Frequent gain- and loss-of-function mutations of the *BjMYB113* gene accounted for leaf color variation in *Brassica juncea*

Guanghui An and Jiongjiong Chen *

**Abstract**

**Background**: Mustard (*Brassica juncea*) is an important economic vegetable, and some cultivars have purple leaves and accumulate more anthocyanins than the green. The genetic and evolution of purple trait in mustard has not been well studied.

**Result**: In this study, free-hand sections and metabolomics showed that the purple leaves of mustard accumulated more anthocyanins than green ones. The gene controlling purple leaves in mustard, *Mustard Purple Leaves* (*MPL*), was genetically mapped and a MYB113-like homolog was identified as the candidate gene. We identified three alleles of the MYB113-like gene, *BjMYB113a* from a purple cultivar, *BjMYB113b* and *BjMYB113c* from green cultivars. A total of 45 single nucleotide polymorphisms (SNPs) and 8 InDels were found between the promoter sequences of the purple allele *BjMYB113a* and the green allele *BjMYB113b*. On the other hand, the only sequence variation between the purple allele *BjMYB113a* and the green allele *BjMYB113c* is an insertion of 1,033-bp fragment in the 3’ region of *BjMYB113c*. Transgenic assay and promoter activity studies showed that the polymorphism in the promoter region was responsible for the up-regulation of the purple allele *BjMYB113a* and high accumulation of anthocyanin in the purple cultivar. The up-regulation of *BjMYB113a* increased the expression of genes in the anthocyanin biosynthesis pathway including *BjCHS*, *BjF3H*, *BjF3’H*, *BjDFR*, *BjANS* and *BjUGFT*, and consequently led to high accumulation of anthocyanin. However, the up-regulation of *BjMYB113* was compromised by the insertion of 1,033-bp in 3’ region of the allele *BjMYB113c*.

**Conclusions**: Our results contribute to a better understanding of the genetics and evolution of the *BjMYB113* gene controlling purple leaves and provide useful information for further breeding programs of mustard.

**Keywords**: Leaf color, MYB transcription factor, BSR-seq, Map-based cloning, *Brassica juncea*

**Background**

Mustard (*Brassica juncea*, 2n = 36, AABB) is an important economic vegetable in the world [1, 2]. This cultivated species contains several different varieties, which exhibit extreme morphologic polymorphisms, with leaves, stems, roots, or seeds as harvest organs. For leafy mustard cultivars, some have green leaves while others have purple leaves. The purple color is caused by the accumulation of anthocyanins in leaf epidermal cells [3]. Anthocyanins are a large group of water-soluble natural pigments and widely distribute in higher plants. They are largely responsible for the red, purple, and blue colors in flowers, fruits, leaves, seeds and other plant organs [4, 5]. In plants, anthocyanins play important roles in attracting pollinators and seed distributors [5]. Anthocyanins can protect plants from UV damage, and they have strong capacity to scavenge reactive oxygen species...
tural crops showed rich color polymorphisms. The gene \( \text{BjPur} \) between a purple and a green cultivar suggested that mechanisms underlying purple mustard remain unclear. However, the function of the \( \text{BjPur} \) gene in mustard has not been verified, and the genetic and molecular proteins, such as NACs, HY5, ERFs, BBX22, and WRKY, were found to be involved in anthocyanin biosynthesis as well. These transcription factors can bind to the promoters of component genes of the MBW complex directly or indirectly to regulate their expressions and consequently anthocyanin accumulation in many plant species such as peach [22], apple [23], pear [24], grape [25], blood orange [26], petunia [27], rice [28], lettuce [29], bokchoy [30], kale [31], cauliflower [32], barrel medic [33], etc. In addition to the MBW complex, other transcription factors, such as NACs, HY5, ERFs, BBX22, and WRKY, were found to be involved in anthocyanins in mustard cultivars.

### Results

**Characterization of anthocyanin in mustard cultivars**

Mustard has both green and purple cultivars (Fig. 1a, b, c). We compared the total anthocyanin contents of one purple mustard cultivar (pl102) and two green cultivars (rt104 and gre101). As expected, the total anthocyanin content in the purple cultivar was considerably higher than those in the two green cultivars (Fig. 2a). Free-hand sections of leaf tissues of the purple cultivar showed that the anthocyanins accumulated mainly in the epidermal cells and in less extent in the mesophyll cells near the epidermises. In contrast, no obvious anthocyanins were found in green cultivars (Fig. 1d).

