein–protein interaction network

Orthologous relationships of eight selected cadmium responsive MYB regulatory genes were determined between *A. thaliana* and ramie using OrthoVeen2 ([Xu et al., 2019](#)). Prediction of interactions between BnGMYB proteins and other proteins based on the Arabidopsis homologs was obtained using the online program STRING version 11.5 with high confidence >0.700 ([Szklarczyk et al., 2017](#)), filtered genes were used to construct the correlation network. The interaction network was visualized in Cytoscape v3.8.2.

RNA isolation and expression analysis of BnR2R3-GMYB genes

Total RNA was extracted using EasySpin Plus plant RNA rapid extraction Kit (Aid-lab Biotechnologies Co., Ltd.). The RNA was reverse-transcribed (Thermo Scientific RevertAid First cDNA Synthesis Kit) into cDNA and quantitative RT-PCR (qPCR) analysis conducted using gene specific primers ([Supplementary Table S8](#)). 18 s gene was used as internal control. The qPCR was conducted using SYBR Green Premix Pro Taq HS qPCR Kit (Accurate Biotechnology Co., Ltd.) on a CFX96 Touch Deep Well Real-Time PCR System (Bio-Rad) according to standard procedure. Relative transcript levels were calculated using the $2^{-\Delta\Delta C_t}$ formula and the result displayed using histograms drawn with GraphPad Prism 8 software and all the histograms merged using Adobe Photoshop (2020) software. All qPCR analyses were performed with three biological and technical replications.

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

PC and XF conceived and designed the experiments; XF, AZ, JC, XW, and PM performed the experiments; XF, CY, KC, and GG, analyzed the data; XF, AA, and PC wrote and reviewed the paper. All authors have read and approved the manuscript.

Funding

This work was supported by the Agricultural Science and Technology Innovation Project of the Chinese Academy of Agricultural Sciences (CAAS-ASTIP-2020), the China Agriculture Research System of MOF and MARA (CARS-16-E03).

Acknowledgments

We would like to thank the Institute of bast fiber crops, Chinese Academy of Agricultural Sciences for kindly supplying the experimental material.

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.

Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fgene.2023.1080909/full#supplementary-material>

SUPPLEMENTARY FIGURE S1

Phylogenetic relationships (neighbor-joining), gene structure and architecture of conserved protein motifs in BnGR2R3-MYB genes.

SUPPLEMENTARY FIGURE S2

Phylogenetic analysis of the R2R3-MYB family in six different plant species.

SUPPLEMENTARY FIGURE S3

Comparative genomic analysis—WDGI. The point with the best homology is red (The highest similarity). The diagonals are not self-alignments of genes, but tandem repeats.

