Alternatively Spliced *BnaPAP2.A7* Isoforms Play Opposing Roles in Anthocyanin Biosynthesis of *Brassica napus* L.

Daozong Chen[1], Yi Liu[1], Shuai Yin[1], Jie Qiu[1], Qingdong Jin[1], Graham J. King[2], Jing Wang[1], Xianhong Ge[1*] and Zaiyun Li[1]

[1] National Key Laboratory of Crop Genetic Improvement, National Center of Oil Crop Improvement (Wuhan), College of Plant Science and Technology, Huazhong Agricultural University, Wuhan, China, [2] Southern Cross Plant Science, Southern Cross University, Lismore, NSW, Australia

Brassica napus L. (rapeseed, oilseed rape, and canola) and varieties of its two diploid parents, *B. oleracea* and *B. rapa*, display a large amount of variation in anthocyanin pigmentation of the leaf, stem, and fruit. Here, we demonstrate that *BnaPAP2.A7*, an ortholog of the *B. oleracea* anthocyanin activator *BoMYB2* that confers purple traits, positively regulates anthocyanin biosynthesis in leaves of *B. napus*. Sequencing of *BnaPAP2.A7* and transgenic analysis suggests that activation of this gene in purple rapeseed may result from a single nucleotide and/or 2bp insertion in its promoter region. *BnaPAP2.A7* gives rise to three splice variants, designated *BnaPAP2.A7-744*, *BnaPAP2.A7-910*, and *BnaPAP2.A7-395* according to the length of the transcripts. While *BnaPAP2.A7-744* encodes a full-length R2R3-MYB, both *BnaPAP2.A7-910* and *BnaPAP2.A7-395* encode truncated proteins that lack both a partial R3 repeat and the complete C terminal domain, and so *in vitro* are unable to interact with the *Arabidopsis* bHLH protein AtTT8. Although expression of either *BnaPAP2.A7-910* or *BnaPAP2.A7-395* in green rapeseed does not result in purple leaves, both genes do modify genome-wide gene expression, with a strong repression of anthocyanin-related genes. We have demonstrated that *BnaPAP.A7* regulates anthocyanin accumulation in leaves of *B. napus* and propose a potential mechanism for modulation of anthocyanin biosynthesis by alternative splicing.

Keywords: *Brassica napus* L., anthocyanin, R2R3-MYB, *BnaPAP2.A7*, alternative splicing

INTRODUCTION

Anthocyanins are a group of plant secondary metabolite flavonoid pigments. They play the role of "nature's Swiss army knife" in conferring tolerance to various stressors as well as in defense against herbivores and pathogens (Gould, 2004). Anthocyanins are biosynthesized from phenylalanine *via* a series of enzymes encoded by structural genes including phenylalanine-ammonia lyase (*PAL*), cinnamate 4-hydrodylase (*C4H*), and 4-coumaryl, CoA ligase (*4CL*), early anthocyanin biosynthesis

(EBGs) genes, chalcone synthase (*CHS*), chalcone isomerase (*CHI*), and flavanone 3-hydroxylase (*F3H*), late anthocyanin genes (LBGs), dihydroflavonol 4-reductase (*DFR*), anthocyanidin synthase (*ANS*), and anthocyanidin 3-O-glucosyltransferase (*UFGT*) (Saito et al., 2013). The expression of these genes is usually regulated at the transcriptional level and primarily by individual MYB transcription factors (TFs). Of these, the MBW transcription factor complex, composed of R2R3-MYB, basic-helix-loop-helix (bHLH), and WD40 proteins, has been shown to control expression of the LBGs (Broun, 2005; Ramsay and Glover, 2005; Gonzalez et al., 2008; Petroni and Tonelli, 2011; Albert et al., 2014; Xu et al., 2015).

Within the MBW complex R2R3-MYBs are the major determinants involved in anthocyanin biosynthesis (Lai et al., 2014; Liu et al., 2015; Jin et al., 2016), either as activators or repressors of structural genes (Pérez-Díaz et al., 2016). They represent the largest subgroup of the plant MYB TF family, and they have two MYB repeats in the N terminal DNA-binding domain that are most similar to R2 and R3 from c-MYB (Feller et al., 2011) as well as an activation or repression domain usually located at the C terminus (Zimmermann et al., 2004; Dubos et al., 2010; Xu et al., 2014). Based on the conservation of the DNA binding domain and amino acid motifs in the C terminal domains, R2R3-MYB proteins have been classified into several subgroups that are involved in plant-specific processes (Dubos et al., 2010). *ZmC1* was the first R2R3-MYB transcription activator identified, reported to regulate the synthesis of anthocyanins in the aleurone of maize (*Zea mays*) kernels (Cone et al., 1986; Paz-Ares et al., 1987). *ZmPl* was later reported to regulate biosynthesis of anthocyanins that give rise to the purple leaf phenotype in maize (Cone et al., 1993). R2R3-MYB transcription activators have subsequently been reported in various plant taxa, such as *AtMYB75* (*PAP1*), *AtMYB90* (*PAP2*), *AtMYB113*, and *AtMYB114* in *Arabidopsis* (Borevitz et al., 2000; Stracke et al., 2001; Zimmermann et al., 2004; Stracke et al., 2007; Gonzalez et al., 2008), *ROSEA1* and *ROSEA2* in *Antirrhinum* (Schwinn et al., 2006; Shang et al., 2011), *VvMYBA1* and *VvMYBA2* in *Vitis* grapevine (Walker et al., 2007), and *Ruby* in *Citrus* (Butelli et al., 2012; Huang et al., 2018).

The first flavonoid-related R2R3-MYB repressor, AmMYB308, was characterized from *Antirrhinum majus* (Tamagnone et al., 1998). R2R3-MYB repressors belong to the same R2R3-MYB family subgroup and are characterized by two conserved motifs in the C-terminal region, C1 (lsrGIDPxT/NHR) and C2/EAR (pdLNLD/EL) (Chen et al., 2019). They have also been identified in many plant taxa, such as *AtMYB4* in *Arabidopsis* (Hemm et al., 2001), PhMYB27 in petunia (Albert et al., 2014), and VvMYBC2-L1/3 and VvMYB4-like in *Vitis* (Cavallini et al., 2015; Pérez-Díaz et al., 2016). Many of these repressors have a function similar either to *Arabidopsis* AtMYB4, which represses the transcription of lignin and general phenylpropanoid genes, or to strawberry FaMYB1 (Aharoni et al., 2001), which negatively regulates transcription of anthocyanin and/or proanthocyanidins (PA) pathway genes (Jin et al., 2000; Cavallini et al., 2015; Jun et al., 2015; Liu et al., 2019). However, the mechanism by which these repressors regulate anthocyanin biosynthesis remains mostly unclear.

Alternative splicing (AS) is a common phenomenon in plants and leads to the production of more than one mRNA transcript from the same gene. AS not only contributes to proteome diversity but can also generate truncated proteins that are potentially regulatory or detrimental to cells (Syed et al., 2012). AS was found to occur in >60% of intron-containing genes in *Arabidopsis* (Marquez et al., 2012), and a similar percentage was estimated in other plant species (Syed et al., 2012). However, to date AS has only been detected in a limited number of R2R3-type MYB genes (Li et al., 2006), none of which have been identified as being involved in the anthocyanin biosynthesis pathway.

Brassica napus L. (rapeseed, oilseed rape, and canola) (AACC, 2n=38) is an important and high yielding global oil crop and, along with *Arabidopsis*, is a member of the Brassicaceae. It is thought to have originated by allopolyploidisation following spontaneous interspecific hybridization between the diploids, *B. rapa* (AA, 2n = 20) and *B. oleracea* (CC, 2n = 18), about 7,500 years ago (Chalhoub et al., 2014). Considerable variation in anthocyanin pigmentation is found in leaves, stems, and fruit of rapeseed as well as in various crop types of *B. oleracea* and *B. rapa* (Izzah et al., 2013). Several R2R3-MYB genes have been reported to regulate the expression of anthocyanin pathway structural genes in *B. rapa* (He et al., 2016) and *B. oleracea* (Yuan et al., 2009; Zhang et al., 2015). The distinct purple color formation found in some cultivars of cauliflower, cabbage, and kale/kohlrabi have been reported to be caused by the insertion of different transposons, point mutations, and/or a 1-bp insertion in the promoter region of the *BoMYB2* gene (Chiu et al., 2010; Chiu and Li, 2012; Yan et al., 2019). However, due to the triplicated genome origins of the *Brassica* lineage, multiple paralogs of these orthologs are usually present within the diploids *B. rapa* (A genome) and *B. oleracea* (C genome), leading to many more copies within the allotetraploid *B. napus* (AC genome). This complexity leads to uncertainty in establishing which functional alleles contribute to a purple phenotype in *B. napus*, with multiple paralog candidates on each of the A or C genomes.

In the present study, we identified a rapeseed R2R3-MYB activator, *BnaPAP2.A7* (BnaA07g25800D). Functional analysis indicated that *BnaPAP2.A7* acts as a positive regulator of anthocyanin accumulation in leaves. In addition, *BnaPAP.A7* produced three AS isoforms that are associated with variation in transcript abundance, protein sequence, and function. The wild-type spliced isoform encodes a protein with a complete R2, R3 repeat and C terminal domain, while the two AS isoforms produced truncated proteins lacking a partial R3 repeat and a complete C terminal domain. Interestingly, transgenic analysis indicated that the two spliced isoforms could significantly downregulate the expression of general phenylpropanoid genes as well as the expression of anthocyanin biosynthesis genes. We discuss the potential mechanism involving AS that is associated with the balance of positive and negative regulation of anthocyanin biosynthesis by the same R2R3-MYB protein.

MATERIALS AND METHODS

Plant Materials and Phenotypic Anatomy

A *B. napus* alien introgression line was generated from the sequential crosses of (*B. rapa* ssp. *chinensis* L. × *Orychophragmus*

violaceus) × *B. napus* by successive phenotypic selection and cytological observation (Li and Heneen, 1999; Xu et al., 2019). After selfing for more than 10 generations, this line possessed the same chromosome complement (2n=38) as *B. napus*, a normal meiotic behavior, and a good seed-set. It was characterized by purple pigmentation of all organs apart from yellow petals and named as PR (purple rapeseed). In addition, two homozygous accessions of *B. napus*, Zhongshuang 11 (ZS11) and Jia9709, which had a wild-type green phenotype were also used in this study for comparative transcriptome analysis and as a recipient for transformation, respectively. All plants were grown in the experimental field and greenhouse in Huazhong Agricultural University, Wuhan. To examine cells with purple pigmentation, young leaves were transversely sectioned by a free-hand section and examined with a Zeiss Axioscope photomicroscope equipped with a MRC digital camera. Transgenic *B. napus* plants were grown in an isolated experimental station. *Arabidopsis thaliana* and *Nicotiana benthamiana* plants were grown in a plant growth chamber at 20–22°C and 70% humidity under a photoperiod of 16-h light/8-h dark.

