The Caragana intermedia CiMYB15 Gene Positively Regulates Flavonoid Metabolism in Arabidopsis Postprint

Authors: Chai Wenjuan, Yang Qi, Li Guojing, Wang Ruigang

Date: 2018-08-06T00:00:00+00:00

Abstract

R2R3-MYB transcription factors are involved in the regulation of plant primary and secondary metabolism. In this study, we identified and cloned an R2R3-MYB gene from the drought transcriptome database of *Caragana intermedia* and designated it as *CiMYB15* (GenBank accession number MH678649); the coding region of *CiMYB15* was transformed into wild-type *Arabidopsis thaliana*, total flavonoid content was determined using spectrophotometry in wild-type and transgenic *Arabidopsis*, and the expression of the *AtCHS* gene in transgenic plants was detected by qRT-PCR. Simultaneously, the promoter sequence of *CiMYB15* was cloned using chromosome walking. The results showed that: (1) The gDNA length of *CiMYB15* is 1960 bp, containing three exons (134, 131, and 521 bp) and two introns (281 and 893 bp); the open reading frame is 786 bp in length, encoding 262 amino acids. (2) A 1580 bp promoter sequence was cloned, which mainly contains wound-responsive elements G-box and P-box, salt-responsive element GT1-motif, drought-inducible response element MBS, fungal elicitor-responsive element BOX-W1, and plant-pathogen interaction element EIER; additionally, it also contains binding sites for MYB transcription factors that regulate flavonoid biosynthesis genes. (3) The expression of *CiMYB15* is induced by UV stress. (4) The total flavonoid content in *CiMYB15*-overexpressing lines was higher than that in the wild type. (5) The expression level of the *AtCHS* gene in overexpressing plants was also higher than that in the wild type. These results demonstrate that *CiMYB15* positively regulates flavonoid metabolism in *Arabidopsis*.

Full Text

CiMYB15 from *Caragana intermedia* Positively Regulates Flavonoid Metabolism in *Arabidopsis*

Chai Wenjuan[1], Yang Qi[1,2,3], Li Guojing[1,2,3], Wang Ruigang[1,2,3*]

[1]College of Life Sciences, Inner Mongolia Agricultural University, Hohhot 010018, China
[2]Inner Mongolia Key Laboratory of Plant Stress Physiology and Molecular Biology, Hohhot 010018, China
[3]Inner Mongolia Scientific Innovation Team of Genetic Resource Utilization and Molecular Improvement of Stress-Resistant Plants, Hohhot 010018, China

Abstract

R2R3-MYB transcription factors participate in regulating primary and secondary metabolism in plants. In this study, we identified and cloned an R2R3-MYB gene from the drought-treated transcriptome of *Caragana intermedia*, designated as *CiMYB15* (GenBank accession number MH678649). The coding region of *CiMYB15* was transformed into wild-type *Arabidopsis thaliana*. Total flavonoid contents in wild-type and transgenic *Arabidopsis* were measured by spectrophotometry, and the expression of the *AtCHS* gene in transgenic plants was detected by qRT-PCR. Meanwhile, the promoter sequence of *CiMYB15* was cloned by genome walking. The results revealed that: (1) The *CiMYB15* genomic DNA (gDNA) was 1960 bp in length, consisting of three exons (134, 131, and 521 bp) and two introns (281 and 893 bp); the open reading frame (ORF) was 786 bp, encoding 262 amino acids. (2) A 1580 bp promoter fragment was cloned, containing mainly wound-inducible elements (G-box and P-box), salt-responsive element (GT1-motif), drought-inducible element (MBS), fungal elicitor-responsive element (BOX-W1), plant-pathogen interaction element (EIER), and MYB transcription factor binding sites for flavonoid biosynthesis genes. (3) The expression of *CiMYB15* was induced by UV stress. (4) Total flavonoid content in *CiMYB15* overexpression lines was higher than that in wild-type plants. (5) The expression level of *AtCHS* was also higher in transgenic plants than in wild-type. These results demonstrate that *CiMYB15* positively regulates flavonoid metabolism in *Arabidopsis*.

Keywords: *Caragana intermedia*, total flavonoids, *CiMYB15*, promoter, chalcone synthase gene

Introduction

The MYB family is an important transcription factor family in plants, named for the presence of one or more MYB domains composed of 51 or 52 amino acids.

MYB transcription factors are generally classified into four categories based on the number of MYB domain repeats: 1R-MYB/MYB-related, R2R3-MYB, R1R2R3-MYB, and 4R-MYB [1]. While R1R2R3-MYB genes are typically studied in animals, R2R3-MYB genes have been extensively investigated in plants [2].

The conserved domain of MYB genes typically exhibits a helix-helix-turn-helix structure. The helix-turn-helix (HTH) spatial structure formed by the second and third helices constitutes a hydrophobic core that can bind and stabilize DNA, with three uniformly distributed tryptophan residues playing a major role [3, 4]. In R2R3-MYB transcription factors, the first tryptophan in the R3 domain is sometimes replaced by phenylalanine (F), another aromatic amino acid, or by hydrophobic amino acids such as leucine (L) or isoleucine (I) [4]. A transcriptional activation domain, generally composed of numerous acidic amino acids, exists between the C-terminus and the conserved DNA-binding domain. Additionally, there is a poorly defined negative regulatory domain [1].