We further used Liquid Chromatography-Electrospray Ionization-Mass Spectrometry (LC-ESI-MS) to analyze the types of anthocyanins and their concentrations in the purple and green cultivars. A total of 18 anthocyanins were identified (Table S1). Of them, six had significantly higher concentration in the purple cultivar than in green one \( p < 0.05 \), including Delphinidin 3,5-O-diglucoside, Delphinidin 3-O-(6′-O-malonyl)-beta-D-glucoside, Cyanidin 3-O-(6′-O-malonyl)-beta-D-glucoside, Pelargonidin 3-O-glucoside, Peonidin 3,5-O-diglucoside, and Peonidin 3-O-glucoside, all emitted dark color (Fig. 2b). We conclude that the purple color in mustard leaves was caused by high accumulation of these anthocyanins.
Map-based cloning of the gene controlling color variation in mustard

As mentioned in Background section, a previous study suggested that an insertion of 1,268-bp fragment in the first intron of a MYB gene was responsible for its loss-of-function and green leaves [3]. Furthermore, some green cultivars such as rt104 and gre101 do not have the 1,268-bp insertion, and therefore should have distinct mechanisms for loss-of-function. To answer these questions, purple cultivar pl102 was crossed with green cultivars rt104 and gre101, respectively. Both F₁ hybrids were purple, and they were selfed to generate two F₂ segregating populations.

Of the 137 individuals in the F₂ segregating population derived from the cross pl102 × rt104, 102 individuals had purple leaves and 35 individuals had green leaves, with an expected ratio of 3:1 ($\chi^2 = 0.0219 < \chi^2(0.05,1) = 3.84$, $P > 0.05$, Table S2), suggesting a single gene controlling the color variation, which was referred to as Mustard Purple Leaves (MPL) in this study. Then, we used BSR-seq to genetically map the gene underlying the color variation. The MPL gene was mapped to chromosome J15 (B05) (Fig. 3a). Screening a total of 1,353 individuals from the F₂ population ultimately mapped the MPL gene between markers AGH260 and AGH263 on J15 (B05), in an interval of approximately 170 kb region (Fig. 3b, c). Only 12 open reading frames (ORFs) were predicted in this interval. The MYB gene, which is an ortholog of MYB113 in Arabidopsis and was previously shown to control color variation in mustard, was also located in the candidate interval and remained a reliable candidate gene (BjMYB113) (Fig. 3d).

As mentioned above, the 1,268-bp insertion, which was suggested to inactivate the candidate gene in a previous study, was absent in the green parent [3]. To verify the BjMYB113 as the candidate gene, we first investigated its sequence variation between the two parents, including its 2,842 bp upstream sequences. A total of 45 SNPs and 8 InDels were found between the two alleles from the purple parent pl102 and the green parent rt104 (Fig. 4).

Interestingly, such sequence variation was not found between the purple parent pl102 and another green cultivar gre101. The BjMYB113 gene from green cultivar gre101 is almost identical to that from the purple cultivar, with only one SNP at -2,269 of promoter region (Fig. 4), and the BjMYB113 gene from either cultivar does not have the 1,268-bp insertion in intron 1. BSR-seq analysis of the pl102 × gre101 segregating population also showed that the gene controlling leaf color was located in the same region as BjMYB113 on chromosome J15 (Fig. S1) (see above). We then sequenced its flanking region and discovered an insertion of 1,033-bp fragment at 2,873 bp downstream its stop codon. Therefore, there are at least four alleles of the BjMYB113 gene, including BjMYB113a (purple allele), BjMYB113b (green allele) and BjMYB113c (green allele with the 1,033-bp...
insertion in its 3’ downstream), as well as an allele identified previously with an insertion of 1,268-bp fragment in the first intron (Fig. 4).

Overexpressing the BjMYB113a, BjMYB113b, or BjMYB113c alleles promoted anthocyanin accumulation in Arabidopsis

The coding sequence of the three alleles BjMYB113a, BjMYB113b and BjMYB113c are highly conserved in different mustard cultivars (Fig. S2). The coding sequence of BjMYB113a and BjMYB113c are identical, while BjMYB113a and BjMYB113b vary by three SNPs, with only one of them leading to amino acid change (M120K) (Fig. S2). To verify the function of these three BjMYB113 alleles, their coding sequence were driven by cauliflower mosaic virus (CaMV) 35 S promoter and transferred into A. thaliana Col-0. All Arabidopsis lines overexpressing BjMYB113 alleles had purple leaves and stems in comparison with green leaves and stems in wild type (Fig. 5c, d, e, f). As expected, the color change was due to high accumulation of anthocyanins and high expression of BjMYB113 (Fig. 5a, b). Above results indicated that the expression level of BjMYB113a (from purple cultivar) was significantly higher than BjMYB113b and BjMYB113c (from green cultivars) (Fig. 6a). We hypothesize that the expression difference between alleles BjMYB113a and BjMYB113b was caused by their promoter regions, and the expression difference between alleles BjMYB113a and BjMYB113c was caused by the 1,033-bp insertion in downstream region of the latter.