Gene Cloning, Plasmid Construction, and Genetic Transformation

According to the reference from Darmor-*bzh*, a 3.66-kb genomic DNA fragment, including the 1.99-kb upstream regulatory sequence of the *BnaA07g25800D*, was amplified from PR using primer set PAP.A7-PF and PAP.A7-gR. PCR products were cloned into pMD^{TM}18-T (Takara, Dalian, China) and transformed into *Escherichia coli* (TransGen, Beijing, China) and sequenced. Because the translated amino acid level sequences of *BnaA07g25800D* have the highest similarity with *AtPAP2*, it was named *BnaPAP2.A7*. To obtain the coding sequence (CDS), the primer set PAP2.A7-CDS-PF and PAP2.A7-CDS-PR were used to amplify CDS from the young leaf cDNA of PR. Three fragments of different lengths were obtained and cloned into pMD^{TM}18-T, transformed into *E. coli* and sequenced. The 3.66-kb full-length gene, including 1.99-kb of upstream regulatory sequence and the 1.67-kb coding region for *BnaPAP2.A7*, was amplified from the plasmid using primers PAP2.A7-KpnI-PF and PAP2.A7-BamHI-gR, and cloned into the KpnI-BamHI sites of binary vector pCAMBIA2301-nos. Subsequently, we amplified the 1.99-kb upstream regulatory sequence from the plasmid using primer set PAP2.A7-KpnI-PF and PAP2.A7-BamHI-PR and cloned this into the *KpnI-BamHI* cloning sites of the binary vector pCAMBIA2301-nos. Three transcripts of *BnaPAP2.A7* were also amplified again from the plasmid using primer set PAP2.A7- BamHI-PF and PAP2.A7-BstEII-PR and cloned into the BamHI - BstEII cloning sites of the vector pCAMBIA2301 behind the 1.99-kb upstream regulatory promoter, respectively. The four plasmids above were finally introduced into the *Agrobacterium tumefaciens* strain GV3101 (pMP90) and transformed into Jia9709 according to Cardoza and Stewart (2003).

Subcellular Localization

Three transcript fragments were amplified and cloned into the *Pml*I-A*si*SI sites of pDOE20 (Gookin and Assmann, 2014) to generate mVenus-BnaPAP2. A7 (YFP- BnaPAP2. A7) fusion protein. The full-length cDNA fragments of *AtPAP2* from *Arabidopsis thaliana* were amplified and cloned into the *Bam*HI-*Xba*I sites of pDOE20 to generate the mTurquoise2-AtPAP2 (CFP-AtPAP2) fusion protein. All the resulting plasmids were introduced into *A. tumefaciens* EHA105, which were then infiltrated into 4-week-old *N. benthamiana* leaves for transient expression. In brief, the *A. tumefaciens* strains, containing the pDOE20 recombinant plasmid that expresses CFP-AtPAP2 and YFP-BnaPAP2. A7, also expresses the silencing suppressor p19 of *Tomato bushy stunt* virus. These were harvested and re-suspended to a final density of 0.2 at 600 nm (OD$_{600}$) in an infiltration buffer (150 mM acetosyringone, 10 mM MgCl$_2$ and 10 mM MES). The *Agrobacterium* suspension was injected into expanded leaves of 4-week-old tobacco plants. The tobacco leaves were imaged under a Zeiss LSM 700 confocal microscope after three days of infiltration. Excitation of YFP was performed using an argon laser line at 488 nm and a cyan fluorescent protein (CFP) with 405 nm (Qin et al., 2017).

Yeast Two-Hybrid Assays

Yeast two-hybrid assays were conducted using the Frozen-EZ Yeast Transformation II Kit (ZYMO RESEARCH, USA, http://www.zymoresearch.com). The PCR products of *AtTT8* from *A. thaliana* were cloned into the pGBKT7 (BD) vectors, and three transcripts from PR leaves cDNA were independently cloned into the pGADT7 (AD) vectors. The PCR products of *AtPAP2* from *A. thaliana* were cloned into the pGADT7 (AD) vectors as a positive control. The transformation of AH109 cells was performed according to the protocol from Frozen-EZ Yeast Transformation II Kit. The transformed yeast cells were spread on the medium SD/-Leu/-Trp. For the yeast two-hybrid assay, some baits contained *AtTT8* plasmids while others did not. The preys used were with transcript or were empty AD vector plasmids introduced into yeast AH109 cells. The colony that was grown on the medium SD/-Leu/-Trp were chosen by selective medium SD/-Leu/-Trp/-His/-Ade to determine the interactions (Qin et al., 2017).

RNA Preparation, Reverse-Transcription PCR, and qRT-PCR Analysis

The seventh leaf of PR, ZS11, and positive and negative transgenic T$_2$ plants were successively collected with three biological repeats and immediately stored in liquid nitrogen for RNA extraction. Total RNA was extracted from various plant tissues using Eastep Super Total RNA Extract Kit (Promega, Shanghai, China) supplemented with RNase-free DNaseI to remove contaminated DNA according to the manufacturer's instructions. First-strand cDNA was synthesized using a RevertAid First Strand cDNA Synthesis Kit (Thermo, USA, https://www.thermofisher.com/cn/zh/home.html). The cDNA was amplified on a CFX96TM Real-time PCR Detection System (Bio-Rad, Germany, http://www.bio-rad.com/). Eight genes PAL1 (BnaA04g21230D), C4H (BnaA04g21230D), CHS (BnaA10g19670D), F3'H (BnaA10g23330D), MYB4 (BnaC03g60080D), ANS (BnaA01g12530D), DFR

(BnaC09g17150D), and *MYBL2* (BnaA02g15320D) were selected for qRT-PCR confirmation of the expression level revealed by RNA-seq analysis. The specific quantitative primers for different genes and their transcripts were designed using Primer 5.0. qRT-PCR assays with three biological replicates and three technical repetitions were performed using a Luna Universal qPCR Master Mix (Biolabs, USA) on a Bio-Rad CFX96 Real-Time Detection System (Bio-Rad, Germany, http://www. bio-rad.com/). The *Bnaactin3* gene was used as an internal control for data normalization, and quantitative variation in the different replicates was calculated using the delta-delta threshold cycle relative quantification method as described previously (Fu et al., 2018).

RNA-Seq and Gene Ontology (GO) Enrichment Analysis

We first re-analyzed the leaf transcriptomic data of PR and ZS11 (Mushtaq et al., 2016) young leaf RNA-seq data (each with two duplicates). In order to investigate the potential gene expression changes in transgenic lines, we also performed RNA-seq analysis. Total RNA of transgenic T_2 lines (three biological duplicates each from independent transformant) containing full-length *BnaPAP2. A7*, *BnaPAP2. A7-395*, and *BnaPAP2. A7-910* as well as negative transgenic T_2 plants (three biological duplicates) were used for the sequencing library construction. The RNA-seq libraries were constructed according to the user manual (Illumina, http:// www.illumina.com/) and sequenced to produce 150-bp paired-end reads. Low-quality reads were removed from the raw reads using the Cutadapt (Martin, 2011) and Trimmomatic (Bolger et al., 2014) software. Clean reads were mapped to the *B. napus* Darmor-*bzh* genome sequence (http://www.genoscope.cns.fr/ brassicanapus/) (Chalhoub et al., 2014) using TopHat2 (Kim et al., 2013). The read counts of each gene were calculated using the HTseq-count function in the HTseq software package (Anders et al., 2015). The R program DEseq2 (http://www.bioconductor. org/packages/release/bioc/html/DESeq2.html) was used to identify differentially expressed genes (DEG)s between purple and green plants based on the criteria: padj < 0.01 & log2FoldChange > 2. Gene expression levels were calculated in terms of Fragments Per Kilobase of exon model per Million mapped reads (FPKM) using Cufflinks (Trapnell et al., 2012). The GO enrichment analysis was performed for DEGs using TBtools software (Chen et al., 2018). The top 20 enriched gene ontology (GO) terms for the comparative transcriptome DEGs of full-length gene, *BnaPAP2.A7-395*, and *BnaPAP2.A7-910*-positive transgenic plants compared with Jia9709 plants were used to produce the GO enrich diagram.

Sample Preparation, Anthocyanin Content Determination, and Metabolic Analysis

Young leaves from three biological replicates of PR, Jia9709, and the positive transgenic lines were collected. For the full-length *BnaPAP2.A7*, *BnaPAP2.A7-395*, and *BnaPAP2.A7-910*-positive lines, three T_2 lines were chosen, and three plants of each were collected as nine biological replicates. All samples were immediately stored in liquid nitrogen for later anthocyanin extraction. A total of 0.5 g of frozen powder was placed into a 5-mL screw-cap tube with 5mL of methanol/water/acetic acid (85:15:0.5; v, MeOH/H2O/AcAc). The sample was then put on

ice in the dark with sonication for 30 min and shaken at 4°C in the dark overnight. The supernatant was collected after centrifuging at 12,000 x g for 10 min and was further diluted and filtered twice using a 0.22 μm PTFE filter (Cameo 25F, Micron Separations Inc., Westboro, MA) before injection for metabolomics analysis (Fu et al., 2018).

For the anthocyanin content analysis, firstly, a dilution series of standard cyanidin ($C_{15}H_{11}ClO_6$; Solarbio, Beijing, China) was measured at 520 nm using a spectrophotometer to obtain a standard curve of anthocyanin content and absorbance. Then, the absorbance of the treated sample was obtained using a spectrophotometer, and the anthocyanin content was calculated from the obtained standard curve.

Metabolomics analysis was performed on a Waters Acquity UPLC system connected to a Synapt G2-XS QToF mass spectrometer, which was equipped with an electrospray ionization source (Waters, Milford, MA) according to methods previously described (Fu et al., 2018). However, we only collected anthocyanin related metabolites, which have an absorption peak at 530 nm. Data processing was performed using the MarkerLynx application manager for Masslynx4.1 software. Principal component analysis (PCA) and cluster analysis of anthocyanin content based on anthocyanin type and content from different plants was performed with MetaboAnalyst (https://www.metaboanalyst.ca/MetaboAnalyst).