The first discovered v-MYB was cloned from avian myeloblastosis virus, containing three MYB domains [5]. Subsequently, increasing numbers of v-MYB homologous genes have been isolated from other animals, as well as from plants and fungi. Currently, multiple plant R2R3-MYB gene families have been identified with numerous members: 125 in *Arabidopsis thaliana* [6], 244 in soybean (*Glycine max*) [7], 109 in rice (*Oryza sativa*) [8], and 192 in black cottonwood (*Populus trichocarpa*) [9]. The highly diverse amino acid sequences in the C-terminal non-conserved domains of R2R3-MYB proteins determine that different MYB transcription factors within the same plant or homologous MYB transcription factors in different plant species have distinct functions. The numerous R2R3-MYB transcription factors are widely involved in regulating plant growth and development, primary and secondary metabolism, and responses to abiotic and biotic stresses [1]. Based on differences in C-terminal non-conserved region amino acids, the *Arabidopsis* R2R3-MYB transcription factor family is divided into 22 subgroups, with members within each subgroup typically having similar functions [6].

Flavonoids are low-molecular-weight polyphenolic secondary metabolites widely present in plants, playing important roles in disease and pest resistance, abiotic stress tolerance, flowering, fruiting, and pigmentation of organs such as seeds [10]. Flavonoid compounds mainly include flavonols, isoflavonoids, condensed tannins, and anthocyanins, with more than 9,000 flavonoid compounds identified to date [11]. Flavonoid biosynthesis exhibits specific spatiotemporal characteristics. The main genes affecting the flavonoid production pathway are of two types: structural genes and regulatory genes. Structural genes mainly include phenylalanine ammonia lyase (*PAL*), cinnamate 4-hydroxylase (*C4H*), 4-hydroxycinnamoyl-CoA ligase (*4CL*), chalcone synthase (*CHS*), and chalcone isomerase (*CHI*), among which *CHS* is a key rate-limiting enzyme in the flavonoid biosynthesis pathway [12].

The regulatory function is primarily performed by the MBW ternary complex

composed of bHLH, MYB, and WD40 proteins [13, 14]. MYB transcription factors mainly regulate flavonoid biosynthesis by binding to cis-acting elements in the promoters of key enzyme genes in the flavonoid biosynthesis pathway, thereby initiating or inhibiting the expression of certain genes. Previous promoter binding gel retardation assays have shown that most catalytic enzyme gene promoters in the phenylpropanoid pathway contain AC-rich motifs, and many R2R3-MYB transcription factors can specifically recognize and bind to AC-rich regions to activate these genes [15-17]. Subgroup 7 members of *Arabidopsis* R2R3-MYB—AtMYB11, AtMYB12, and AtMYB111—regulate flavonol metabolism by binding to genes encoding *CHS*, *CHI*, flavanone-3-hydroxylase (*F3H*), and flavonol synthase (*FLS*). The *myb11 myb12 myb111* triple mutant cannot accumulate flavonoids due to the absence of subgroup 7 MYB genes, though anthocyanin synthesis remains unaffected [18]. The *AcMYB1* gene, isolated from the flavonoid metabolic branch pathway in onion (*Allium cepa* L.), is an R2R3-MYB gene. Transient overexpression and RNAi experiments demonstrated that *AcMYB1* positively regulates anthocyanin synthesis in onion [19]. The *MsMYB12* and *MsMYB22* genes play key roles in proanthocyanidin and flavonol synthesis in red-fleshed apple (*Malus sieversii* f. *niedzwetzkyana*). MsMYB12 interacts with bHLH3 and bHLH33 to regulate proanthocyanidin synthesis, while MsMYB22 directly binds to the *FLS* gene promoter to initiate flavonol synthesis [20].

Caragana intermedia, belonging to the genus *Caragana* in the Fabaceae family and commonly known as "ningtiao," has a well-developed root system, small leaves, and is often hairy, providing excellent water retention. With low habitat requirements, it is commonly used as a windbreak and sand-fixation tree in northwest and north China [21]. Additionally, due to its rich secondary metabolites and flavonoid compounds characteristic of Fabaceae, it is also used medicinally [10, 22]. In this study, we obtained a sequence encoding an R2R3-MYB transcription factor from the drought transcriptome database of *C. intermedia*, cloned the cDNA, gDNA, and promoter sequences of this MYB gene, constructed an overexpression vector, and transformed it into wild-type *Arabidopsis*. Our findings demonstrate that this gene is involved in regulating flavonoid metabolism in *Arabidopsis*.

Materials and Methods

1.1 Plant Materials and Cultivation Seeds of *Caragana intermedia* were collected from Helinger County, Hohhot City, Inner Mongolia. Fully developed, naturally matured, and pest-free seeds were selected and sown in cultivation pots containing a substrate mixture of nutrient soil and vermiculite at a ratio of 1:3 to 1:5. Plants were grown for 20-25 days in a greenhouse at 22°C under a 16 h light/8 h dark photoperiod with light intensity of 7000-8000 lux. *Arabidopsis thaliana* Columbia ecotype (Col-0) was provided by our laboratory. *Arabidopsis* seeds were sterilized with 70% ethanol and 100% ethanol for 10 minutes each,

dried in a clean bench, resuspended in sterile water, and stratified at 4°C in darkness for 3 days before sowing in pots for subsequent experiments.

Healthy and uniform *C. intermedia* seedlings (30 days old) were selected for UV-B stress experiments and RNA/DNA extraction. Samples were collected at 0, 0.5, 1, 3, 6, 9, and 12 h. Shoots were excised, rapidly frozen in liquid nitrogen, and stored at -80°C. Healthy, unstressed, bolting *Arabidopsis* plants were selected, and flowers and siliques that had already set pods were removed.