To test above hypothesis, promoter activity of the three BjMYB113 alleles was analyzed. GUS gene driven by the promoter of BjMYB113a, BjMYB113b, and BjMYB113c, respectively, was transformed into Arabidopsis Col-0. Six transgenic lines were obtained for each allele, and the GUS activities of these transgenic lines were measured. The GUS activity driven by promoter of BjMYB113a or BjMYB113c was significantly higher than that driven by the promoter of BjMYB113b, but no activity difference was observed between the promoters of BjMYB113a and BjMYB113c (Fig. 6b). Therefore, the polymorphisms in the promoter region accounts for the expression difference between BjMYB113a and BjMYB113b, and the low expression of BjMYB113c was likely caused by the 1,033-bp insertion in its downstream region rather than the point mutation in its promoter region.

Forty-five SNPs and eight InDels were found between the promoter regions of BjMYB113a and BjMYB113b. The motif analysis was carried out for these two promoters. The activation of the BjMYB113a in purple cultivar pl102 was likely due to the point mutations at -1,317 (C to G), -1,576 (A to
These four SNPs produced new transcription factor binding sites, including a MBSI motif (AAAAAAC(G/C)GTTA), a bHLH motif (CANNTG), and two TATA-box motifs (ATTATA).

The 1,033-bp insertion suppresses the expression of \textit{BjMYB113c}

To verify the effects of the 1,033-bp insertion on the expression of \textit{BjMYB113}, the \textit{BjMYB113c} allele, including its native promoter and 3’ region was transformed to
Arabidopsis Col-0. No color change was observed in positive transgenic lines (Fig. 5i). In contrast, transgenic line of a construct containing BjMYB113a gene with its native promoter and 3’downstream sequences did change the color from green to purple (Fig. 5g). Note the main sequence difference between above two constructs was the 1,033-bp insertion. We conclude that high expression of BjMYB113a was due to mutations in its promoter region, and that the low expression of BjMYB113c was caused by the suppression effects of the 1,033-bp fragment inserted in its downstream region.

**Fig. 5** The total anthocyanin content and the expression of BjMYB113 in Arabidopsis transgenic lines. **a** The total anthocyanin content in Arabidopsis transgenic lines. Col-0 is green wild-type Arabidopsis as the control, "BjMYB113a OX" refers to overexpression transgenic lines of BjMYB113a (CDS), "BjMYB113a COM" refers to complementary transgenic line of BjMYB113a (with native promotor and 3’region sequence). **b** The expression of BjMYB113 in the Arabidopsis transgenic lines. **c** The wild-type Arabidopsis Col-0. **d-f** BjMYB113a (CDS), BjMYB113b (CDS) and BjMYB113c (CDS) overexpression transgenic lines. **g-i** BjMYB113a, BjMYB113b and BjMYB113c complementary transgenic line. Scale bar represents 1 cm.
High expression of the BjMYB113a activated the anthocyanin biosynthetic genes

As an important component in the MBW transcription factor complex, activated BjMYB113 gene may up-regulate multiple genes in the anthocyanin biosynthesis pathway. qRT-PCR was carried out to investigate the expression difference of genes, including BjPAL (BjuA036480), BjC4H (BjuB015902), BjCHS (BjuA041225), BjCHI (BjuA004576), BjF3H (BjuA035478), BjF3’H (BjuA047311), BjDFR (BjuA033678), BjANS (BjuB044852), and BjUGFT (BjuA047199). The qRT-PCR results showed that the expression of BjCHS, BjF3H, BjF3’H, BjDFR, BjANS, and BjUGFT in purple cultivar pl102 were significantly higher than that in green cultivars rt104 and gre101 (Fig. 7).

The evolution of the BjMYB113 gene

To understand the evolution of the BjMYB113 gene in mustard, approximately 3 kb promoter sequences of 58 mustard accessions including 5 purple and 53 green ones were PCR amplified and sequenced. All accessions were also genotyped for the 1,268-bp insertion in intron 1 and the 1,033-bp insertion in its downstream. A neighbor-joining tree was constructed using the full-length gene sequences, with BoMYB113 from Brassica oleracea as an outgroup (Fig. 8). Two major clades were found for these gene sequences from mustard, varying at least 50 nucleotides in promoter sequences between members from the two clades. The six sequences from Clade 2 are identical, all with the inactivated promoter as in green cultivar rt101 (BjMYB113b). The Clade 1 contains five distinct sequences including the sequence of BjMYB113a, which has the activated promoter and lacks the 1,033-bp insertion in its downstream. The second sequence (with three cultivars) is BjMYB113c with one nucleotide different from BjMYB113a in promoter sequence and has the 1,033-bp insertion in its downstream. The third sequence (two cultivars) varies only two nucleotides from BjMYB113a and also has the 1,033-bp insertion in intron 1.