References

- Agarwal, P., Mitra, M., Banerjee, S., and Roy, S. (2020). MYB4 transcription factor, a member of R2R3-subfamily of MYB domain protein, regulates cadmium tolerance via enhanced protection against oxidative damage and increases expression of PCS1 and MT1C in *Arabidopsis*. *Plant Sci.* 297, 110501. doi:10.1016/j.plantsci.2020.110501
- Akagi, T., Ikegami, A., Tsujimoto, T., Kobayashi, S., Sato, A., Kono, A., et al. (2009). DkMyb4 is a Myb transcription factor involved in proanthocyanidin biosynthesis in persimmon fruit. *Plant Physiol.* 151 (4), 2028–2045. doi:10.1104/pp.109.146985
- Arce-Rodríguez, M. L., and Ochoa-Alejo, N. (2017). An r2r3-MYB transcription factor regulates capsaicinoid biosynthesis. *Plant Physiol.* 174 (3), 1359–1370. doi:10.1104/pp.17.00506
- Ashburner, M., Ball, C. A., Blake, J. A., Botstein, D., Butler, H., Cherry, J. M., et al. (2000). Gene ontology: Tool for the unification of biology. The gene ontology consortium. *Nat. Genet.* 25 (1), 25–29. doi:10.1038/75556
- Bac-Molenaar, J. A., Fradin, E. F., Rienstra, J. A., Vreugdenhil, D., and Keurentjes, J. J. (2015). GWA mapping of anthocyanin accumulation reveals balancing selection of MYB90 in *Arabidopsis thaliana*. *PLoS one* 10 (11), e0143212. doi:10.1371/journal.pone.0143212
- 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, W202–W208. Web Server issue. doi:10.1093/nar/gkp335
- Baldoni, E., Genga, A., and Cominelli, E. (2015). Plant MYB transcription factors: Their role in drought response mechanisms. *Int. J. Mol. Sci.* 16 (7), 15811–15851. doi:10.3390/ijms160715811
- Ballester, A. R., Molthoff, J., de Vos, R., Hekkert, B., Orzaez, D., Fernández-Moreno, J. P., et al. (2010). Biochemical and molecular analysis of pink tomatoes: Deregulated expression of the gene encoding transcription factor SlMYB12 leads to pink tomato fruit color. *Plant Physiol.* 152 (1), 71–84. doi:10.1104/pp.109.147322
- Baudry, A., Heim, M. A., Dubreucq, B., Caboche, M., Weisshaar, B., and Lepiniec, L. (2004). TT2, TT8, and TTG1 synergistically specify the expression of BANYULS and proanthocyanidin biosynthesis in *Arabidopsis thaliana*. *Plant J. Cell. Mol. Biol.* 39 (3), 366–380. doi:10.1111/j.1365-313X.2004.02138.x
- Cannon, S. B., Mitra, A., Baumgarten, A., Young, N. D., and May, G. (2004). The roles of segmental and tandem gene duplication in the evolution of large gene families in *Arabidopsis thaliana*. *BMC plant Biol.* 4, 10. doi:10.1186/1471-2229-4-10
- Cavallini, E., Matus, J. T., Finezzo, L., Zenoni, S., Loyola, R., Guzzo, F., et al. (2015). The phenylpropanoid pathway is controlled at different branches by a set of R2R3-MYB C2 repressors in grapevine. *Plant Physiol.* 167 (4), 1448–1470. doi:10.1104/pp.114.256172
- Chen, T., Li, W., Hu, X., Guo, J., Liu, A., and Zhang, B. (2015). A cotton MYB transcription factor, GbMYB5, is positively involved in plant adaptive response to drought stress. *Plant Cell. Physiol.* 56 (5), 917–929. doi:10.1093/pcp/pcv019
- Chen, Q., Liu, Y., Maere, S., Lee, E., Van Isterdael, G., Xie, Z., et al. (2015). A coherent transcriptional feed-forward motif model for mediating auxin-sensitive PIN3 expression during lateral root development. *Nat. Commun.* 6, 8821. doi:10.1038/ncomms9821
- Chen, S., Niu, X., Guan, Y., and Li, H. (2017). Genome-wide analysis and expression profiles of the MYB genes in *brachypodium distachyon*. *Plant Cell. Physiol.* 58 (10), 1777–1788. doi:10.1093/pcp/pcx115