1.2 Total RNA and DNA Extraction Total RNA from *C. intermedia* or *Arabidopsis* was extracted using TRIzol reagent (Invitrogen). DNA from *C. intermedia* was extracted using the CTAB method [23]. RNA quality was assessed by 1% agarose gel electrophoresis, and quantification was performed using a Q5000 micro-volume UV spectrophotometer (Quawell). Total RNA samples with clear bands, no smearing, and acceptable quantification results were used for first-strand cDNA synthesis.

1.3 First-Strand cDNA Synthesis and Gene Cloning After DNase treatment of total RNA samples, 1000 ng of *C. intermedia* total RNA or 500 ng of *Arabidopsis* total RNA was used for first-strand cDNA synthesis with RTase M-MLV reverse transcriptase (TaKaRa, Dalian Bao Biological Engineering Co., Ltd.) according to the kit protocol. Based on the *CiMYB15* gene sequence from the transcriptome, specific primers were designed for the ORF region using Primer Premier 5.0 software. Forward and reverse primers were appended with *Spe* I and *Sal* I restriction sites (indicated by lowercase letters) for subsequent overexpression vector construction (Table 1). The ORF was amplified from *C. intermedia* first-strand cDNA. Primer STAR high-fidelity polymerase, rTaq polymerase, and DNA markers (DL5000, 1 kb DNA ladder) were purchased from TaKaRa (Dalian Bao Biological Engineering Co., Ltd.). The PCR reaction mixture contained: 5× Primer STAR buffer (+Mg^2) 10 L, forward (*CiMYB15-HA5′*) and reverse (*CiMYB15-HA3′*) primers (10 mol/L) 2 L each, dNTPs (2.5 mmol/L) 4 L, template 1 L, Primer STAR polymerase 0.5 L, and sterile water 30.5 L. Cycling conditions were: 98°C for 1 min; 35 cycles of 98°C for 10 s, 60°C for 15 s, and 72°C for 1 min; and a final extension at 72°C for 10 min. Amplified products were analyzed by 1% agarose gel electrophoresis and sequenced by Shanghai Sangon Biotech Co., Ltd., where all primers were synthesized.

Table 1 Primers used in gene cloning and functional analysis

Primer name	Primer sequence	Usage
CiMYB15-HA5′	GCactagtTAAGATTCAGAGCTCTCAGCAATTGT	Amplification of ORF and gDNA

Primer name	Primer sequence	Usage
CiMYB15-HA3'	GCgtcgacATGGTTAGAGCTCCTTGCTTCGAA	Amplification of ORF and gDNA
q-*CiMYB15-5'*	CTGTAGACTGCGCTGGATTAACTATC	Real-time quantitative PCR
q-*CiMYB15-3'*	GTTCTTCCTGGTAACTTTGCTGAA	Real-time quantitative PCR
AtCHS-5'	GTACCTGTTTCCAAGCAACTCATGC	Real-time quantitative PCR
AtCHS-3'	CGCAGTCTACAGCTCTTTCCACTAC	Real-time quantitative PCR
AtCHI-5'	GAGGAAGTACCAATTACCAGCATCA	Real-time quantitative PCR
AtCHI-3'	GGTGCCATAGACGGACATT	Real-time quantitative PCR
AtF3H-5'	TCCATACTCGCTCAACACG	Real-time quantitative PCR
AtF3H-3'	CTCTATCTGTCAAGTGGAAGG	Real-time quantitative PCR
AtFLS-5'	GAAAACGCAACCGTAAGAG	Real-time quantitative PCR
AtFLS-3'	AGAGGCTTATGAGTTTGGC	Real-time quantitative PCR
AtDFR-5'	TGTAGCAGCAAGGTAATGG	Real-time quantitative PCR
AtDFR-3'	GGATTCTCTCGGATGGATTAG	Real-time quantitative PCR
AtEF1 -F	CGCCGATGTGAACAATGAC	Real-time quantitative PCR

Primer name	Primer sequence	Usage
AtEF1 -R	TGGTGGTCGGTCCATTCAT	Real-time quantitative PCR
CiEF1 -F	GAGAGAGCGCGGTGATAAGG	Real-time quantitative PCR
CiEF1 -R	AGAAGGGTGCCAAATGATGAG	Real-time quantitative PCR

1.4 Construction of *CiMYB15* Overexpression Recombinant Vector Recombinant plasmids with correct sequences were double-digested with *Spe* I and *Sal* I. The correct digested fragment, i.e., the ORF region of *CiMYB15*, was ligated to the linearized plant expression vector pCanG-HA driven by the CaMV35S promoter. The recombinant overexpression vector was verified by *Sal* I and *Sac* I digestion. The recombinant vector was introduced into *Agrobacterium tumefaciens* strain GV3101 by electroporation, and positive clones were verified by colony PCR.