Phylogenetic Analysis

Protein sequences of PAP1, PAP2, MYB113, and MYB114 from *Arabidopsis thaliana* and homologs in *B. rapa*, *B. oleracea*, and *B. oleracea* were aligned using MEGA7 (Kumar et al., 2016), as were evolutionary analyses. The evolutionary history was inferred by using the Maximum Likelihood method based on the JTT matrix-based model (Jones et al., 1992). The tree with the highest log likelihood (-874.7994) is shown. Initial tree(s) for the heuristic search were obtained automatically by applying Neighbor-Join and BioNJ algorithms to a matrix of pairwise distances estimated using a JTT model, and then selecting the topology with superior log likelihood value. The tree is drawn to scale, with branch lengths proportional to the number of substitutions per site. The analysis involved 19 amino acid sequences. All positions containing gaps and missing data were eliminated. There were a total of 130 positions in the final dataset.

RESULTS

Characterization of Sequence Variation and Expression of *BnPAP2.A7* in Purple and Green Rapeseed

At the seedling stage, purple rapeseed (PR) displays distinctive purple leaves and petioles (**Figure 1A**) compared to the green rapeseed ZS11 (**Figure 1B**). In addition, the purple phenotype of PR is not visible when planted in a glass greenhouse (**Figure 1C**). This result suggests that ultraviolet radiation, which is largely blocked by glass, may be required for purple color formation. As shown in leaf sections, the purple pigments mainly

accumulated in the epidermal cells (**Figure 1D**). Comparative leaf transcriptome analysis revealed significant upregulation of homologous late-biosynthesis genes (LBGs) involved in the anthocyanin biosynthesis pathway, including *DFR*, *ANS*, and UGFT, in PR compared with ZS11 (Mushtaq et al., 2016). These results suggested that allelic variation of one of the MYB-bHLH-WD40 (MBW) components may result in the activation of LBGs and lead to the purple color formation of PR. We therefore identified all homologous genes, classified as *Arabidopsis*, as MBW components within the reference *B. napus* genome (Darmor-*bzh*) (**Table S1**) and discovered that two of the eight homologous genes of *AtPAP2*, *BnaCnng28030D* and *BnaA07g25800D*, were differentially expressed between PR and ZS11. However, only *BnaA07g25800D* was highly expressed in PR, which is consistent with its purple-color phenotype (**Figure 2A**). Quantitative RT-PCR (qRT-PCR) analysis further confirmed that *BnaA07g25800D* was expressed at a significantly higher level in the leaves of PR than in those of ZS11 (**Figure 2B**). We also found that *BnaA07g25800D* was mostly absent in the leaves of greenhouse grown plants but was highly expressed in the leaves of plants grown outdoors (**Figure 2C**). BnaA07g25800D was thus considered the key transcription factor involved in the regulation of anthocyanin biosynthesis in leaves of PR.

To determine the full-length sequences of the corresponding alleles from PR, ZS11, and J9709, we isolated the gene using overlapping primers that bound specifically to the genomic region of *BnaA07g25800D*. J9709 also has a green phenotype and is particularly susceptible to genetic transformation. The PCR products represented about 3.66 kb of genomic fragments, comprised of a 1.67 kb transcribed region and a 1.99 kb promoter region. Compared with ZS11, the PR allele included a 1bp insertion (T) at -1,339 and a 2bp insertion (TC) at -611bp within the promoter region (**Figure 3A**). In the transcribed region, there is a 2 bp (CA) insertion in ZS11 at +176 in intron 1, and a 1bp (T) insertion in intron 2 at +1,083 (**Figure 3A**). The sequence of the J9709 allele is identical to ZS11 in both the promoter and transcribed regions and is identical to the PR allele within the introns (**Figure 3A**). The translated amino acid sequences of the three alleles were identical. As the translated amino acid sequences of *BnaA07g25800D* were most similar to those of AtPAP2, it was named *BnaPAP2.A7*. Phylogenetic analysis revealed that *BnaPAP2.A7*, *BraA07g032100*, and *Bo6g100940* (*BoMYB2*) are classified within the same subgroup (**Figure 3B**). This suggests that these genes may have similar functions.

BnaPAP2.A7 Produces Three Transcript Isoforms *via* AS

Interestingly, when we isolated the full-length cDNA of *BnaPAP2.A7* from PR we found three alternative splicing (AS)
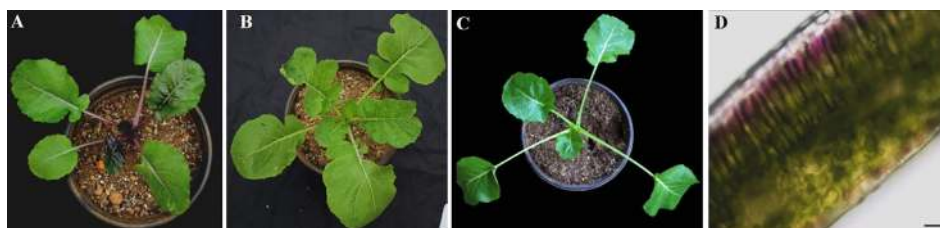

FIGURE 1 | Characterization of anthocyanin accumulation in leaves of purple rapeseed. **(A–C)** Different phenotype of young plant of PR **(A)** and ZS11 **(B)** planted outdoor and PR in the glass greenhouse **(C)**. **(D)** Hand section of a young leaf of PR. Scale bar: 10μm.
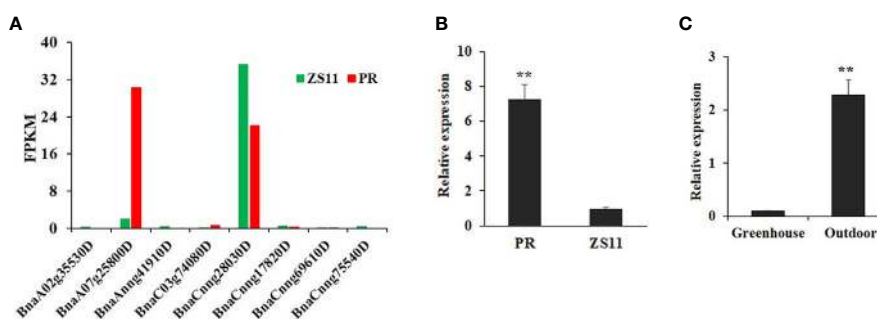

FIGURE 2 | Characterization of the expression of *BnPAP2.A7* in purple and green rapeseed. **(A)** The fragments per kilobase of exon model per million mapped reads (FPKM) of eight *AtPAP2* homologous genes in leaves of PR and ZS11. **(B)** qRT-PCR analysis of *BnaPAP2.A7* in leaves of PR and ZS11 planted outside. **(C)** qRT-PCR analysis of *BnaPAP2.A7* in leaves of PR planted outside and in glass greenhouse. **P < 0.01, one-way ANOVA test. Error bars represent the SD of three replicates.
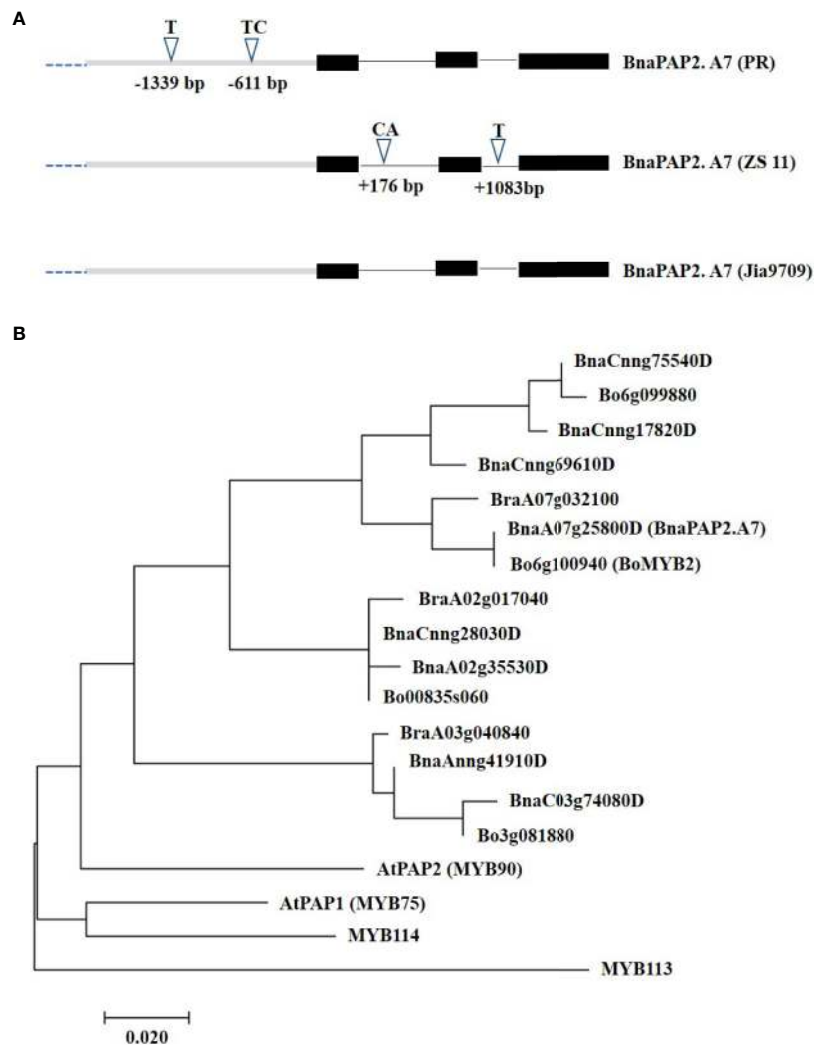
FIGURE 3 | Allelic variation and phylogenetic analysis of *BnaPAP2. A7*. **(A)** Allelic variation of *BnaPAP2. A7* including 1.99 kb promoter region and 1.67 kb coding region between PR and ZS11. **(B)** Phylogenetic tree derived from amino acid sequences of R2R3-MYBs in *B. napus*, *B. rapa*, and *B. oleracea* showed homology to *Arabidopsis AtPAP2*.

isoforms of 910 bp, 395 bp, and 744 bp (**Figure 4A**). To analyze their relative expression, RT-PCR products amplified by primers for coding regions were randomly cloned and detected by colony PCR and agarose gel electrophoresis. Based on 900 clones, the results indicated transcript abundance decreasing from isoform 744 (51.1%) > isoform 395 (40.1%) > isoform 910 (8.8%) (**Figure 4B**). qRT-PCR analysis using specific primer pairs for each isoform also revealed a similar expression trend (**Figure 4C**). When these isoforms were compared with the predicted coding DNA sequence (CDS) of BnaA07g25800D, the 744 isoform with three exons was determined to be the normal transcript. However, 910 contained these three exons and an additional 166 bp of the second intron, whilst 395 contained the first two exons and the first 349 bp of the third exon (**Figure 5A**). At the amino acid level, the 744 isoform-translated proteins have 244 amino

acids and the complete R2, R3, and C-terminal domains. However, due to AS, the 910 and 395 isoforms contain premature stop codons and produce proteins composed of 103 and 108 amino acids, respectively (**Figure 5A**). The latter two proteins thus contain a complete R2 repeat but lack both a section of R3 from the middle region of the second α-helix as well as the complete original C-terminal domain (**Figure 5B**).