Fig. 7 The relative expression of nine anthocyanin biosynthesis genes in three Brassica juncea cultivars pl102, rt104 and gre101. Statistical analyzes were conducted by Student’s t test, *P < 0.05; **P < 0.01
033-bp insertion. The fourth sequence (five cultivars) has the same promoter sequence as BjMYB113a and with the 1,033-bp insertion. The fifth sequence (one cultivar LY) varies a 38-bp insertion in the promoter sequence and has the 1,268-bp insertion in intron 1. Therefore, all promoter sequences from Clade 1 are most likely activated ones (see above). Surprisingly, of the 52 cultivars in Clade 1, 36 have green leaves but they do not have the large insertions in either intron 1 or in its 3’ downstream. Three genotypes of 36 green cultivars were randomly chosen as representatives to investigate the expression level of BjMYB113. The expression level of BjMYB113 from three green cultivars were significantly lower than BjMYB113a from purple cultivar, suggesting that there might be other mutation events leading to the inactivation of the BjMYB113 gene or loss-of-function mutations in other genes of the anthocyanin biosynthesis pathway (Fig. S3).

Discussion

**MYB plays a central role in color variation in plant species**

In this study, we showed that the BjMYB113 gene was responsible for the color variation in mustard, and point mutations and small InDels in the promoter region accounted for its activation and accumulation of the anthocyanins. Color polymorphisms might be caused by any genes in the anthocyanin biosynthesis pathways. However, the most frequent causal genes for natural color variations were MYB-encoding genes. For example, the RLL2 gene which encodes an R2R3-MYB transcription factor, regulates multiple genes from the anthocyanin biosynthesis pathways and promote the high-level accumulation of anthocyanins in lettuce leaves [29]. In *B. oleracea* species, the BoMYB2 gene has been activated independently for at least three times [31, 32]. The blood orange arose by insertion of a Copia-like retrotransposon to a MYB gene Ruby [26]. The OsC1 mutations in the coding region vary degrees of apiculus coloration in rice [36].

While wild species in natural populations have green leaves, many horticultural cultivars have red/purple leaves, such as red leaf plum, purple lettuce and purple mustard. With the assumption that their wild ancestors had green leaves, the red or purple cultivars arose from mutations and artificially selected and maintained during domestication or breeding. We showed in this study that
the BjMYB113 gene was activated due to the accumulated point mutations and/or InDels. It will be interestingly to investigate which point mutation(s)/InDel(s) played the critical role in its activation.

In this study, we also found a mutation event that inactivated the BjMYB113 gene, a large insertion in its downstream. BLASTN search of the inserted sequences showed that it had a high copy number in the genome of mustard. The inserted sequences do not show similarity to known transposable elements, and the sequences per se do not show a typical transposon structure such as target site duplication (TSD) or terminal inverted repeats (TIR).

The mechanism for the inactivation of the BjMYB113 gene by the insertion remains unknown. One possibility is that the insertion may result in methylation in this region to downregulate its expression. We investigated cytosine methylation status in this region, but no significant difference in methylation ratio was found between BjMYB113a and BjMYB113c (Fig. S5). Another possibility is that the insertion disrupts an enhancer motif of the gene. It is also possible that the insertion might have altered the chromatin structure and consequently downregulates gene expression in its vicinity.

Our sequence analysis suggested that at least 36 cultivars with green leaves have the activated promoter of the BjMYB113a allele from the purple cultivar, but they do not have the 1,033-bp insertion in its 3’ downstream or 1,268-bp insertion in intron 1 that could silence the gene. We sequenced the entire BjMYB113 gene from some cultivars, but no sequences variations were identified. Therefore, these green phenotypes were most likely caused by loss-of-function mutations of other genes in the anthocyanin biosynthesis pathway rather than the BjMYB113 gene.

In a recent report, a 1,268-bp insertion in the first intron suppressed the BjMYB113 gene, leading to a low anthocyanin accumulation [3]. The only difference between the promoter sequences of the BjMYB113 gene from that cultivar LY and the BjMYB113a allele is a 38-bp insertion, and it is still unclear whether the low expression of BjMYB113 gene in the cultivar is due to the 38-bp insertion in its promoter or the insertion of the 1, 268-bp in its intron 1 (Fig. 4).

**Conclusions**
We genetically cloned an R2R3-MYB coding gene controlling purple leaves in mustard. BjMYB113 was activated due to promoter variations, leading to the high expression of genes in the anthocyanin biosynthesis pathway and the high accumulation of anthocyanin in the purple cultivar. However, a large insertion in its 3’region or in its first intron compromised the high expression in the active allele leading to green color. These results indicated that BjMYB113 transcription factor, as the important member of the MBW ternary complex, has experienced both gain-of-function and loss-of-function mutations during artificial selection and domestication. Our results contribute to a better understanding of the genetics and evolution of the BjMYB113 gene.
controlling purple leaves and provide useful information for further breeding programs of mustard, which will be of broad interest to biologists.

**Methods**

**Plant materials and genetic segregating populations**
Purple mustard cultivar pl102 and green cultivars rt104 and gre101 were chosen to study the genetics of purple color leaves. These three mustard cultivars are brown mustard (*Brassica juncea*). Purple cultivar pl102 was crossed with green cultivars rt104 and gre101, respectively, to develop two F2 segregating populations. These F2 segregating populations along with its parents were acquired from professor Hanhui Kuang’s lab (College of Horticulture and Forestry Sciences, Huazhong Agricultural University) in this study. Parental seed lines were originally acquired at National Center for Vegetable Improvement (Central China). Seeds were germinated and seedings grew in a greenhouse under 16/8 h photoperiod at 25 °C ± 2 °C. All plants were planted to the field on campus of Huazhong Agriculture University, Wuhan, China.