1.5 Genetic Transformation of Wild-Type *Arabidopsis* and Screening of Transgenic Homozygous Lines Wild-type *Arabidopsis* was transformed by floral dip to obtain T generation transgenic plants. Transgenic plants were selected on 1/2 MS medium containing kanamycin (25 mg/L). Total RNA was extracted from wild-type and transgenic *Arabidopsis*, reverse-transcribed to cDNA, and the expression of *CiMYB15* in transgenic plants was detected by real-time quantitative PCR or RT-PCR. Primers for qRT-PCR (q-*CiMYB15-5* and q-*CiMYB15-3*) were designed based on the *CiMYB15* cDNA sequence. Using 16-fold diluted *Arabidopsis* cDNA as template and SYBR@ Green I fluorescent dye (TaKaRa), the expression level of *CiMYB15* in wild-type and overexpression lines was detected on a Roche LightCycler480 real-time PCR system. *Arabidopsis* elongation factor *AtEF1* (elongation factor 1-alpha) was used as the reference gene. The PCR reaction mixture contained: SYBR@ Green I 10 L, forward and reverse primers (q-*CiMYB15-5* , q-*CiMYB15-3* , 10 mol/L) 4 L each, sterile water 4.2 L, and template 5 L. Cycling conditions were: 95°C for 30 s; 40 cycles of 95°C for 5 s, 60°C for 30 s, and 72°C for 15 s. Melting curve analysis was performed after amplification, and data were analyzed using the 2^(-ΔCT) method. Each sample had three technical replicates.

1.6 Amplification of *CiMYB15* Promoter and Prediction of *cis*-Acting Elements The Genome Walking Kit (TaKaRa) was used to obtain the promoter sequence of *CiMYB15*. Based on the *CiMYB15* gDNA sequence, three antisense-specific primers (SP1, SP2, SP3) with high annealing temperatures (65-70°C) were designed, spaced approximately 100 bp apart. These were used

with four uniquely designed degenerate primers (AP1, AP2, AP3, AP4) with lower annealing temperatures for thermal asymmetric PCR. Using *C. intermedia* gDNA as template, three rounds of nested PCR were performed to obtain the flanking sequence of the known sequence. The third-round PCR products were gel-purified, ligated into the pMD19-T vector, transformed into *E. coli* competent cells, and positive clones were selected by colony PCR for sequencing. The obtained sequence was analyzed online using the PlantCARE database to predict promoter sequences.

1.7 Determination of Total Flavonoid Content Total flavonoids were extracted by ultrasonic-assisted extraction with 70% methanol. Healthy, one-month-old *Arabidopsis* plants without yellow leaves or lesions were selected, cut into pieces, and ground into homogenate at room temperature. Exactly 0.5 g of tissue was placed in a tube, and 70% methanol was added at a solid-liquid ratio of 1:20. Ultrasonic extraction was performed for 70 min at 60°C and 80 Hz. After overnight incubation at 4°C in darkness, samples were centrifuged and the supernatant was collected for total flavonoid content determination by spectrophotometry [24]. A standard curve was prepared using rutin standards purchased from Guizhou Dida Biological Technology Co., Ltd. Exactly 1.00 mg of rutin standard was dissolved in 70% methanol and diluted to 5 mL. Aliquots of 0, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, and 0.8 mL of rutin standard solution were mixed with 0.08 mL of 5% NaNO solution, vortexed, and incubated for 6 min. Then 0.08 mL of 10% Al(NO) solution was added, vortexed, and incubated for 6 min. Finally, 0.8 mL of 4% NaOH solution and 0.24 mL of ddH O were added, vortexed, and incubated for 15 min. Absorbance was measured at 510 nm and recorded, and a standard curve was generated in Excel. Experimental samples were processed using the same steps as for the standard curve.

Results

2.1 Cloning and Sequence Analysis of *CiMYB15* A MYB sequence was selected from the drought transcriptome database of *C. intermedia*. Analysis revealed that this sequence contained a complete open reading frame (ORF). Using specific primers and *C. intermedia* cDNA and gDNA as templates, PCR amplification was performed (Fig. 1a [Figure 1: see original paper], 1b).

Fig. 1 Gel electrophoresis of PCR products of *CiMYB15*. 1, 2: Amplified cDNA; 3, 4: Amplified gDNA; M: 1 kb DNA ladder.

Sequencing results showed that the ORF was 786 bp in length, with a start codon of ATG and a stop codon of TGA, encoding 262 amino acids (Fig. 2 [Figure 2: see original paper]). The gDNA was 1960 bp in length, containing three exons (134, 131, and 521 bp) and two introns (281 and 893 bp) (Fig. 2). BLAST searches in NCBI using the nucleotide and encoded amino acid sequences revealed that the gene showed highest similarity to *AtMYB15* in *Arabidopsis*, and

was therefore named *CiMYB15*.

Fig. 2 cDNA, gDNA, and deduced amino acid sequence of *CiMYB15*. Underlined portions indicate intron regions.

The predicted molecular weight of CiMYB15 protein was 30.24 kDa, with a theoretical isoelectric point of 5.84. Secondary structure prediction using HNN online tool showed that random coils accounted for the highest proportion at 66.79%, while -helices and -sheets accounted for 25.19% and 8.02%, respectively. Hydrophobicity analysis of the amino acid sequence using the ProtScale tool in the ExPASy database revealed that the grand average of hydropathy (GRAVY) was -0.909, indicating that the protein is overall hydrophilic.