- 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 (8), 1194–1202. doi:10.1016/j.molp.2020.06.009
- Chen, X., Wang, P., Gu, M., Lin, X., Hou, B., Zheng, Y., et al. (2021). R2R3-MYB transcription factor family in tea plant (*Camellia sinensis*): Genome-wide characterization, phylogeny, chromosome location, structure and expression patterns. *Genomics* 113 (3), 1565–1578. doi:10.1016/j.ygeno.2021.03.033
- Chu, C. C., Lee, W. C., Guo, W. Y., Pan, S. M., Chen, L. J., Li, H. M., et al. (2005). A copper chaperone for superoxide dismutase that confers three types of copper/zinc superoxide dismutase activity in *Arabidopsis*. *Plant Physiol.* 139 (1), 425–436. doi:10.1104/pp.105.065284
- Clemens, S., and Ma, J. F. (2016). Toxic heavy metal and metalloid accumulation in crop plants and foods. *Annu. Rev. plant Biol.* 67, 489–512. doi:10.1146/annurev-arplant-043015-112301
- Crooks, G. E., Hon, G., Chandonia, J. M., and Brenner, S. E. (2004). WebLogo: A sequence logo generator. *Genome Res.* 14 (6), 1188–1190. doi:10.1101/gr.849004
- Du, H., Yang, S. S., Liang, Z., Feng, B. R., Liu, L., Huang, Y. B., et al. (2012). Genome-wide analysis of the MYB transcription factor superfamily in soybean. *BMC plant Biol.* 12, 106. doi:10.1186/1471-2229-12-106
- Du, H., Feng, B. R., Yang, S. S., Huang, Y. B., and Tang, Y. X. (2012). The R2R3-MYB transcription factor gene family in maize. *PLoS one* 7 (6), e37463. doi:10.1371/journal.pone.0037463
- Dubos, C., Stracke, R., Grotewold, E., Weisshaar, B., Martin, C., and Lepiniec, L. (2010). MYB transcription factors in *Arabidopsis*. *Trends plant Sci.* 15 (10), 573–581. doi:10.1016/j.tplants.2010.06.005
- Fasani, E., DalCorso, G., Costa, A., Zenoni, S., and Furini, A. (2019). The *Arabidopsis thaliana* transcription factor MYB59 regulates calcium signalling during plant growth and stress response. *Plant Mol. Biol.* 99 (6), 517–534. doi:10.1007/s11103-019-00833-x
- Feller, A., Machemer, K., Braun, E. L., and Grotewold, E. (2011). Evolutionary and comparative analysis of MYB and bHLH plant transcription factors. *Plant J. Cell. Mol. Biol.* 66 (1), 94–116. doi:10.1111/j.1365-313X.2010.04459.x
- Feng, X., Gao, G., Yu, C., Zhu, A., Chen, J., Chen, K., et al. (2021). Transcriptome and metabolome analysis reveals anthocyanin biosynthesis pathway associated with ramie (*Boehmeria nivea* (L.) Gaud.) leaf color formation. *BMC genomics* 22 (1), 684. doi:10.1186/s12864-021-08007-0
- Fichman, Y., Zandalinas, S. I., Sengupta, S., Burks, D., Myers, R. J., Jr., Azad, R. K., et al. (2020). MYB30 orchestrates systemic reactive oxygen signaling and plant acclimation. *Plant Physiol.* 184 (2), 666–675. doi:10.1104/pp.20.00859
- Gahlaut, V., Jaiswal, V., Kumar, A., and Gupta, P. K. (2016). Transcription factors involved in drought tolerance and their possible role in developing drought tolerant cultivars with emphasis on wheat (*Triticum aestivum* L.). *TAG Theor. Appl. Genet. Theor. und angewandte Genet.* 129 (11), 2019–2042. doi:10.1007/s00122-016-2794-z
- Gao, G., Xiong, H., Chen, J., Chen, K., Chen, P., Yu, C., et al. (2018). Hydroponic method for ramie and removal of nitrogen and phosphorus from livestock wastewater. *Int. J. phytoremediation* 20 (6), 545–551. doi:10.1080/15226514.2017.1393395