Truncated Proteins Encoded by Two Spliced Isoforms Are Localized in the Nucleus but Have Lost the Ability to Interact With the bHLH Protein

To explore the subcellular localization of the proteins encoded by the three transcripts, a pDOE20 plasmid co-expressing BnaPAP2.A7-yellow fluorescent protein (YFP) and AtPAP2-cyan fluorescent
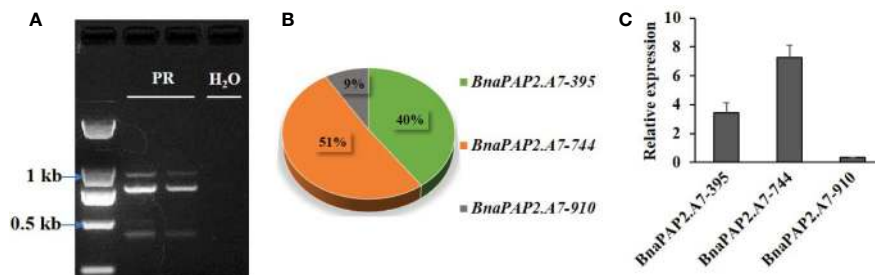
FIGURE 4 | Characterization the expression of three alternative spliced transcripts produced by *BnaPAP2. A7*. **(A)** Detection of three isoforms by agarose gel electrophoresis. **(B)** Expression abundance analysis of three isoforms by colony PCR and agarose gel electrophoresis. **(C)** qRT-PCR analysis of three isoforms.
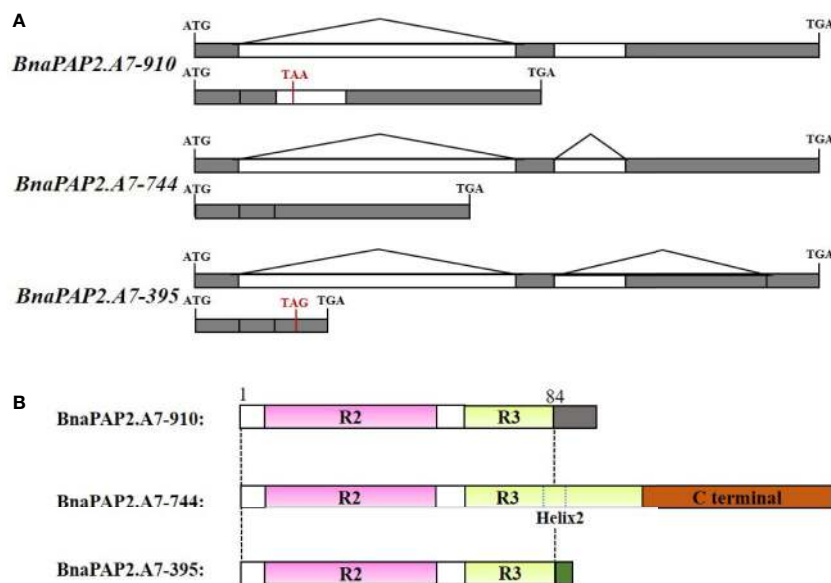

FIGURE 5 | Production, gene structure, and proteins sequence of three isoforms of *BnaPAP2. A7.PR*. **(A)** Production and gene structure of the three isoforms. **(B)** Schematic diagram showing the protein sequence encoded by three isoforms. Gray and green boxes represent the new short C-terminal sequence that are completely different from those of original C-terminal.

protein (CFP) fusion proteins were introduced into tobacco leaves for transient expression analysis. The results demonstrated that the proteins produced by the three isoforms of *BnaPAP2.A7* co-localized with AtPAP2 in the nucleus (**Figure 6A**). We wished to investigate whether the loss of the R3 and C-terminal domain affects the affinity between BnaPAP2.A7 and bHLH, which is important in subsequent MBW complex formation. We therefore cloned *AtTT8* and performed a yeast two-hybrid assay of the interaction between AtTT8 and proteins encoded by the three isoforms, with AtPAP2 used as a positive control. AtTT8 strongly interacted with the 744-encoded protein in yeast cells, but no binding was detected between AtTT8 and proteins produced by 910 or 395 (**Figure 6B**). We therefore speculate that only the 744 isoform-encoded protein participates in the formation of MBW complexes and regulates anthocyanin biosynthesis.

Transgenic Analysis of Full-Length *BnaPAP2.A7* and Three AS Transcripts in Green Rapeseed

To test our speculation further, we constructed pCAMBIA2301-based expression vectors containing full-length genomic fragments of *BnaPAP2.A7.PR* that included both the 1.99 kb promoter and 1.67 kb transcribed regions as well as the three isoforms driven by the same 1.99 kb promoter, and we transformed these into the green-leaved Jia9709. This gave rise to 15 positively transformed plants harboring the full-length sequence of *BnaPAP2.A7.PR*, 56 for the 395 isoform, and 33 for the 910 isoform. However, no transformed plant harboring the 744 isoform was obtained, despite attempts using various infection conditions (**Table S2**) and different expression vectors. Of the 15 positive plants carrying the full-length *BnaPAP2.A7.PR* gene, six displayed
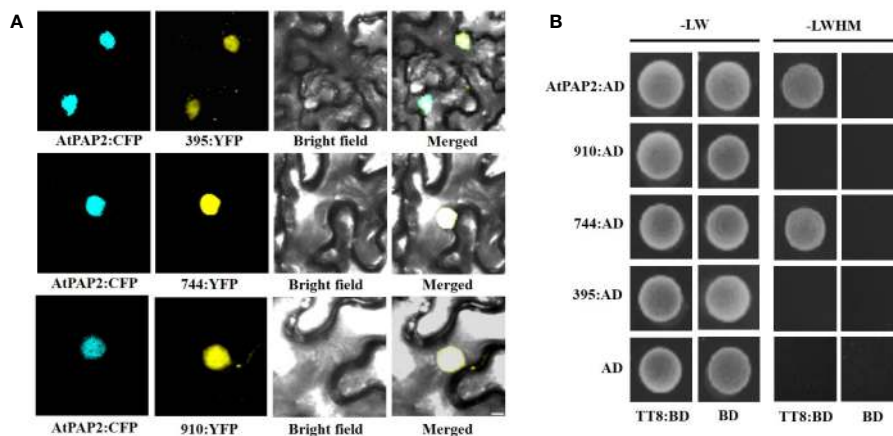
FIGURE 6 | Subcellular localization of proteins encoded by three isoforms and their ability to bind to bHLH protein revealed by Yeast two-hybrid analyses. **(A)** BiFC interaction assays of BnaPAP2.A7-395, BnaPAP2.A7-744, and BnaPAP2.A7-910 with AtPAP2. Partial YFP fusion constructs were expressed transiently in *N. benthamiana* leaves. AtPAP2-CFP was used as a known nuclear marker. Scale bar: 100 μm. **(B)** Yeast two-hybrid analyses of interactions between three transcripts and AtTT8. 395: BnaPAP2.A7-395; 744: BnaPAP2.A7-744; 910: BnaPAP2.A7-910. LW: Leucine and Tryptophan; LWHM: Leucine, Tryptophan, Histidine and Methionine.

purple leaves in the T_0 generation, three of which (WZ1, WZ2, and WZ5) were used for further study. A deeper purple color was observed in the subsequent T_1 (data not shown) and T_2 generations of all three lines (**Figure 7A**). However, no purple color was observed in the leaves or other organs of transgenic plants harboring the 910 or 395 isoforms in the T_1 (data not shown) and T_2 generations (**Figure 7A**). RT-PCR analysis confirmed that Jia9709 carrying full-length *BnaPAP2.A7.PR* also produced three isoforms, each having a significantly higher transcript level in the T_2 transgenic plants than those in control plants (**Figure 7B**). Meanwhile, the 910 and 395 isoforms also had a significantly higher transcript level in transgenic plants harboring the 910 and 395 isoforms, respectively (**Figures 7C, D**).

In the T_2 generation, the WZ1, WZ2, and WZ5 lines had an average anthocyanin content of 3.13~4.1 mg/g fresh weight—significantly higher than that presented by Jia9709 (0.61~1 mg/g). However, the leaf anthocyanin content of the 910 (0.66~0.85 mg/g) and 395 (66~0.89 mg/g) lines were similar to the control line (**Figure S1A**). Non-target metabolic analysis identified similar numbers of anthocyanin related metabolites in the control (921 metabolites) as well as in the full-length *BnaPAP2.A7.PR* (939), 910 isoform (912), and 395 isoform (905) transgenic lines (**Supplementary datasheet 1**). However, metabolite-based PCA analysis revealed that the three transgenic lines and controls each formed distinct separated groups of plants (**Figure S1B**). Cluster analysis shows that while plants of the 395 and 910 transgenic lines had similar metabolite profiles, each were significantly different from those of the *BnaPAP2.A7.PR* full-length transgenic lines and controls (**Figure S1C**). Taken together, these results indicated that the proteins encoded by 910 and 395 failed to promote anthocyanin synthesis. Therefore, it is reasonable to consider that the 744 isoform-encoded proteins may participate in the formation

of MBW complexes and regulate anthocyanin biosynthesis, although no plants transformed with this isoform were obtained.