2.2 Homology and Phylogenetic Analysis of CiMYB15 The CiMYB15 amino acid sequence was queried in NCBI BLAST, revealing high similarity with MYB amino acid sequences from other legume species. Multiple sequence alignment and conserved domain analysis of CiMYB15, *Arabidopsis* AtMYB15 (XP_020889427.1), soybean GmMYB29 (NP_001241360.1), and *Medicago truncatula* MtMYB51 (ABR28339.1) showed that CiMYB15 belongs to the typical R2R3-MYB transcription factor family, with the R2 domain spanning amino acids 13-63 and the R3 domain spanning amino acids 66-114 (Fig. 3 [Figure 3: see original paper]). A phylogenetic tree was constructed using MEGA5.0 software to cluster CiMYB15 with MYB proteins from other species. The results indicated that *C. intermedia* CiMYB15 clustered with *M. truncatula* MtMYB51 from the legume family, showing the closest phylogenetic relationship with 72% similarity; the closest relationship with the model plant *Arabidopsis* was with AtMYB15, with 50% similarity (Fig. 4 [Figure 4: see original paper]). AtMYB15 belongs to subgroup 2 of the *Arabidopsis* R2R3-MYB transcription factor family, whose members all contain the conserved C-terminal motif IDxSFW−MxFWFD, which is also present in CiMYB15 [6] (Fig. 3).

Fig. 3 Alignment of amino acid sequences of CiMYB15 and other plant MYBs. Sequences from top to bottom are from *C. intermedia*, *A. thaliana*, *G. max*, and *M. truncatula*. Black solid and dashed lines indicate MYB R2 and MYB R3 domains, respectively; black dots mark the IDxSFW−MxFWFD conserved motif.

Fig. 4 Phylogenetic analysis of CiMYB15 and other plant MYBs. The phylogenetic tree was constructed by the neighbor-joining method with 1000 bootstrap replications.

2.3.1 Cloning of *CiMYB15* Promoter Based on the cloned *CiMYB15* gDNA sequence, three antisense-specific primers were designed to obtain the upstream promoter sequence of *CiMYB15* using the Genome Walking Kit. The degenerate primers AP1, AP2, AP3, and AP4 provided in the kit were paired with the three specific primers for three rounds of nested PCR amplification.

Sequencing results indicated that the pairing of degenerate primer AP2 with specific primers successfully amplified a sequence upstream of the start codon ATG, with a length of 1580 bp (Fig. 5a [Figure 5: see original paper]). Specific primers were designed to verify the *CiMYB15* promoter (Fig. 5b).

Fig. 5 PCR products of *CiMYB15* promoter. 1: Product of the first extension; 2: Product of the second extension; 3, 4: Products of the third extension; 5, 6: Products of specific primers; M1: 1 kb DNA ladder; M2: DL5000.

2.3.2 Prediction of *cis*-Acting Elements in *CiMYB15* Promoter The *CiMYB15* promoter sequence was analyzed online (PlantCARE) to identify potential *cis*-acting elements. The analysis revealed that the *CiMYB15* promoter sequence contained the core promoter elements TATA-box and CAAT-box (Table 2). In addition to core elements, light-responsive *cis*-acting elements were most abundant, including ATC-motif, Box-4, MRE, GAG-motif, GAP-box, and MNF1. Furthermore, the promoter sequence contained *cis*-acting elements related to biotic and abiotic stresses, such as fungal elicitor-responsive element BOX-W1, plant-pathogen interaction element EIER, anaerobic induction-responsive element ARE, wound-inducible elements G-box and P-box, salt-responsive element GT1-motif, drought-inducible element MBS, salicylic acid-responsive element TCA-element, and jasmonic acid-responsive elements CGTCA-motif and TGACG-motif. Among these, the MBSI element is a MYB transcription factor binding site that regulates flavonoid biosynthesis genes. *Cis*-acting elements related to plant growth regulation were also present, such as meristem-specific expression element CAT-box and gibberellin-responsive element GARE-motif.

Table 2 Prediction analysis of the *CiMYB15* promoter

cis-Acting element name	Biological function	Sequence	Original species
TATA-box	Core promoter element	CAAAT	*Arabidopsis thaliana*
CAAT-box	Core promoter element	TGGTTT	*Arabidopsis thaliana*
ATC-motif	Light-responsive element	AGTAATCT	*Zea mays*
BOX-4	Light-responsive element	ATTAAT	*Spinacia oleracea*
BOX-W1	Fungal elicitor-responsive element	AACCTAA	*Petroselinum crispum*
CCAAT-box	MYBHv1 binding site	TTGACC	*Petunia hybrida*
TCA-element	Salicylic acid-responsive element	CAACGG	*Petroselinum crispum*

cis-Acting element name	Biological function	Sequence	Original species
CGTCA-motif	Jasmonic acid-responsive element	CCATCTTTTT	*Hordeum vulgare*
TGACG-motif	Jasmonic acid-responsive element	GAGAAGAATA	*Brassica oleracea*
EIER	Plant-pathogen interaction element	CGTCA	*Hordeum vulgare*
G-box	Wound-inducible/light-responsive element	TGACG	*Hordeum vulgare*
P-box	Wound-inducible element (cooperates with G-box)	TTCGACC	*Nicotiana tabacum*
CAT-box	Meristem-specific expression element	CACATGG	*Arabidopsis thaliana*
GAG-motif	Light-responsive element	AAAAAAC(A/C)GTTA	*Arabidopsis thaliana*
GAP-box	Light-responsive element	GTGCCC	*Zea mays*
GARE-motif	Gibberellin-responsive element	AGAGAGT	*Zea mays*
GT1-motif	Salt-responsive element	AAATGGACCA	*Oryza sativa*
MBS	Drought-inducible element	TCTGTTG	*Arabidopsis thaliana*
MBSI	Flavonoid biosynthesis gene regulation element	GGTTAA	*Arabidopsis thaliana*
MNF1	Light-responsive element	GTCAT	*Brassica oleracea*

2.4 Expression Analysis of *CiMYB15* Under UV Treatment Based on promoter and bioinformatics analysis of *CiMYB15*, UV-B stress was applied to *C. intermedia* to investigate its potential function. The transcriptional changes of *CiMYB15* under UV stress were detected by qRT-PCR. The results showed that *CiMYB15* expression was strongly induced by UV stress, exhibiting an initial increase followed by a decrease, peaking at 6 h with approximately 190-fold higher expression than the untreated control (Fig. 6 [Figure 6: see original paper]).