- Guo, Y., Zhu, C., Zhao, S., Zhang, S., Wang, W., Fu, H., et al. (2019). De novo transcriptome and phytochemical analyses reveal differentially expressed genes and characteristic secondary metabolites in the original oolong tea (*Camellia sinensis*) cultivar 'Tieguanyin' compared with cultivar 'Benshan'. *BMC genomics* 20 (1), 265. doi:10.1186/s12864-019-5643-z
- Holub, E. B. (2001). The arms race is ancient history in *Arabidopsis*, the wildflower. *Nat. Rev. Genet.* 2 (7), 516–527. doi:10.1038/35080508
- Hsu, Y. T., and Kao, C. H. (2003). Role of abscisic acid in cadmium tolerance of rice (*Oryza sativa* L.) seedlings. *Plant, Cell. & Environ.* 26 (6), 867–874. doi:10.1046/j.1365-3040.2003.01018.x
- Huang, K. Y., Zhu, A. G., Chen, X. R., Shi, Y. L., Tang, Q., Wang, X. F., et al. (2019). Comparative transcriptomics reveals the selection patterns of domesticated ramie. *Ecol. Evol.* 9 (12), 7057–7068. doi:10.1002/ece3.5271
- Huang, D., Tang, Z., Fu, J., Yuan, Y., Deng, X., and Xu, Q. (2020). CsMYB3 and CsRuby1 form an 'activator-and-repressor' loop for the regulation of anthocyanin biosynthesis in citrus. *Plant Cell. Physiol.* 61 (2), 318–330. doi:10.1093/pcp/pcz198
- Jian, W., Cao, H., Yuan, S., Liu, Y., Lu, J., Lu, W., et al. (2019). SIMYB75, 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, 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 (7), R46. doi:10.1186/gb-2004-5-7-r46
- Jiang, C. K., and Rao, G. Y. (2020). Insights into the diversification and evolution of r2r3-MYB transcription factors in plants. *Plant Physiol.* 183 (2), 637–655. doi:10.1104/pp.19.01082
- 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
- Kim, D., Ntui, V. O., and Xiong, L. (2016). *Arabidopsis* YAK1 regulates abscisic acid response and drought resistance. *FEBS Lett.* 590 (14), 2201–2209. doi:10.1002/1873-3468.12234
- Kiprioti, M., Alexopoulou, E., Vafeiadakis, T., and Kipriotis, H. (2015). *Crops XII, Ramie and kenaf as feed crops*.
- 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. Cell. Mol. Biol.* 16 (2), 263–276. doi:10.1046/j.1365-313x.1998.00278.x
- Krzywinski, M., Schein, J., Birol, I., Connors, J., Gascoyne, R., Horsman, D., et al. (2009). Circos: An information aesthetic for comparative genomics. *Genome Res.* 19 (9), 1639–1645. doi:10.1101/gr.092759.109
- Kumar, S., Stecher, G., and Tamura, K. (2016). MEGA7: Molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Mol. Biol. Evol.* 33 (7), 1870–1874. doi:10.1093/molbev/msw054
- Li, Z., Peng, R., Tian, Y., Han, H., Xu, J., and Yao, Q. (2016). Genome-wide identification and analysis of the MYB transcription factor superfamily in *Solanum lycopersicum*. *Plant Cell. Physiol.* 57 (8), 1657–1677. doi:10.1093/pcp/pcw091
- Li, J., Han, G., Sun, C., and Sui, N. (2019). Research advances of MYB transcription factors in plant stress resistance and breeding. *Plant Signal. Behav.* 14 (8), 1613131. doi:10.1080/15592324.2019.1613131
- Liu, C., Xie, T., Chen, C., Luan, A., Long, J., Li, C., et al. (2017). Genome-wide organization and expression profiling of the R2R3-MYB transcription factor family in pineapple (*Ananas comosus*). *BMC genomics* 18 (1), 503. doi:10.1186/s12864-017-3896-y
- Liu, C., Zeng, L., Zhu, S., Wu, L., Wang, Y., Tang, S., et al. (2018). Draft genome analysis provides insights into the fiber yield, crude protein biosynthesis, and vegetative growth of domesticated ramie (*Boehmeria nivea* L. Gaud.). *DNA Res. Int. J. Rapid Publ. Rep. Genes. genomes* 25 (2), 173–181. doi:10.1093/dnares/dsx047