The Potential Negative Regulation of Anthocyanin Biosynthesis by Two Spliced Isoforms

In order to analyze the potential function of the 395 and 910 isoforms, we performed RNA-seq analysis of full-length, 395 and, 910 transgenic plants as well as plants subjected to transformation but with no inserts, which served as controls. Compared to the control plants this identified 391 (374 up/17 down) differentially expressed genes (DEGs) in transgenic plants harboring the full-length gene, 3,820 (1,348 up/2,472 down) in the 395 isoform and 3,598 (1,238 up/2,360 down) in the 910 isoform (**Supplementary datasheet 2**). Gene ontology (GO) enrichment analysis revealed a substantial difference between the DEGs of the three types of transgenic plants (**Figures 8A–C**). DEGs of full-length *BnaPAP2. A7.PR* transgenic plants were enriched in several GO categories involved in flavonoid, anthocyanin, and proanthocyanidin biosynthetic and metabolic processes. However, DEGs in the 395 and the 910 transgenic plants were enriched in GO categories mainly involved in the phenol-containing compound catabolic process, the cinnamic acid biosynthetic and metabolic process, and other molecular function such as flavonoid binding (**Figures 8A–C**). In particular, all of the 16 DEGs involved in anthocyanin biosynthesis showed significantly higher transcript abundance in full-length *BnaPAP2. A7.PR* transgenic plants. These included F3'H, FLS, DFR, and UGFT as well as the anthocyanin transporter TT19 and transcription factors such as *PAP2* and *TT8* (**Supplementary datasheet 3**; **Figure S2A**). In contrast, most of the DEGs involved in anthocyanin biosynthesis in the 395 (45/57) and 910 (33/39) transgenic lines were significantly
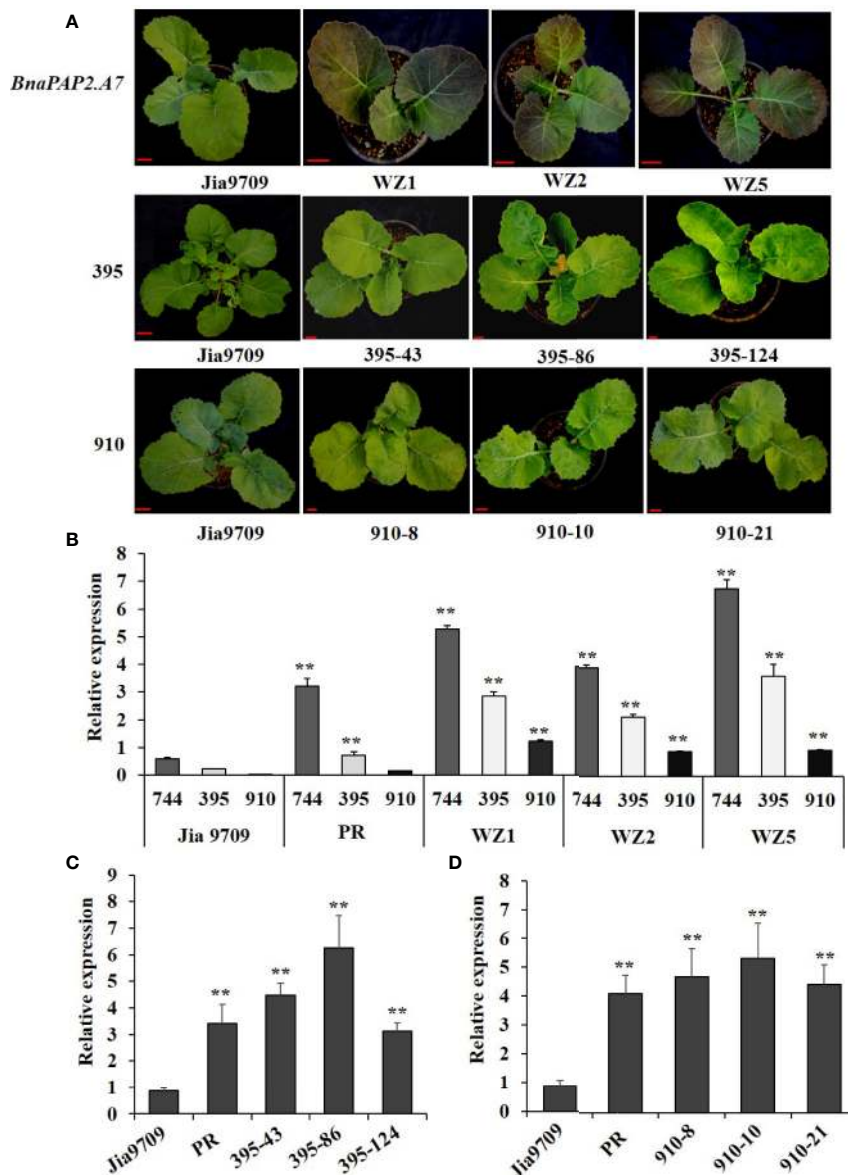
FIGURE 7 | Transgenic analysis of full-length *BnPAP2.A7* and three isoforms in green rapeseed. **(A)** Phenotypes of transgenic lines at the seedling stage containing full-length PR *BnaPAP2. A7*, 395, and 910 isoform driven by the same 1.99kb promoter of *BnaPAP2. A7*. Scale bar: 1 cm. **(B)** qRT-PCR analysis of three isoforms in transgenic plants with full-length *BnaPAP2. A7*. **(C, D)** qRT-PCR analysis of 395 **(C)** and 910 **(D)** isoforms in respective transgenic plants. **P < 0.01, one-way ANOVA test. Error bars represent the SD of three replicates.

downregulated, including *PAL*, *C4H*, and *4CL* and almost all EBGs (**Supplementary datasheet 3**, **Figures S2B, C**). Additionally, transcription factors such as MYB111 and MYB12 were significantly downregulated while MYBL2 was significantly upregulated (**Supplementary datasheet 3**). These results suggest that the two alternatively spliced transcripts disturbed the pattern of genome-wide gene transcription and repressed the expression of many anthocyanin related genes.

In order to confirm the transcriptional pattern revealed by the RNA-seq analysis, we performed qRT-PCR analysis for eight genes, including two phenylpropanoid pathway genes, two EBGs,

two LBGs, as well as two transcription factors (**Figure 9**). All genes showed a similar relative expression level in transgenic lines and J9709 as revealed by RNA-seq analysis. The expression of *PAL1* (BnaA04g21230D) in full-length transgenic lines was similar as in the J9709 control, but it was significantly downregulated in the 395 and 910-isoform transgenic lines. *C4H* (BnaA04g21230D), *CHS* (BnaA10g19670D), *F3'H* (BnaA10g23330D), and *MYB4* (BnaC03g60080D) were significantly upregulated in full-length transgenic lines but were significantly downregulated in the 395 and 910 transgenic lines. *ANS* (BnaA01g12530D) and *DFR* (BnaC09g17150D) were
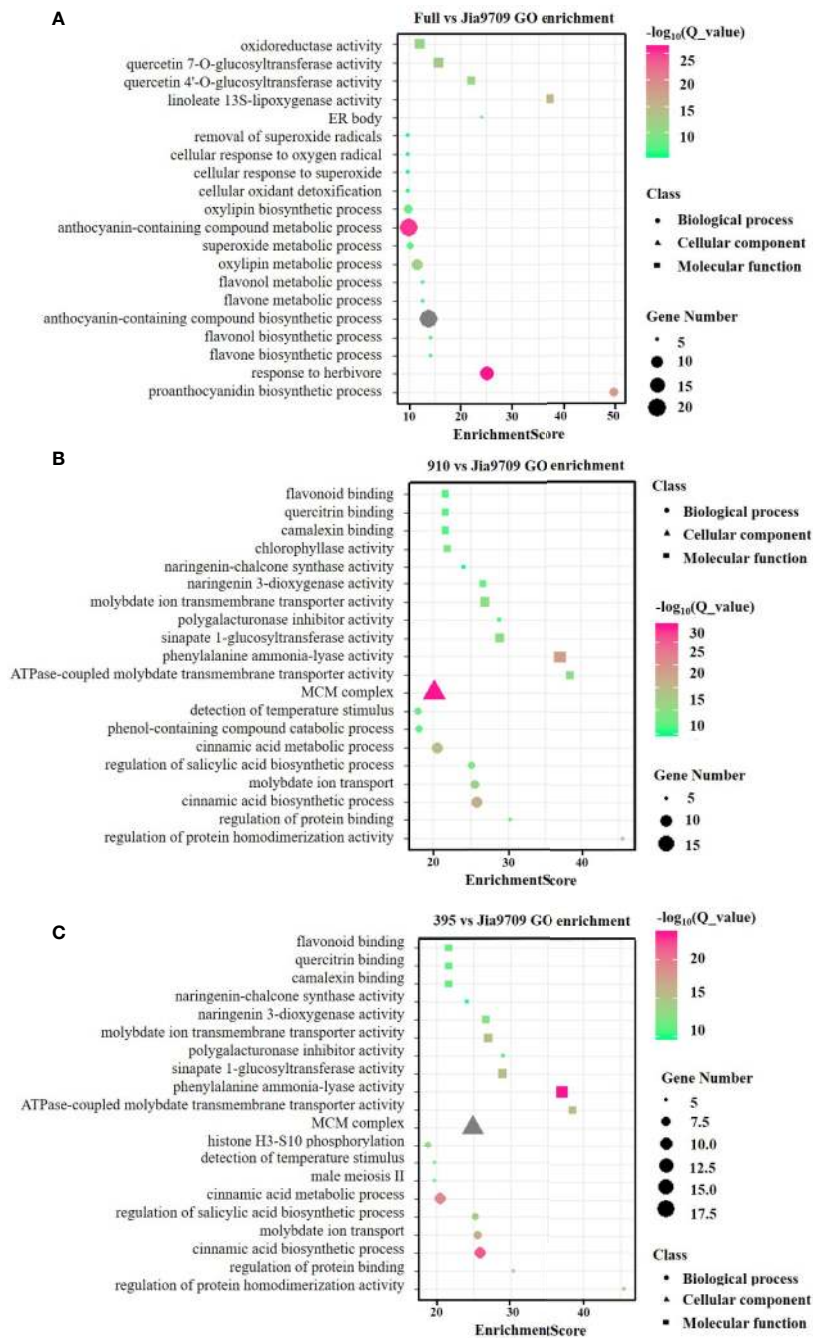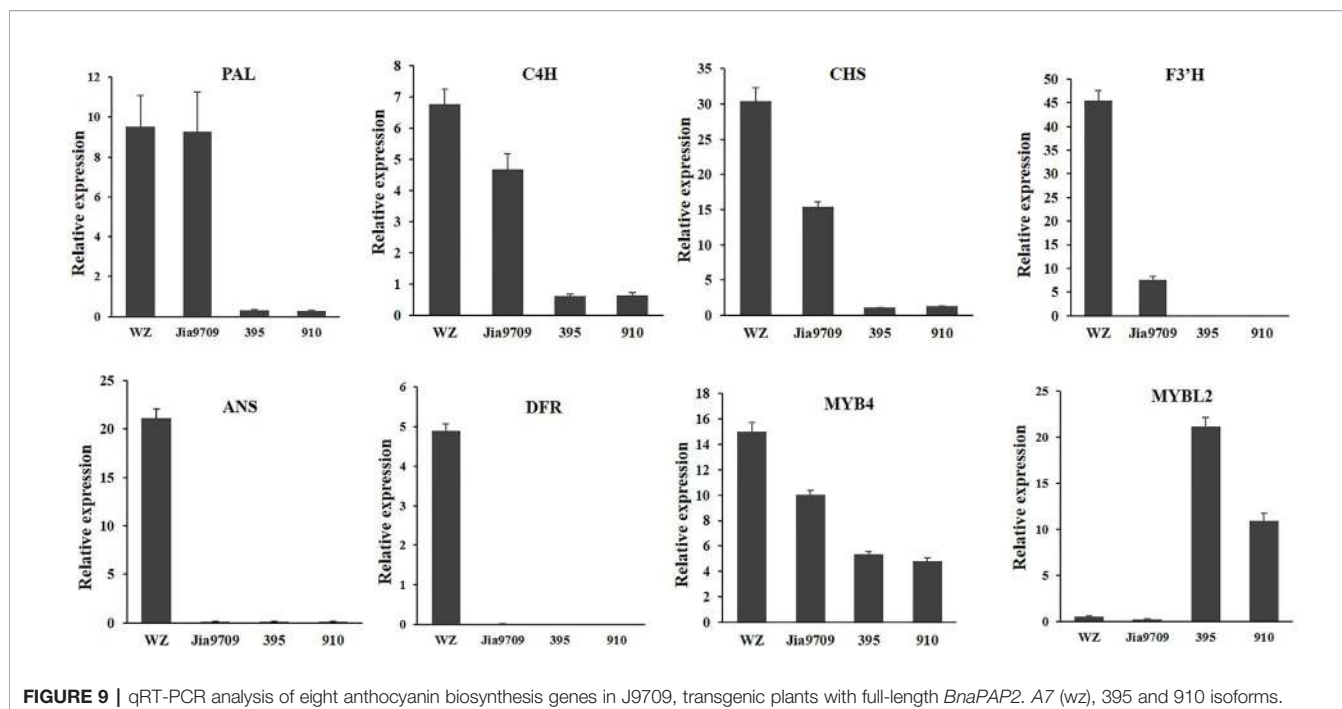
FIGURE 8 | GO enrichment analysis of the DEGs in transgenic lines. **(A-C)** Top twenty enriched GO categories of DEGs in full-length *BnPAP2.A7* transgenic lines **(A)**, 395 **(B)**, and 910 **(C)** isoform transgenic lines.