Fig. 6 *CiMYB15* gene expression analysis under UV-B treatment by qRT-PCR. *CiEF1* was selected as the reference gene. Results were calculated using

the 2^(-ΔΔCT) method.

2.5 Construction of pCanG-*CiMYB15* Recombinant Expression Vector and Identification of Overexpression Homozygous Lines To investigate the function of *CiMYB15* from *C. intermedia*, a recombinant binary expression vector pCanG-*CiMYB15* driven by the 35S strong promoter was constructed by double digestion. The *CiMYB15* coding region was amplified using specific primers *CiMYB15-HA5'* and *CiMYB15-HA3'* , yielding a 786 bp fragment (Fig. 1a). The fragment was ligated into the pEASY-Blunt-Simple cloning vector for sequencing, then digested with *Spe* I and *Sal* I and ligated to the linearized overexpression vector pCanG-HA. The *Spe* I and *Sal* I sites flank the target gene in the recombinant plasmid. Gel electrophoresis of the double-digested pCanG-*CiMYB15* vector detected an ~800 bp band, confirming successful construction of the overexpression vector (Fig. 7a [Figure 7: see original paper]).

Fig. 7 Construction of recombinant vector pCanG-*CiMYB15* and expression analysis of *CiMYB15* in transgenic lines. (a) Identification of the recombinant vector pCanG-*CiMYB15*: 1, vector control; 2, recombinant vector digested by *Spe* I and *Sal* I. (b) RT-PCR assay of *CiMYB15* transgenic lines: C-, negative control using wild-type *Arabidopsis* cDNA as template; C+, positive control using *C. intermedia* cDNA as template; other numbers indicate overexpression lines. (c) Expression level of *CiMYB15* in transgenic lines analyzed by qRT-PCR (*AtEF1* was selected as reference gene; results were calculated using the 2^(-ΔCT) method; Y-axis is logarithmic). M: 1 kb DNA ladder.

The pCanG-*CiMYB15* overexpression vector was transformed into wild-type *Arabidopsis* by floral dip to obtain T generation transgenic plants. Transgenic plants were screened on 1/2 MS medium containing 25 mg/L kanamycin to obtain five T generation homozygous lines. The expression of the target gene *CiMYB15* in transgenic plants was detected at the transcriptional level by RT-PCR and qRT-PCR (Fig. 7b, 7c). Three lines with relatively high expression levels—*CiMYB15*-3, *CiMYB15*-8, and *CiMYB15*-19—were selected for subsequent experiments.

2.6 Increased Total Flavonoid Content in *CiMYB15* Overexpression Plants Analysis of the *CiMYB15* promoter sequence revealed the presence of MYB transcription factor binding sites that regulate flavonoid biosynthesis genes, appearing three times. This suggested that *CiMYB15* might be involved in flavonoid compound metabolism. Therefore, total flavonoid content in plants was determined by methanol-assisted ultrasonic aluminum nitrate colorimetry. A rutin standard curve was first established using the aluminum nitrate colorimetric method, showing a linear relationship between rutin concentration (x) and absorbance (y) described by $y = 1.227x - 0.0038$ ($R^2 = 0.9984$), with good linearity in the range of 0-0.08 mg/mL (Fig. 8a [Figure 8: see original paper]). Based on this formula, the total flavonoid concentration in the filtrate was cal-

culated from the OD values measured for wild-type and each transgenic line, then converted to total flavonoid content per gram of fresh plant weight. The results showed that total flavonoid content in *CiMYB15* transgenic lines was significantly higher than in wild-type *Arabidopsis* (wild-type: 42.8 ± 6.17 g/g FW; *CiMYB15*-3: 52.77 ± 4.97 g/g FW; *CiMYB15*-8: 68.44 ± 9.17 g/g FW; *CiMYB15*-19: 53.65 ± 9.74 g/g FW) (Fig. 8b [Figure 8: see original paper]). Notably, *CiMYB15*-8, which showed the highest expression level of *CiMYB15*, also had the highest flavonoid content among the three lines, indicating that *CiMYB15* positively regulates flavonoid production with a dosage effect.

Fig. 8 Total flavonoid content of wild-type and *CiMYB15* transgenic *Arabidopsis* lines.

2.7 Overexpression of *CiMYB15* Affects Expression of Flavonoid Biosynthesis-Related Genes in *Arabidopsis* Since flavonoid content was altered in transgenic plants, qRT-PCR was performed using wild-type and transgenic *Arabidopsis* cDNA to further confirm the effect of *CiMYB15* on flavonoid biosynthesis pathway genes. The results showed that the expression of *AtCHS*, a key enzyme gene in the early stage of the flavonoid biosynthesis pathway, was significantly upregulated in *CiMYB15* overexpression lines (Fig. 9a [Figure 9: see original paper]), while the expression of other genes (*CHI*, *F3H*, *FLS*, and *DFR*) showed no obvious changes compared to wild-type (Fig. 9b [Figure 9: see original paper]). This suggests that *CiMYB15* may positively regulate flavonoid metabolism by specifically affecting the expression of the *CHS* gene at the transcriptional level.