- Liu, C., Hao, J., Qiu, M., Pan, J., and He, Y. (2020). Genome-wide identification and expression analysis of the MYB transcription factor in Japanese plum (*Prunus salicina*). *Genomics* 112 (6), 4875–4886. doi:10.1016/j.ygeno.2020.08.018
- 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 (9), 1431–1441. doi:10.1093/pcp/pcx075
- Luan, M. B., Jian, J. B., Chen, P., Chen, J. H., Chen, J. H., Gao, Q., et al. (2018). Draft genome sequence of ramie, *Boehmeria nivea* (L.) Gaudich. *Mol. Ecol. Resour.* 18 (3), 639–645. doi:10.1111/1755-0998.12766
- Lynch, M., and Conery, J. S. (2000). The evolutionary fate and consequences of duplicate genes. *Sci. (New York, NY)* 290 (5494), 1151–1155. doi:10.1126/science.290.5494.1151
- Ma, Q., Dai, X., Xu, Y., Guo, J., Liu, Y., Chen, N., et al. (2009). Enhanced tolerance to chilling stress in *Osmunda* transgenic rice is mediated by alteration in cell cycle and ectopic expression of stress genes. *Plant Physiol.* 150 (1), 244–256. doi:10.1104/pp.108.133454
- Ma, Q., Xia, Z., Cai, Z., Li, L., Cheng, Y., Liu, J., et al. (2018). GmWRKY16 enhances drought and salt tolerance through an ABA-mediated pathway in *Arabidopsis thaliana*. *Front. Plant Sci.* 9, 1979. doi:10.3389/fpls.2018.01979
- Mehrtrens, 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 (2), 1083–1096. doi:10.1104/pp.104.058032
- Ming, R., VanBuren, R., Wai, C. M., Tang, H., Schatz, M. C., Bowers, J. E., et al. (2015). The pineapple genome and the evolution of CAM photosynthesis. *Nat. Genet.* 47 (12), 1435–1442. doi:10.1038/ng.3435
- Miura, K., Renhu, N., and Suzaki, T. (2020). The PHD finger of *Arabidopsis* SIZ1 recognizes trimethylated histone H3K4 mediating SIZ1 function and abiotic stress response. *Commun. Biol.* 3 (1), 23. doi:10.1038/s42003-019-0746-2
- Mu, L., Cai, M., Wang, Z., Liu, J., Liu, T., Wanapat, M., et al. (2020). Assessment of ramie leaf (*Boehmeria nivea* L. gaud.) as an animal feed supplement in P.R. China. *Trop. animal health Prod.* 52 (1), 115–121. doi:10.1007/s11250-019-01997-w
- Muñoz-Gómez, S., Suárez-Baron, H., Alzate, J. F., González, F., and Pabón-Mora, N. (2021). Evolution of the subgroup 6 r2r3-MYB genes and their contribution to floral color in the perianth-bearing piperales. *Front. Plant Sci.* 12, 633227. doi:10.3389/fpls.2021.633227
- Ng, D. W., Aboensinghe, J. K., and Kamali, M. (2018). Regulating the regulators: The control of transcription factors in plant defense signaling. *Int. J. Mol. Sci.* 19 (12), 3737. doi:10.3390/ijms19123737
- Nguyen, N. H., and Cheong, J. J. (2018). H2A.Z-containing nucleosomes are evicted to activate AtMYB44 transcription in response to salt stress. *Biochem. biophysical Res. Commun.* 499 (4), 1039–1043. doi:10.1016/j.bbrc.2018.04.048
- Ogata, K., Kanei-Ishii, C., Sasaki, M., Hatanaka, H., Nagadoi, A., Enari, M., et al. (1996). The cavity in the hydrophobic core of Myb DNA-binding domain is reserved for DNA recognition and trans-activation. *Nat. Struct. Biol.* 3 (2), 178–187. doi:10.1038/nsb0296-178
- Paz-Ares, J., Ghosal, D., Wienand, U., Peterson, P. A., and Saedler, H. (1987). The regulatory c1 locus of *Zea mays* encodes a protein with homology to myb proto-oncogene products and with structural similarities to transcriptional activators. *EMBO J.* 6 (12), 3553–3558. doi:10.1002/j.1460-2075.1987.tb02684.x