significantly upregulated in full-length transgenic lines but were almost silenced in J9709 as well as in the 395 and 910 transgenic lines. Interestingly, *MYBL2* (BnaA02g15320D) was expressed at a very low level both in J9709 as well as in full-length transgenic lines but was significantly upregulated in the 395 and 910-isoform transgenic lines (**Figure 9**).

DISCUSSION

In this work, we have demonstrated that one of the eight homologs of *AtPAP2* in rapeseed, *BnaPAP2.A7*, is a key gene responsible for purple leaf formation. Phylogenetic analysis indicated that the A genome *BnaPAP2.A7* is an orthologous

FIGURE 9 | qRT-PCR analysis of eight anthocyanin biosynthesis genes in J9709, transgenic plants with full-length *BnaPAP2. A7* (wz), 395 and 910 isoforms.

gene of the C genome *BoMYB2*, which has previously been shown to be responsible for purple traits in cauliflower, kale, kohlrabi, and cabbage (Chiu et al., 2010; Yan et al., 2019). Moreover, BoMYB2 has been shown to be involved in MYB–bHLH–WD40 formation by direct interaction with BobHLH1. The expression of *BoMYB2* in *Arabidopsis* significantly upregulates a subset of anthocyanin structural genes, including *F3'H*, *DFR*, and *ANS* (Chiu and Li, 2012). Here, both the RNA-seq based transcript and qRT-PCR analysis indicated a significantly higher expression level of *BnaPAP2.A7* in purple leaves of PR than that in green leaves of ZS11 (**Figure 2**). Transgenic analysis in J9709 of the full-length *BnaPAP2.A7.PR* allele that included the 1.99 kb promoter region further confirmed elevated expression in leaves; significant upregulation of late anthocyanin biosynthesis genes appears to contribute to anthocyanin accumulation (**Figure 7**). Yeast two-hybrid studies also indicated that BnaPAP2.A7 could interact with AtTT8 *in vitro* (**Figure 6**). As found with BoMYB2, in leaves of *B. napus*, BnaPAP2.A7 appears to be involved in the formation of MYB–bHLH–WD40 complexes, affecting transcriptional regulation of late anthocyanin biosynthesis genes. However, it should be noted that only transgenic plants with *BnaPAP2.A7.PR* had significant purple color in leaves compared to the presence of purple color in the leaves, stem, and silique of PR. These results suggest that tissue-specific accumulation of anthocyanins in *B. napus* may be controlled by multiple copies of R2R3-MYB. Consistent with this, a paralog of *BnaPAP2.A7* on chromosome C6 has been shown to be responsible for the purple stem trait of PR, based on Bulked Segregant Analysis (BSA) mapping and cloning (Chen DZ et al., unpublished data).

Transgenic analysis indicated that the 1.99 kb promoter region from PR is capable of activating the expression of *BnaPAP2.A7* in different genetic backgrounds. However, only a single bp insertion (T) and 2bp insertion (TC) were present in the promoter region of *BnaPAP2.A7. PR* compared to ZS11 (**Figures 3** and **7**). In addition, *BnaPAP.A7* was highly expressed in both PR and transgenic J9709 in outdoor compared to greenhouse environments, which suggests that its expression may be UV dependent (**Figure 2**). However, the mechanism for such activation remains unknown since no known motif was identified at these sites. Similarly, the activation of the *BoMYB2* in purple cabbage was recently revealed to be caused by a point mutation at -408 and/or the 1-bp insertion at -95 in the promoter region (Yan et al., 2019). However, this was caused by the insertion of 4 kb Harbinger DNA transposon in purple cauliflower and by the insertion of a 7.6kb CACTA DNA transposon in both purple kale and kohlrabi (Chiu et al., 2010; Yan et al., 2019). In fact, the activation of the R2R3-MYB related to the anthocyanin biosynthesis is often caused by insertion of a transposon, as found for example in *Antirrhinum majus* (Lister et al., 1993), blood orange (Butelli et al., 2012), tea (Sun et al., 2016), and *Phalaenopsis* orchids (Hsu et al., 2019). Whether there are other possible elements that could enhance the expression of *BnaPAP2.A7* in PR beyond the 2 kb upstream is worthy of further study.

In this study, we found that two of the three alternative transcripts of *BnaPAP.A7* encoded truncated proteins lacking a partial R3 repeat and complete C terminal domain. Plant R2R3-MYBs contain two repeats (R2 and R3), each forming a helix-turn-helix motif. It has been proposed that the third helices of

both the R2 and R3 domains are critical for the DNA-binding activity (Gabrielsen et al., 1991; Jia et al., 2004), whereas the first two α-helices of the R3 domain are involved in interactions with the bHLH protein (Zimmermann et al., 2004). Here, the two truncated proteins had both lost part of the second and third helices of the R3 repeat (**Figure 5**). Although there is evidence that the two truncated proteins localized in the nucleus, they had lost the ability to bind AtTT8 (**Figure 6**). Unexpectedly, the two proteins appeared to contribute to strong repression of genes involved in the phenylpopanid pathway and anthocyanin biosynthesis (**Figure S2**). This suggests that the two isoforms encoded by spliced transcripts act as repressors of anthocyanin biosynthesis, in direct contrast to the activator role encoded by the normally spliced isoform. Similarly, different alleles of Ruby2, a novel R2R3-MYB regulatory gene, have recently been shown to have opposite effects on the regulation of anthocyanin biosynthesis in *Citrus* (Huang et al., 2018). While *Ruby2*^{Full} encodes an anthocyanin activator, *Ruby2*^{Short} encodes a repressor with partial R2 and R3 repeats due to an alternative downstream ATG start codon and a 54 bp in-frame deletion. *Ruby2*^{Short} has lost the ability to activate anthocyanin biosynthesis genes but retains the ability to interact with bHLH1, thus acting as a passive competitor in the regulatory complex (Huang et al., 2018). However, because the two rapeseed isoforms are unable to interact with AtTT8, it is unclear how they inhibit the expression of genes related to anthocyanin biosynthesis.