**Analysis of total anthocyanin contents**
Method of total anthocyanin content analysis followed the guidelines and protocols described previously [45]. For anthocyanin extraction, 0.1 g tissues of each sample were incubated in 300 µL of extraction buffer (methanol containing 1 % HCl) overnight at 4 °C in the dark. After extraction, 200 µL of water and 200 µL of chloroform were added to each sample, and absorbances were read at 530 and 657 nm. The quantity of anthocyanin was determined by calculating absorbance at 530 nm (A530) – at 530 and 657 nm. The quantity of anthocyanin was determined by calculating absorbance at 530 nm (A530) – at 530 and 657 nm.

**Flavonoid metabolite profiling**
The relative quantities of flavonoid metabolites in *Brassica juncea* leaf samples were analyzed with a liquid chromatography-electrospray ionization-tandem mass spectrometry (LC-ESI-MS/MS) system by Met-Ware (Wuhan, China). The liquid chromatography–electrospray ionization-tandem mass spectrometry (LC-ESI-MS/MS) system was used for the relative quantification of anthocyanin metabolites in *Brassica juncea* leaves samples. The protocols were described detailedly in Methods S1.

**Bulked Segregant Analysis and RNA-seq (BSR-seq)**
Bulked Segregant Analysis in combination with RNA-seq (BSR-seq) was used to map genes controlling purple leaves [46]. A total of 20 purple individuals from an F2 segregating population were mixed as the purple-pool, and 20 green individuals were mixed as the green-pool. Total RNAs were extracted from the two pools using RNAiso plus (Takara, Japan). RNA-seq was performed on Illumina Hiseq2500 platform (Novogene, China), and approximately 5 GB clean data were obtained for each pool. The data were mapped to *Brassica juncea* reference genome [47], using the Bowtie software [48]. SNP callings were performed using SAMTools [49]. Low-quality SNPs with map quality value < 30, reads depth < 10× or base quality value < 20 were excluded. The key parameter of Δ(SNP-index) was used to identify the target region for purple/green traits [50]. The Δ(SNP-index) was calculated by subtracting the SNP-index value of the green-pool from the SNP-index value of purple-pool. Cleaved amplified polymorphic sequence (CAPS) markers were designed in the candidate region, and were used to screen the population to fine map the casual gene. The primers used in map-based cloning were shown in Table S3. The RNA-seq data supporting the results of this study is available in the NCBI SRA (Sequence Read Archive, [http://www.ncbi.nlm.nih.gov/sra/](http://www.ncbi.nlm.nih.gov/sra/)) repository under BioProject PRJNA672814.

**RNA extraction and quantitative Real-Time reverse transcription polymerase chain reaction (qRT-PCR) analysis**
Total RNAs were extracted from leaves using RNAiso plus (Takara, Japan). The cDNA was synthesized using TransScript One-Step gDNA Removal and cDNA Synthesis SuperMix (TransScript, China) with Oligo-dT18 primer. qRT-PCR analysis followed the guidelines and protocols described previously [51]. All reactions were performed using the SYBR Premix (5.0 µL of 2x SYBR Premix Go Taq II, 0.5 µL of primers, 1.0 µL of cDNA, and 3.5 µL of ddH2O). Melting curve analysis of qRT-PCR samples revealed that there was only one product for each gene primer reaction. The PCR products were sequenced to confirm the specific amplification. A house-keeping gene *BjEF-1-a* was used as an internal standard in tissues. The relative expression levels were counted using the formula 2−△△Cq as described in Bio-Rad protocol, and statistical differences were calculated using student’s test. Three biological replications and three technical replications were performed in qRT-PCR. The primers used in qRT-PCR analysis were shown in Table S3.

**Plasmid construction and plant transformation**
The full-length *BjMYB113* cDNA was cloned into the pRI101-GFP vector with CaMV 35 S promoter to construct the overexpression vector. The full-length *BjMYB113* gene sequence (including native promoter and 3’region sequence) was amplified and cloned into the pRI101-GFP to construct the complementary vector. Approximately 3 kb promoter region of three *BjMYB113*
alleles were amplified and recombined into the pCAMBIA1301-GUS vector for checking GUS activity. All vectors were constructed using homologous recombination. Positive plasmids were verified by sequencing and then transformed into Agrobacterium tumefaciens GV3101 using thermal transformation method. The vectors were transformed into Arabidopsis using floral-dip method [52]. All primers used for vector construction were shown in Table S3.