Fig. 9 Detection of flavonoid metabolism-related genes in wild-type and *CiMYB15* transgenic lines. *AtEF1* was selected as the reference gene. Results were calculated using the 2^(-ΔΔCT) method.

Discussion

The plant R2R3-MYB transcription factor family, with its numerous members, is widely involved in plant growth and development, substance metabolism, and environmental stress responses. Flavonoid compounds are not only essential for plant environmental adaptation but also promote human health as medicinal supplements. The regulation of the flavonoid biosynthesis pathway depends on the MBW ternary complex composed of bHLH, MYB, and WD40 proteins.

In this study, we cloned the *CiMYB15* gene from the drought transcriptome database of *Caragana intermedia*. Multiple sequence alignment and phylogenetic analysis revealed that the C-terminus of CiMYB15 contains the IDxSFW−MxFWFD motif characteristic of subgroup 2 members of the R2R3-MYB family. Subgroup 2 members in *Arabidopsis* include *AtMYB13*, *AtMYB14*, and *AtMYB15*, which are primarily involved in resistance to abiotic stress [16]. Previous studies have shown that UV treatment can induce flavonol accumulation in

plants to minimize UV damage, and the key enzyme gene *CHS* in the flavonoid biosynthesis pathway is upregulated upon UV treatment [11]. In this study, we detected the expression of *CiMYB15* in UV-treated *C. intermedia* seedlings and found that its expression was strongly induced by UV stress, consistent with the expression pattern of the *CHS* gene.

Overexpression of *Arabidopsis AtMYB15* enhances drought resistance, and *AtMYB15* overexpression lines show more sensitive phenotypes to ABA during seed germination and stomatal movement compared to wild-type [25]. Additionally, AtMYB15 acts as a transcriptional repressor in cold signaling pathways, directly binding to the promoter region of *CBF* genes to negatively regulate their expression; meanwhile, MPK6 specifically interacts with AtMYB15, phosphorylating serine at position 168 and preventing AtMYB15 from binding to the *CBF3* promoter region [26]. In other species, homologous genes of *AtMYB15* are mainly involved in secondary metabolism and may regulate plant responses to abiotic or biotic stresses by controlling secondary metabolite production. For example, grape (*Vitis vinifera*) VvMYB14 and VvMYB15 coordinately regulate stilbene biosynthesis by binding to the promoter region of *STS* (stilbene synthase) genes; phenylalanine can form flavonoid compounds under the action of MYBF1, MYBPA1, and MYBPA2, or form stilbenes under the regulation of MYB15 [27]. White clover (*Trifolium repens*) TaMYB14 transcription factor can activate proanthocyanidin synthesis, thereby increasing anthocyanin accumulation in plants [28]. Soybean GmMYB29 is also a homolog of AtMYB15; overexpression or silencing of *GmMYB29* in soybean hairy roots can increase or decrease isoflavone content, and GmMYB29 regulates the expression of certain flavonoid pathway enzyme genes such as *PAL*, *4CL*, *CHS*, and *IFS* [29]. The *CiMYB15* gene from *C. intermedia* in this study also participates in regulating flavonoid metabolism. Overexpression of *CiMYB15* in *Arabidopsis* increased total flavonoid accumulation in transgenic plants, and the expression of *AtCHS* was upregulated in three overexpression lines. Unlike GmMYB29, CiMYB15 did not significantly affect the expression of other flavonoid biosynthesis enzyme genes (*CHI*, *F3H*, *FLS*, and *DFR*) in *Arabidopsis*. This suggests that the CiMYB15 transcription factor may specifically recognize and bind to the *CHS* gene promoter. The functional differences among homologous genes from different species likely arise from the highly diverse amino acid sequences in the non-conserved C-terminal regions of MYB proteins. Whether *CiMYB15* can regulate isoflavonoid metabolism like GmMYB29 may require further verification by transforming the gene into legume plants.

References

[1] Dubos C, Stracke R, Grotewold E, et al. MYB transcription factors in *Arabidopsis*. *Trends Plant Sci*, 2010, 15(10): 573-581.
[2] Feller A, Machemer K, Braun E L, et al. Evolutionary and comparative analysis of MYB and bHLH plant transcription factors. *The Plant Journal*, 2011,