A possible explanation may be found by considering that many anthocyanin biosynthesis genes were significantly downregulated in the two AS transcript transgenic lines, which likely resulted from the repression of AtMYB4-like repressors (Chen et al., 2019). AtMYB4, the first transcription factor found to function as a repressor of the phenylpropanoid pathway, is light-responsive and affects one or more phenylpropanoid structural genes, especially for the C4H (Borevitz et al., 2000; Jin et al., 2000; Hemm et al., 2001). In *B. rapa* subsp. *rapa* cv. Tsuda, BrMYB4, a homologous gene of AtMYB4, has also been shown to bind to the promoter region of BrC4H directly and repress expression of C4H, with negative regulation of all anthocyanin biosynthesis. In addition, suppression of *ANS*, *DFR*, and UFGT has also been consistently observed in all transgenic tobacco and other plants by heterologous expression of AtMYB4-like repressors (Pérez-Díaz et al., 2016; Xu et al., 2017; Anwar et al., 2018). However, RNA-seq and qRT-PCR analysis did not detect activation of MYB4 homologous genes by the 395 and 910 isoforms. This indicates that, although the 395 and 910 isoforms repressed the anthocyanin biosynthesis genes, this was not by activating MYB4. Interestingly, an R3 MYB repressor, MYBL2, was upregulated in the 395 and 910 transgenic plants (**Supplementary Datasheet 3**; **Figure 9**). In *Arabidopsis*, AtMYBL2, which contains an R3 domain and part of an R2 domain, is also a negative regulator of anthocyanin biosynthesis (Dubos et al., 2008; Matsui et al., 2008). Phylogenetic analyses indicate that MYBL2-like R3-MYB repressors are closely related to R2R3-MYB repressors and might have originated *via* tandem duplication followed by

large deletions in the R2 domain (Chen et al., 2019). It is possible that 395 and 910 isoforms could activate MYBL2 and so repress the expression of anthocyanin biosynthesis genes. A further study should be performed to test the overexpression of 395 and 910 isoforms in PR or other model plants such as tobacco to confirm their negative roles in anthocyanin biosynthesis. Further study should also focus on the potential binding activity of the two rapeseed truncated proteins using methods such as ChiP-seq.

In summary, our study has identified and characterized *BnaPAP2.A7*, an R2R3-MYB activator of anthocyanin biosynthesis, in *B. napus*. We have shown that it plays a key role in the regulation of purple color formation in leaves. We have also uncovered a potential new negative feedback loop controlling flavonoid biosynthesis involving alternatively spliced transcripts, which will contribute to our understanding of the homeostatic mechanism of flavonoid accumulation in plants.

DATA AVAILABILITY STATEMENT

Genomic sequences of BnaPAP2.A7 from PR, ZS11 and J9709 were deposited in NCBI, with accession number 2292245, 2292253, and MT438412. The coding sequence of two alternative spliced isoforms were deposited in NCBI, with accession number 2244186 and 2244486. Transcriptome data were deposited in NCBI under BioProject ID: PRJNA554517. The sequences of primer sets in this study were listed in **Table S3**.

AUTHOR CONTRIBUTIONS

XG, JW, and ZL conceived and designed the experiments. DC, YL, SY, QJ, and JQ performed the experiments. DC, XG, and JW analyzed the data, DC, XG, GK, and YL wrote the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2020.00983/full#supplementary-material

REFERENCES

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

Albert, N. W., Davies, K. M., Lewis, D. H., Zhang, H., Montefiori, M., Brendolise, C., et al. (2014). A conserved network of transcriptional activators and repressors regulates anthocyanin pigmentation in eudicots. *Plant Cell* 26, 962–980. doi: 10.1105/tpc.113.122069

Anders, S., Pyl, P. T., and Huber, W. (2015). HTSeq-a Python framework to work with high-throughput sequencing data. *Bioinformatics* 31, 166–169. doi: 10.1093/bioinformatics/btu638

Anwar, M., Wang, G., Wu, J., Waheed, S., Allan, AC., and Zeng, L. (2018). Ectopic overexpression of a novel R2R3-MYB, NtMYB2 from Chinese narcissus represses anthocyanin biosynthesis in tobacco. *Molecules* 23, 781. doi: 10.3390/molecules23040781

Bolger, A. M., Lohse, M., and Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30, 2114–2120. doi: 10.1093/bioinformatics/btu170

Borevitz, J. O., Xia, Y., Blount, J., Dixon, R. A., and Lamb, C. (2000). Activation tagging identifies a conserved MYB regulator of phenylpropanoid biosynthesis. *Plant Cell* 12, 2383–2394. doi: 10.1105/tpc.12.12.2383

Broun, P. (2005). Transcriptional control of flavonoid biosynthesis: a complex network of conserved regulators involved in multiple aspects of differentiation in Arabidopsis. *Curr. Opin. Plant Biol.* 8, 272–279. doi: 10.1016/j.pbi.2005.03.006

Butelli, E., Licciardello, C., Zhang, Y., Liu, J., Mackay, S., Bailey, P., et al. (2012). Retrotransposons control fruit-specifc, cold-dependent accumulation of anthocyanins in blood oranges. *Plant Cell* 24, 1242–1255. doi: 10.1105/tpc.111.095232

Cardoza, V., and Stewart, C. N. (2003). Increased Agrobacterium-mediated transformation and rooting efficiencies in canola (*Brassica napus* L.) from hypocotyl segment explants. *Plant Cell Rep.* 21, 599–604. doi: 10.1007/s00299-002-0560-y

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, 1448–1470. doi: 10.1104/pp.114.256172

Chalhoub, B., Denoeud, F., Liu, S. Y., Parkin, I.A.P., Tang, H.B., Wang, X.Y., et al. (2014). Early allopolyploid evolution in the post-Neolithic Brassica napus oilseed genome. *Science* 345, 950. doi: 10.1126/science.1253435

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

Chen, L., Hu, B., Qin, Y., Hu, G., and Zhao, J. (2019). Advance of the negative regulation of anthocyanin biosynthesis by MYB transcription factors. *Plant Physiol. Biochem.* 136, 178–187. doi: 10.1016/j.plaphy.2019.01.024

Chiu, L. W., and Li, L. (2012). Characterization of the regulatory network of BoMYB2 in controlling anthocyanin biosynthesis in purple cauliflower. *Planta* 236, 1153–1164. doi: 10.1007/s00425-012-1665-3

Chiu, L. W., Zhou, X., Burke, S., Wu, X., Prior, R. L., and Li, L. (2010). The purple cauliflower arises from activation of a MYB transcription factor. *Plant Physiol.* 154, 1470–1480. doi: 10.1104/pp.110.164160

Cone, K. C., Burr, F. A., and Burr, B. (1986). Molecular analysis of the maize anthocyanin regulatory locus cl. *Proc. Natl. Acad. Sci. U.S.A.* 83, 9631–9635. doi: 10.1073/pnas.83.24.9631

Cone, K. C., Cocciolone, S. M., Moehlenkamp, C. A., Weber, T., Drummond, B. J., Tagliani, L. A., et al. (1993). Role of the regulatory gene pl in the photocontrol of maize anthocyanin pigmentation. *Plant Cell* 5, 1807–1816. doi: 10.1105/tpc.5.12.1807

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

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

Feller, A., Machemer, K., Braun, E. L., and Grotewold, E. (2011). Evolutionary and comparative analysis of MYB and bHLH plant transcription factors. *Plant J.* 66, 94–116. doi: 10.1111/j.1365-313X.2010.04459.x

Fu, W., Chen, D., Pan, Q., Li, F., Zhao, Z., Ge, X., et al. (2018). Production of red-flowered oilseed rape via the ectopic expression of *Orychophragmus violaceus* OvPAP2. *Plant Biotechnol. J.* 16, 367–380. doi: 10.1111/pbi.12777

Gabrielsen, O. S., Sentenac, A., and Fromageot, P. (1991). Specific DNA binding by c-Myb: evidence for a double helix-turn-helix-related motif. *Science* 6, 1140–1143. doi: 10.1126/science.1887237

Gonzalez, A., Zhao, M., Leavitt, J. M., and Lloyd, A. M. (2008). Regulation of the anthocyanin biosynthetic pathway by the TTG1/bHLH/Myb transcriptional complex in Arabidopsis seedlings. *Plant J.* 53, 814–827. doi: 10.1111/j.1365-313X.2007.03373.x

Gookin, T. E., and Assmann, S. M. (2014). Significant reduction of BiFC non-specific assembly facilitates in planta assessment of heterotrimeric G-protein interactors. *Plant J.* 80, 553–567. doi: 10.1111/tpj.12639

Gould, K. S. (2004). Nature's Swiss army knife: the diverse protective roles of anthocyanins in leaves. *J. BioMed. Biotechnol.* 2004, 314–320. doi: 10.1155/S1110724304406147

He, Q., Zhang, Z. F., and Zhang, L. G. (2016). Anthocyanin Accumulation, Antioxidant Ability and Stability, and a Transcriptional Analysis of Anthocyanin Biosynthesis in Purple Heading Chinese Cabbage (*Brassica rapa* L. ssp. pekinensis). *J. Agr. Food Chem.* 64, 132–145. doi: 10.1021/acs.jafc.5b04674

Hemm, M. R., Herrmann, K. M., and Chapple, C. (2001). AtMYB4: a transcription factor general in the battle against UV. *Trends Plant Sci.* 6, 135–136. doi: 10.1016/S1360-1385(01)01915-X

Hsu, C. C., Su, C. J., Jeng, M. F., Chen, W. H., and Chen, H. H. (2019). A HORT1 Retrotransposon Insertion in the PeMYB11 Promoter Causes Harlequin/Black Flowers in Phalaenopsis Orchids. *Plant Physiol.* 180, 1535–1548. doi: 10.1104/pp.19.00205

Huang, D., Wang, X., Tang, Z., Yuan, Y., Xu, Y., He, J., et al. (2018). Subfunctionalization of the Ruby2-Ruby1 gene cluster during the domestication of citrus. *Nat. Plant* 4, 930–941. doi: 10.1038/s41477-018-0287-6

Izzah, N. K., Lee, J., Perumal, S., Park, J. Y., Ahn, K., Fu, D. H., et al. (2013). Microsatellite-based analysis of genetic diversity in 91 commercial Brassica oleracea L. cultivars belonging to six varietal groups. *Genet. Res. Crop Evo.* 60, 1967–1986. doi: 10.1007/s10722-013-9966-3

Jia, L., Clegg, M. T., and Jiang, T. (2004). Evolutionary dynamics of the DNA-binding domains in putative R2R3-MYB genes identified from rice subspecies indica and japonica genomes. *Plant Physiol.* 134, 575–585. doi: 10.1104/pp.103.027201

Jin, H., Cominelli, E., Bailey, P., Parr, A., Mehrtens, F., Jones, J., et al. (2000). Transcriptional repression by AtMYB4 controls production of UV-protecting sunscreens in Arabidopsis. *EMBO J.* 19, 6150–6161. doi: 10.1093/emboj/19.22.6150

Jin, W., Wang, H., Li, M., Wang, J., Yang, Y., Zhang, X., et al. (2016). The R2R3 MYB transcription factor PavMYB10.1 involves in anthocyanin biosynthesis and determines fruit skin colour in sweet cherry (*Prunus avium* L.). *Plant Biotechnol. J.* 14, 2120–2133. doi: 10.1111/pbi.12568