Promoter activity analysis

The quantitative GUS activity was measured using the Lu’s methods with slight modification [53]. GUS activity was detected in 1-month-old Arabidopsis leaf tissues (10 mg) from three independent transgenic lines and six individuals in each line. Total proteins were extracted using 300 µL GUS extraction buffer (50 mM phosphate buffer, pH 7.0; 10 mM EDTA, pH 8.0; 0.1 % Sodium Dodecyl Sulfate; 10 mM β-mercaptoethanol). BCA Protein Assay (Beyotime Biotechnology, China) was used to measure the protein concentrations. Extraction (100 µL) was added to 900 µL GUS extraction buffer containing 1 mM 4-methylumbelliferyl glucuronide (MUG, Sigma) and incubated at 37 °C. The 900 µL stop solution (1 M Sodium Carbonate) immediately added into 100 µL the above reaction mixture and 60 min later, respectively. Fluorescence of 4-methylumbelliferone (MU) was monitored using Tecan Infinite™ at 455 nm emission and 365 nm excitation. GUS activity was expressed as µmoles 4-Dodecyl Sulfate; 10 mM β-buffer, pH 7.0; 10 mM EDTA, pH 8.0; 0.1 % Sodium Dodecyl Sulfate; 10 mM β-mercaptoethanol. GUS activity was expressed as µmoles 4-Dodecyl Sulfate; 10 mM β-buffer, pH 7.0; 10 mM EDTA, pH 8.0; 0.1 % Sodium Dodecyl Sulfate; 10 mM β-mercaptoethanol. GUS activity was expressed as µmoles 4-Dodecyl Sulfate; 10 mM β-buffer, pH 7.0; 10 mM EDTA, pH 8.0; 0.1 % Sodium Dodecyl Sulfate; 10 mM β-mercaptoethanol.

Sequence analysis and neighbor-joining tree

The promoter sequence of BjMYB113 alleles was analyzed using PlantCARE (http://bioinformatics.psb.ugent. be/webtools/plantcare/html/).

Sequence alignments were conducted using Muscle program and manually adjusted in GeneDoc (http://www.nrbsc.org/gfx/genedoc/). The neighbor-joining tree was constructed using MEGA 7.0 [54] and bootstrap values were calculated using 1,000 times.

The gene sequences (BjMYB113a, -b, and -c) supporting the results of this study are available in the NCBI GenBank (https://www.ncbi.nlm.nih.gov/genbank/) under accession number MW166171-MW166173 of Banklit2394807. The primers used in gene sequences analysis were shown in Table S3.

Methylation sensitive digestion

Method of methylation sensitive digestion followed the guidelines and protocols described previously [55]. Genomic DNA was extracted from purple mustard cultivar pl102 and green cultivar gre101. The quality and integrity of extracted genomic DNA was evaluated by spectrophotometric analysis using NanoDrop (ThermoFisher Scientific, U.S.A.). Genomic DNA was digested by the methylation-sensitive endonuclease McrBC according to the manufacturer’s instructions (New England Biolab Inc., U.S.A.). Then, qRT-PCR analysis was performed [51]. Each sample was measured in three independent experiments. The mean Ct values were used to calculate ΔC_T as follows: ΔC_T = [C_T (McrBC treatment) – C_T (Control)] and the methylation percentage was calculated as methylation% = 100 – (100 × 2^ΔC_T).

Abbreviations

ANS: Anthocyanidin synthase; BSR-seq: Bulked Segregant Analysis and RNA-seq; CaMV: Cauliflower mosaic virus; CAPS: Cleaved amplified polymorphic sequence; CHI: Chalcone isomerase; CHS: Chalcone synthase; DGEs: Differentially expressed genes; DFR: Dihydroflavonol 4-reductase; Flavanone 3-hydroxylase, GUS: Glucuronidase; InDel: Insertions and Deletions; LC-ESI-MS: Liquid Chromatography-Electrospray Ionization-Mass Spectrometry; MPL: Mustard Purple Leaves; MU: 4-methylumbelliferone; MUG: 4-methylumbelliferyl glucuronide; ORFs: Open reading frames; PAL: Phenylalanine-ammonia lyase; qRT-PCR: Quantitative Real-Time reverse transcription polymerase chain reaction; ROS: Reactive oxygen species; SNPs: Single nucleotide polymorphisms; TIR: Terminal inverted repeats; TSD: Target site duplication.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12870-021-03084-5.

Additional file 1: Figure S1. Map-based cloning of the MPL gene in F2 population of pl102 × gre101.

Additional file 2: Figure S2. The alignment of protein sequences of BjMYB113 and its most similar R2R3-MYB transcription factors homologous genes.

Additional file 3: Figure S3. The relative expression of BjMYB113 of purple cultivar pl102 and three green cultivars.

Additional file 4: Figure S4. Analysis of the methylation status in the 3’ regions of BjMYB113a and BjMYB113c.

Additional file 5: Table S1. Different types of anthocyanin and their concentrations in purple and green cultivars with three biological replications.

Additional file 6: Table S2. The Chi-square (χ2) of two purple/green segregating populations.

Additional file 7: Table S3. The primers used in this research.