- Pierrugues, O., Brutesco, C., Oshiro, J., Gouy, M., Deveaux, Y., Carman, G. M., et al. (2001). Lipid phosphate phosphatases in *Arabidopsis*. Regulation of the AtLPP1 gene in response to stress. *J. Biol. Chem.* 276 (23), 20300–20308. doi:10.1074/jbc.M009726200
- Riechmann, J. L., and Ratcliffe, O. J. (2000). A genomic perspective on plant transcription factors. *Curr. Opin. plant Biol.* 3 (5), 423–434. doi:10.1016/s1369-5266(00)00107-2
- Sacks, D., Baxter, B., Campbell, B. C. V., Carpenter, J. S., Cognard, C., Dippel, D., et al. (2018). Multisociety consensus quality improvement revised consensus statement for endovascular therapy of acute ischemic stroke. *Int. J. stroke official J. Int. Stroke Soc.* 13 (6), 612–632. doi:10.1177/1747493018778713
- Schilling, S., Kennedy, A., Pan, S., Jermini, L. S., and Melzer, R. (2020). Genome-wide analysis of M1KC-type MADS-box genes in wheat: Pervasive duplications, functional conservation and putative neofunctionalization. *New Phytologist* 225 (1), 511–529. doi:10.1111/nph.16122
- Seo, P. J., Lee, S. B., Suh, M. C., Park, M. J., Go, Y. S., and Park, C. M. (2011). The MYB96 transcription factor regulates cuticular wax biosynthesis under drought conditions in *Arabidopsis*. *Plant Cell.* 23 (3), 1138–1152. doi:10.1105/tpc.111.083485
- She, W., Zhu, S., Jie, Y., Xing, H., and Cui, G. (2015). Expression profiling of cadmium response genes in ramie (*Boehmeria nivea* L.) root. *Bull. Environ. Contam. Toxicol.* 94 (4), 453–459. doi:10.1007/s00128-015-1502-z
- Shim, J. S., and Choi, Y. D. (2013). Direct regulation of WRKY70 by AtMYB44 in plant defense responses. *Plant Signal. Behav.* 8 (6), e20783. doi:10.4161/psb.24509
- Si, W., Hang, T., Guo, M., Chen, Z., Liang, Q., Gu, L., et al. (2019). Whole-genome and transposed duplication contributes to the expansion and diversification of TLC genes in maize. *Int. J. Mol. Sci.* 20 (21), 5484. doi:10.3390/ijms20215484
- Smirnova, E., Marquis, V., Poirier, L., Aubert, Y., Zumsteg, J., Ménard, R., et al. (2017). Jasmonic acid oxidase 2 hydroxylates jasmonic acid and represses basal defense and resistance responses against botrytis cinerea infection. *Mol. plant* 10 (9), 1159–1173. doi:10.1016/j.molp.2017.07.010
- Smita, S., Katiyar, A., Chinnusamy, V., Pandey, D. M., and Bansal, K. C. (2015). Transcriptional regulatory network analysis of MYB transcription factor family genes in rice. *Front. Plant Sci.* 6, 1157. doi:10.3389/fpls.2015.01157
- Stracke, R., Werber, M., and Weisshaar, B. (2001). The R2R3-MYB gene family in *Arabidopsis thaliana*. *Curr. Opin. plant Biol.* 4 (5), 447–456. doi:10.1016/s1369-5266(00)00199-0
- Szklarczyk, D., Morris, J. H., Cook, H., Kuhn, M., Wyder, S., Simonovic, M., et al. (2017). The STRING database in 2017: Quality-controlled protein-protein association networks, made broadly accessible. *Nucleic Acids Res.* 45 (D1), D362–D368. doi:10.1093/nar/gkw937
- Tehseen, M., Cairns, N., Sherson, S., and Cobbett, C. S. (2010). Metallochaperone-like genes in *Arabidopsis thaliana*. *Metallomics Integr. biometal Sci.* 2 (8), 556–564. doi:10.1039/c003484c
- Thiel, T., Graner, A., Waugh, R., Grosse, I., Close, T. J., and Stein, N. (2009). Evidence and evolutionary analysis of ancient whole-genome duplication in barley predating the divergence from rice. *BMC Evol. Biol.* 9, 209. doi:10.1186/1471-2148-9-209