Jones, D. T., Taylor, W. R., and Thornton, J. M. (1992). The rapid generation of mutation data matrices from protein sequences. *Comput. Appl. Biosci.* 8, 275–282. doi: 10.1093/bioinformatics/8.3.275

Jun, J. H., Liu, C., Xiao, X., and Dixon, R. A. (2015). The transcriptional repressor MYB2 regulates both spatial and temporal patterns of proanthocyandin and anthocyanin pigmentation in Medicago truncatula. *Plant Cell* 27, 2860–2879. doi: 10.1105/tpc.15.00476

Kim, D., Pertea, G., Trapnell, C., Pimentel, H., Kelley, R., and Salzberg, S. L. (2013). TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biol.* 14, R36. doi: 10.1186/gb-2013-14-4-r36

Kumar, S., Stecher, G., and Tamura, K. (2016). MEGA7: Molecular Evolutionary Genetics Analysis version 7.0 for bigger datasets. *Mol. Biol. Evol.* 33, 1870–1874. doi: 10.1093/molbev/msw054

Lai, B., Li, X. J., Hu, B., Qin, Y. H., Huang, X. M., Wang, H. C., et al. (2014). LcMYB1 is a key determinant of differential anthocyanin accumulation among genotypes, tissues, developmental phases and ABA and light stimuli in Litchi chinensis. *PloS One* 9, e86293. doi: 10.1371/journal.pone.0086293

Li, Z., and Heneen, W. K. (1999). Production and cytogenetics of intergeneric hybrids between the three cultivated Brassica diploids and *Orychophragmus violaceus*. *Theor. Appl. Genet.* 99, 694–704. doi: 10.1007/s001220051286

Li, J., Li, X., Guo, L., Lu, F., Feng, X., He, K., et al. (2006). A subgroup of MYB transcription factor genes undergoes highly conserved alternative splicing in Arabidopsis and rice. *J. Exp. Bot.* 57, 1263–1273. doi: 10.1093/jxb/erj094

Lister, C., Jackson, D., and Martin, C. (1993). Transposon-induced inversion in antirrhinum modifies nivea gene-expression to give a novel flower color pattern under the control of cycloidea (radialis). *Plant Cell* 5, 1541–1553. doi: 10.1105/tpc.5.11.1541

Liu, J. Y., Osbourn, A., and Ma, P. D. (2015). MYB Transcription Factors as Regulators of Phenylpropanoid Metabolism in Plants. *Mol. Plant* 8, 689–708. doi: 10.1016/j.molp.2015.03.012

Liu, Y., Lin-Wang, K., Espley, R. V., Wang, L., Li, Y., Liu, Z., et al. (2019). StMYB44 negatively regulates anthocyanin biosynthesis at high temperatures in tuber flesh of potato. *J. Exp. Bot.* 70, 3809–3824. doi: 10.1093/jxb/erz194

Marquez, Y., Brown, J. W., Simpson, C., Barta, A., and Kalyna, M. (2012). Transcriptome survey reveals increased complexity of the alternative splicing landscape in Arabidopsis. *Genome Res.* 22, 1184–1195. doi: 10.1101/gr.134106.111

Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet J.* 17, 10–12. doi: 10.14806/ej.17.1.200

Matsui, K., Umemura, Y., and Ohme-Takagi, M. (2008). AtMYBL2, a protein with a single MYB domain, acts as a negative regulator of anthocyanin biosynthesis in Arabidopsis. *Plant J.* 55, 954–967. doi: 10.1111/j.1365-313X.2008.03565.x

Mushtaq, M. A., Pan, Q., Chen, D., Zhang, Q., Ge, X., and Li, Z. (2016). Comparative Leaves Transcriptome Analysis Emphasizing on Accumulation of Anthocyanins in Brassica: Molecular Regulation and Potential Interaction with Photosynthesis. *Front. Plant Sci.* 7:311. doi: 10.3389/fpls.2016.00311

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 protooncogene products and with structural similarities to transcriptional activators. *EMBO J.* 6, 3553–3558. doi: 10.1002/j.1460-2075.1987.tb02684.x

Pérez-Díaz, J. R., Pérez-Díaz, J., Madrid-Espinoza, J., González-Villanueva, E., Moreno, Y., and Ruiz-Lara, S. (2016). New member of the R2R3-MYB transcription factors family in grapevine suppresses the anthocyanin accumulation in the flowers of transgenic tobacco. *Plant Mol. Biol.* 90, 63–76. doi: 10.1007/s11103-015-0394-y

Petroni, K., and Tonelli, C. (2011). Recent advances on the regulation of anthocyanin synthesis in reproductive organs. *Plant Sci.* 181, 219–229. doi: 10.1016/j.plantsci.2011.05.009

Qin, Z., Wu, J., Geng, F., Feng, N., Chen, J., Kong, C., et al. (2017). Regulation of FT splicing by an endogenous cue in temperate grasses. *Nat. Commun.* 8, 14320. doi: 10.1038/ncomms14320

Ramsay, N. A., and Glover, B. J. (2005). MYB-bHLH-WD40 protein complex and the evolution of cellular diversity. *Trends Plant Sci.* 10, 63–70. doi: 10.1016/j.tplants.2004.12.011

Saito, K., Yonekura-Sakakibara, K., Nakabayashi, R., Higashi, Y., Yamazaki, M., Tohge, T., et al. (2013). The flavonoid biosynthetic pathway in Arabidopsis: structural and genetic diversity. *Plant Physiol. Biochem.* 72, 21–34. doi: 10.1016/j.plaphy.2013.02.001

Schwinn, K., Venail, J., Shang, Y., Mackay, S., Alm, V., Butelli, E., et al. (2006). A small family of MYB-regulatory genes controls floral pigmentation intensity and patterning in the genus Antirrhinum. *Plant Cell* 18, 831–851. doi: 10.1105/tpc.105.039255

Shang, Y., Venail, J., Mackay, S., Bailey, P. C., Schwinn, K. E., Jameson, P. E., et al. (2011). The molecular basis for venation patterning of pigmentation and its effect on pollinator attraction in flowers of Antirrhinum. *New Phytol.* 189, 602–615. doi: 10.1111/j.1469-8137.2010.03498.x

Stracke, R., Werber, M., and Weisshaar, B. (2001). The R2R3-MYB gene family in Arabidopsis thaliana. *Curr. Opin. Plant Biol.* 4, 447–456. doi: 10.1016/S1369-5266(00)00199-0

Stracke, R., Ishihara, H., Huep, G., Barsch, A., Mehrtens, F., Niehaus, K., et al. (2007). Differential regulation of closely related R2R3-MYB transcription factors controls flavonol accumulation in different parts of the *Arabidopsis thaliana* seedling. *Plant J.* 50, 660–677. doi: 10.1111/j.1365-313X.2007.03078.x

Sun, B., Zhu, Z., Cao, P., Chen, H., Chen, C., Zhou, X., et al. (2016). Purple foliage coloration in tea (*Camellia sinensis* L.) arises from activation of the R2R3-MYB transcription factor CsAN1. *Sci. Rep.* 6, 32534. doi: 10.1038/srep32534

Syed, N. H., Kalyna, M., Marquez, Y., Barta, A., and Brown, J. W. (2012). Alternative splicing in plants–coming of age. *Trends Plant Sci.* 17, 616–623. doi: 10.1016/j.plants.2012.06.001

Tamagnone, L., Merida, A., Parr, A., Mackay, S., Culianez-Macia, F. A., Roberts, K., et al. (1998). The AmMYB308 and AmMYB330 transcription factors from Antirrhinum regulate phenylpropanoid and lignin biosynthesis in transgenic tobacco. *Plant Cell* 10, 135–154. doi: 10.1105/tpc.10.2.135

Trapnell, C., Roberts, A., Goff, L., Pertea, G., Kim, D., Kelley, D. R., et al. (2012). Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. *Nat. Prot.* 7, 562–578. doi: 10.1038/nprot.2012.016

Walker, A. R., Lee, E., Bogs, J., McDavid, D. A., Thomas, M. R., and Robinson, S. P. (2007). White grapes arose through the mutation of two similar and adjacent regulatory genes. *Plant J.* 49, 772–785. doi: 10.1111/j.1365-313X.2006.02997.x

Xu, W., Grain, D., Bobet, S., Le Gourrierec, J., Thévenin, J., Kelemen, Z., et al. (2014). Complexity and robustness of the flavonoid transcriptional regulatory network revealed by comprehensive analyses of MYB-bHLH-WDR complexes and their targets in Arabidopsis seed. *New Phytol.* 202, 132–144. doi: 10.1111/nph.12620

Xu, W., Dubos, C., and Lepiniec, L. (2015). Transcriptional control of flavonoid biosynthesis by MYB-bHLH-WDR complexes. *Trends Plant Sci.* 20, 176–185. doi: 10.1016/j.tplants.2014.12.001

Xu, C., Huang, Q., Ge, X., and Li, Z. (2019). Phenotypic, cytogenetic, and molecular marker analysis of Brassica napus introgressants derived from an intergeneric hybridization with Orychophragmus. *PloS One* 14, e0210518. doi: 10.1371/journal.pone.0210518

Xu, H., Wang, N., Liu, J., Qu, C., Wang, Y., Jiang, S., et al. (2017). The molecular mechanism underlying anthocyanin metabolism in apple using the MdMYB16 and MdBHLH33 genes. *Plant Mol. Biol.* 94, 149–165.

Yan, C. H., An, G. H., Zhu, T., Zhang, W. Y., Zhang, L., Peng, L. Y., et al. (2019). Independent activation of the *BoMYB2* gene leading to purple traits in *Brassica oleracea*. *Theor. Appl. Genet.* 132, 895–906. doi: 10.1007/s00122-018-3245-9

Yuan, Y., Chiu, L. W., and Li, L. (2009). Transcriptional regulation of anthocyanin biosynthesis in red cabbage. *Planta* 230, 1141–1153. doi: 10.1007/s00425-009-1013-4

Zhang, Y., Hu, Z., Zhu, M., Zhu, Z., Wang, Z., Tian, S., et al. (2015). Anthocyanin accumulation and molecular analysis of correlated genes in purple kohlrabi (*Brassica oleracea* var. gongylodes L.). *J. Agric. Food Chem.* 63, 4160–4169. doi: 10.1021/acs.jafc.5b00473

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.* 40, 22–34. doi: 10.1111/j.1365-313X.2004.02183.x