Additional file 8: Method S1. Anthocyanin metabolite profiling.

Acknowledgements

We thank Dr. Changchun Yu (Huazhong Agricultural University) for the vector pRI101-GFP and pCAMBIA1301-GUS; Dr. Weiyi Zhang (Huazhong Agricultural University) for technical support of bioinformatics; Dr. Hanhui Kuang (Huazhong Agricultural University) for plant materials and helpful suggestions on this work.

Authors’ contributions

J.C. designed the project. G.A. performed the experiments. G.A. wrote the manuscript with the help from J.C. The author(s) read and approved the final manuscript.

Funding

This work was supported by grants from the National Key Research and Development Program of China [2018YFD0100080], and the Fundamental Research Funds for the Central Universities [2662018PY070]. All funders were not involved in the design of the study, data analysis, and writing the manuscript.
Availability of data and materials
The data supporting the result of this study are within the paper and its additional files. All sequencing datasets are deposited in the National Centre for Biotechnology Information (NCBI) under the BioProject ID PRJNA672814 with the Sequence Read Archive (SRA) accession SRR1219405-SRR1219408. The gene sequences (BMVY1113_g - b and - c) supporting the results of this study are available in the NCBI GenBank (https://www.ncbi.nlm.nih.gov/genbank/) under accession number MW166171 - MW166173 of Bankit2394807.

Declarations

Ethics approval and consent to participate
Not applicable.

Competing interests
The authors declare no conflict of interests.

Received: 12 January 2021 Accepted: 4 June 2021
Published online: 29 June 2021

References

1. Xie Q, Hu Z, Zhang Y, Tian S, Wang Z, Zhao Z, Yang Y, Chen G. Accumulation and molecular regulation of anthocyanin in purple tumorous stem mustard (Brassica juncea var. tumida Tsen et Lee). J Agraic Food Chem. 2014;62(31):7813–21.
2. Liu X, Lu Y, Yuan Y, Liu S, Guan C, Chen S, Liu Z. De novo transcriptome of Brassica juncea seed coat and identification of genes for the biosynthesis of flavonoids. PlOc One. 2013;8(8):e71110.
3. Heng S, Cheng Q, Zhang T, Liu X, Huang H, Yao P, Liu Z, Wan Z, Fu T. Fine-mapping of the B‘pur gene for purple leaf color in Brassica juncea. Theor Appl Genet. 2020. https://doi.org/10.1007/s00122-020-03604-0.
4. Kong J-M, Chia L-S, Goh N-K, Chia T-F, Brouillard R. Analysis and biological activities of anthocyanins. Phytochemistry. 2003;64(5):923–33.
5. Stuutman J, Hoballah ME, Broger L, Moore J, Basten C, Kuhlemeier C. Dissection of floral pollination syndromes in Petunia. Genetics. 2004;168(3):1585–99.
6. Neill SO, Gould KS. Anthocyanins in leaves: light attenuators or antioxidants? Funct Plant Biol. 2003;30(8):865–73.
7. Gould KS. Nature’s swiss army knife: the diverse protective roles of anthocyanins in leaves. J Biomed Biotechnol. 2004;2004(5):314–20.
8. Albert NW, Lewis DH, Zhang H, Irving LJ, Jameson PE, Davies KM. Light-induced vegetative anthocyanin pigmentation in Petunia. J Exp Bot. 2009;60(7):2101–202.
9. Olsen KM, Slimestad R, Lea US, Brede C, Lovdal T, Ruoff P, Verheul M, Lillo C. Temperature and nitrogen effects on regulators and products of the flavonoid pathway: experimental and kinetic model studies. Plant Cell Environ. 2009;32(3):286–99.
10. Zhang Y, Butelli E, De Stefano R, Schoonbeek H, Magusin A, Pagliarani C, Hellens RP, Putterill J, Stevenson DE, Kutty-Amma S, Allan AC. Red colouration in apple fruit is due to the activity of the MYB transcription factor MdMYB810. Plant J. 2007;49(3):414–27.
11. Li C, Wu J, Hu KD, Wei SW, Sun HY, Hu LY, Han Z, Yao GF, Zhang H. PyWRKY26 and PyBHLH3 cotarget the PyMYB114 promoter to regulate anthocyanin biosynthesis and transport in red-skinned pears. Hortic Res. 2020;7:37.
12. Li C, Weinberg M. Anthocyanins and their role in cancer prevention. Curr Opin Plant Biol. 2014;18(5):62–7.
13. Zhang Y, Butelli E, Martin C. Engineering anthocyanin biosynthesis in plants. Annu Rev Plant Biol. 2017;68(21):5745–80.
14. Espley RV, Hellens RP, Putterill J, Stevenson DE, Kutty-Amma S, Allan AC. Red colouration in apple fruit is due to the activity of the MYB transcription factor MdMYB810. Plant J. 2007;49(3):414–27.
15. Butelli E, Lobb A, Tornielli B, Tornielli M, Finezzo L, Begheldo M, Ruperti B, Tornielli L. Molecular analysis of the anthocyanin2 gene of petunia and its role in the evolution of flower color. Plant Cell. 1999;11(8):1433–44.
16. Sweeney MT, Thomson MJ, Pfeil BE, McCouch S. Caught red-handed: Ac encodes a basic helix-loop-helix protein conditioning red perciparc in rice. Plant Cell. 2006;18(2):283–94.
17. Su W, Liu Y, Wu C, Yue Z, He S, Lavelle D, Zhang W, Zhang L, An G, et al. Characterization of four polymorphic genes controlling red leaf colour in lettuce that have undergone disruptive selection since domestication. Biotechnol J. 2020;18(2):479–90.
18. Zhang Y, Chen G, Dong T, Pan Y, Zhao Z, Tian S, Hu Z. Anthocyanin accumulation and transcriptional regulation of anthocyanin biosynthesis in purple bok choy (Brassica rapa var. chinensis). J Agric Food Chem. 2014;62(51):12366–74.
19. Yan C, An G, Zhu T, Zhang W, Zhang L, Peng L, Chen J, Kuang H. Independent activation of the OmMYB2 gene leading to purple traits in Brassica oleracea. Theor Appl Genet. 2019;132(4):895–906.
20. Chiu LW, Zhou X, Burke S, Wu X, Prior RL, Li L. The purple cauliflower arises from activation of a MYB transcription factor. Plant Physiol. 2010;154(3):1470–80.
21. Li P, Chen B, Zhang G, Chen L, Dong Q, Wen J, Mysore KS, Zhao J. Regulation of anthocyanin and proanthocyanidin biosynthesis by Medicagro truncatula bHLH transcription factor MTT8. New Phytol. 2016;201(3):905–21.
22. Amato A, Cavallini E, Zenoni S, Finezzo L, Begheldo M, Ruperti B, Tonielli GB. A grapevine TTG2-like WRKY transcription factor is involved in regulating vacuolar transport and flavonoid biosynthesis. Front Plant Sci. 2017;8:71997.
23. Pent PK, Xu YJ, Yang X, Zhou X, Deng XB, Wu HH, Guo ZJ. Constitutive expression of rice WRKY30 gene increases the endogenous jasmonic acid accumulation, PR gene expression and resistance to fungal pathogens in rice. Plant Sci. 2012;216(3):1485–98.
24. Saitho K, Onishi K, Mikami I, Thidar K, Sano Y. Allelic diversification at the OsC1 gene for purple leaf color in rice. Genetics. 2004;168(2):997–1007.
25. Jung S, Venkatesh J, Kang MY, Kwon JK, Kang BC. A non-LTR retrotransposon activates anthocyanin biosynthesis by regulating a MYB transcription factor in Capsicum annuum. Plant Sci. 2019;287:110181.
38. Radicella JP, Brown D, Tolar LA, Chandler VL. Allelic diversity of the maize B regulatory gene: different leader and promoter sequences of two B alleles determine distinct tissue specificities of anthocyanin production. Genes Dev. 1992;6(11):2152–64.