66(1): 94-116.
[3] Ogata K, Hojo H, Aimoto S, et al. Solution structure of a DNA-binding unit of Myb: a helix-turn-helix-related motif with conserved tryptophans forming a hydrophobic core. *Proceedings of the National Academy of Sciences*, 1992, 89(14): 6428-6432.
[4] Ogata K, Morikawa S, Nakamura H, et al. Solution structure of a specific DNA complex of the Myb DNA-binding domain with cooperative recognition helices. *Cell*, 1994, 79(4): 639-648.
[5] Klempnauer K H, Gonda T J, Bishop J M. Nucleotide sequence of the retroviral leukemia gene v-myb and its cellular progenitor c-myb: the architecture of a transduced oncogene. *Cell*, 1982, 31(2 Pt 1): 453-463.
[6] Stracke R, Werber M, Weisshaar B. The R2R3-MYB gene family in *Arabidopsis thaliana*. *Current Opinion in Plant Biology*, 2001, 4(5): 447-456.
[7] Du H, Yang S S, Liang Z, et al. Genome-wide analysis of the MYB transcription factor superfamily in soybean. *BMC Plant Biology*, 2012, 12: 106-117.
[8] Yanhui C, Xiaoyuan Y, Kun H, et al. The MYB transcription factor superfamily of *Arabidopsis*: expression analysis and phylogenetic comparison with the rice MYB family. *Plant Molecular Biology*, 2006, 60(1): 107-124.
[9] Wilkins O, Nahal H, Foong J, et al. Expansion and diversification of the *Populus* R2R3-MYB family of transcription factors. *Plant Physiology*, 2009, 149(2): 981-993.
[10] Yang F Y, Liu K, Cui S, et al. Metabolism of flavonoids and the antioxidant capacity of transgenic *Arabidopsis thaliana* expressing *CiCHS* gene. *Acta Botanica Boreali-Occidentalia Sinica*, 2018, 38(3): 393-400.
[11] España L, Heredia-Guerrero J A, Reina-Pinto J J, et al. Transient silencing of CHALCONE SYNTHASE during fruit ripening modifies tomato epidermal cells and cuticle properties. *Plant Physiology*, 2014, 166(3): 1371-1386.
[12] Yu H N, Wang L, Sun B, et al. Functional characterization of a chalcone synthase from the liverwort *Plagiochasma appendiculatum*. *Plant Cell Reports*, 2015, 34(2): 233-245.
[13] Baudry A, Heim M A, Dubreucq B, et al. TT2, TT8, and TTG1 synergistically specify the expression of BANYULS and proanthocyanidin biosynthesis in *Arabidopsis thaliana*. *The Plant Journal*, 2004, 39(3): 366-380.
[14] Debeaujon I, Nesi N, Perez P, et al. Proanthocyanidin-accumulating cells in *Arabidopsis* testa: regulation of differentiation and role in seed development. *The Plant Cell*, 2003, 15(11): 2514-2531.
[15] Zhou J, Lee C, Zhong R, et al. MYB58 and MYB63 are transcriptional activators of the lignin biosynthetic pathway during secondary cell wall formation in *Arabidopsis*. *The Plant Cell*, 2009, 21(1): 248-266.
[16] Legay S, Sivadon P, Blervacq A S, et al. EgMYB1, an R2R3 MYB transcription factor from eucalyptus negatively regulates secondary cell wall formation in *Arabidopsis* and poplar. *New Phytologist*, 2010, 188(3): 774-786.
[17] Lacombe E, Van Doorsselaere J, Boerjan W, et al. Characterization of *cis*-elements required for vascular expression of the cinnamoyl CoA reductase gene and for protein-DNA complex formation. *The Plant Journal*, 2000, 23(5): 663-676.

[18] Stracke R, Ishihara H, Huep G, et al. Differential regulation of closely related R2R3-MYB transcription factors controls flavonol accumulation in different parts of the *Arabidopsis thaliana* seedling. *The Plant Journal*, 2007, 50(4): 660-677.
[19] Schwinn K E, Ngo H, Kenel F, et al. The onion (*Allium cepa* L.) R2R3-MYB gene MYB1 regulates anthocyanin biosynthesis. *Frontiers in Plant Science*, 2016, 7: 1865.
[20] Wang N, Xu H, Jiang S, et al. MYB12 and MYB22 play essential roles in proanthocyanidin and flavonol synthesis in red-fleshed apple (*Malus sieversii* f. *niedzwetzkyana*). *The Plant Journal*, 2017, 90(2): 276-292.
[21] Yang Q, Yin J, Li G, et al. Reference gene selection for qRT-PCR in *Caragana korshinskii* Kom. under different stress conditions. *Molecular Biology Reports*, 2014, 41(4): 2325-2334.
[22] Yu X M, Yue W R, Zhang Y N, et al. Heterologous expression of *CkLEA1* gene enhanced tolerance to abiotic stress in *Arabidopsis*. *China Biotechnology*, 2016, 36(10): 28-34.
[23] Attitalla I H. Modified CTAB method for high quality genomic DNA extraction from medicinal plants. *Pakistan Journal of Biological Sciences*, 2011, 14(21): 998-999.
[24] Chukwumah Y, Walker L T, Verghese M. Peanut skin color: a biomarker for total polyphenolic content and antioxidative capacities of peanut cultivars. *International Journal of Molecular Sciences*, 2009, 10(11): 4941-4952.
[25] Ding Z, Li S, An X, et al. Transgenic expression of MYB15 confers enhanced sensitivity to abscisic acid and improved drought tolerance in *Arabidopsis thaliana*. *Journal of Genetics and Genomics*, 2009, 36(1): 17-29.
[26] Kim S H, Kim H S, Bahk S, et al. Phosphorylation of the transcriptional repressor MYB15 by mitogen-activated protein kinase 6 is required for freezing tolerance in *Arabidopsis*. *Nucleic Acids Research*, 2017, 45(11): 6613-6627.
[27] Holl J, Vannozzi A, Czemmel S, et al. The R2R3-MYB transcription factors MYB14 and MYB15 regulate stilbene biosynthesis in *Vitis vinifera*. *The Plant Cell*, 2013, 25(10): 4135-4149.
[28] Hancock K R, Collette V, Fraser K, et al. Expression of the R2R3-MYB transcription factor TaMYB14 from *Trifolium arvense* activates proanthocyanidin biosynthesis in the legumes *Trifolium repens* and *Medicago sativa*. *Plant Physiology*, 2012, 159(3): 1204-1220.
[29] Chu S, Wang J, Zhu Y, et al. An R2R3-type MYB transcription factor, GmMYB29, regulates isoflavone biosynthesis in soybean. *PLoS Genetics*, 2017, 13(5): e1006770.

Note: Figure translations are in progress. See original paper for figures.