- Wang, D., Zhang, Y., Zhang, Z., Zhu, J., and Yu, J. (2010). KaKs_Calculator 2.0: A toolkit incorporating gamma-series methods and sliding window strategies. *Genomics, proteomics Bioinforma.* 8 (1), 77–80. doi:10.1016/S1672-0229(10)60008-3
- Wang, Y., Tang, H., Debarry, J. D., Tan, X., Li, J., Wang, X., et al. (2012). MCScanX: A toolkit for detection and evolutionary analysis of gene synteny and collinearity. *Nucleic acids Res.* 40 (7), e49. doi:10.1093/nar/gkr1293
- Wani, S. H., Tripathi, P., Zaid, A., Challa, G. S., Kumar, A., Kumar, V., et al. (2018). Transcriptional regulation of osmotic stress tolerance in wheat (*Triticum aestivum* L.). *Plant Mol. Biol.* 97 (6), 469–487. doi:10.1007/s11103-018-0761-6
- Wang, H., Sun, S., Zhou, Z., Qiu, Z., and Cui, X. (2020). Rapid analysis of anthocyanin and its structural modifications in fresh tomato fruit. *Food Chem.* 333, 127439. doi:10.1016/j.foodchem.2020.127439
- Wang, Y., Li, F., He, Q., Bao, Z., Zeng, Z., An, D., et al. (2021). Genomic analyses provide comprehensive insights into the domestication of bast fiber crop ramie (*Boehmeria nivea*). *Plant J. Cell. Mol. Biol.* 107, 787–800. doi:10.1111/tpj.15346
- Waterhouse, A. M., Procter, J. B., Martin, D. M., Clamp, M., and Barton, G. J. (2009). Jalview Version 2--a multiple sequence alignment editor and analysis workbench. *Bioinforma. Oxf. Engl.* 25 (9), 1189–1191. doi:10.1093/bioinformatics/btp033
- Wilkins, M. R., Gasteiger, E., Bairoch, A., Sanchez, J. C., Williams, K. L., Appel, R. D., et al. (1999). Protein identification and analysis tools in the ExPASy server. *Methods Mol. Biol. Clift. NJ* 112, 531–552. doi:10.1385/1-59259-584-7:531
- Williams, C. E., and Grotwold, 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 (1), 563–571. doi:10.1074/jbc.272.1.563
- Wu, Y., Wen, J., Xia, Y., Zhang, L., and Du, H. (2022). Evolution and functional diversification of R2R3-MYB transcription factors in plants. *Hortic. Res.* 9, uhac058. doi:10.1093/hr/uhac058
- Xu, L., Dong, Z., Fang, L., Luo, Y., Wei, Z., Guo, H., et al. (2019). OrthoVenn2: A web server for whole-genome comparison and annotation of orthologous clusters across multiple species. *Nucleic acids Res.* 47 (W1), W52–W58. doi:10.1093/nar/gkz333
- Xu, H., Liu, P., Wang, C., Wu, S., Dong, C., Lin, Q., et al. (2022). Transcriptional networks regulating suberin and lignin in endodermis link development and ABA response. *Plant physiol.* 190 (2), 1165–1181. doi:10.1093/plphys/kiac298
- Yamaguchi-Shinozaki, K., and Shinozaki, K. (2006). Transcriptional regulatory networks in cellular responses and tolerance to dehydration and cold stresses. *Annu. Rev. plant Biol.* 57, 781–803. doi:10.1146/annurev.arplant.57.032905.105444
- Yoshida, K., Ma, D., and Constabel, C. P. (2015). The MYB182 protein down-regulates proanthocyanidin and anthocyanin biosynthesis in poplar by repressing both structural and regulatory flavonoid genes. *Plant physiol.* 167 (3), 693–710. doi:10.1104/pp.114.253674
- Zhang, P., Wang, R., Ju, Q., Li, W., Tran, L. P., and Xu, J. (2019). The r2r3-MYB transcription factor MYB49 regulates cadmium accumulation. *Plant physiol.* 180 (1), 529–542. doi:10.1104/pp.18.01380
- Zhang, S., Chen, Y., He, X., Du, J., Zhang, R., Ma, Y., et al. (2020). Identification of MYB transcription factors regulating theanine biosynthesis in tea plant using omics-based gene coexpression analysis. *J. Agric. food Chem.* 68 (3), 918–926. doi:10.1021/acs.jafc.9b06730
- Zhao, J., Zhang, W., Zhao, Y., Gong, X., Guo, L., Zhu, G., et al. (2007). SAD2, an importin-like protein, is required for UV-B response in Arabidopsis by mediating MYB4 nuclear trafficking. *Plant Cell.* 19 (11), 3805–3818. doi:10.1105/tpc.106.048900
- Zhao, J., Zhou, H., and Li, Y. (2013). UBIQUITIN-SPECIFIC PROTEASE16 interacts with a HEAVY METAL ASSOCIATED ISOPRENYLATED PLANT PROTEIN27 and modulates cadmium tolerance. *Plant Signal. Behav.* 8 (10), e25680. doi:10.4161/psb.25680
- Zhou, B., Leng, J., Ma, Y., Fan, P., Li, Y., Yan, H., et al. (2020). BrmiR828 targets BrPAP1, BrMYB82, and BrTAS4 involved in the light induced anthocyanin biosynthetic pathway in Brassica rapa. *Int. J. Mol. Sci.* 21 (12), 4326–4342. doi:10.3390/ijms21124326
- Zimmermann, I. M., Heim, M. A., Weisshaar, B., and Uhrig, J. F. (2004). Comprehensive identification of *Arabidopsis thaliana* MYB transcription factors interacting with R/B-like BHLH proteins. *Plant J. Cell. Mol. Biol.* 40 (1), 22–34. doi:10.1111/j.1365-313X.2004.02183.x