39. Li Y, Xiong H, Guo H, Zhou C, Xie Y, Zhao L, Gu J, Zhao S, Ding Y, Liu L. Identification of the vernalization gene VRN-B1 responsible for heading date variation by QTL mapping using a RIL population in wheat. BMC Plant Biol. 2020;20(1):1331.

40. Li R, Hou Z, Gao L, Xiao D, Hou X, Zhang C, Yan J, Song L. Conjunctive analyses of BSA-Seq and BSR-Seq to reveal the molecular pathway of leafy head formation in Chinese cabbage. Plants (Basel). 2019;8(12):603.

41. Sun L, Zhang Y, Cui H, Zhang L, Sha T, Wang C, Fan C, Luan F, Wang X. Linkage mapping and comparative transcriptome analysis of firmness in watermelon (Citrullus lanatus). Front Plant Sci. 2020;11:831.

42. Gao J, Dai G, Zhou W, Liang H, Huang J, Qiang D, Chen W, Wu H, Yang X, Li D, et al. Mapping and identifying a candidate gene PmH1, a recessive gene regulating purple leaf in rice, by using bulked segregant and transcriptome analysis with next-generation sequencing. Int J Mol Sci. 2019;20(18):4335.

43. Zhao Z, Sheng X, Yu H, Wang J, Shen Y, Gu H. Identification of candidate genes involved in curd riceneys in cauliflower. Int J Mol Sci. 2020;21(6):1999.

44. Tang QW, Tian MY, An GH, Zhang WY, Chen JJ, Yan CH. Rapid identification of the purple stem (Ps) gene of Chinese kale (Brassica oleracea var. alboglabra) in a segregation distortion population by bulked segregant analysis and RNA sequencing. Mol Breed. 2017;37(12):153.

45. Liu S, Yeh CT, Tang HW, Nettleton D, Schnable PS. Gene mapping via bulked segregant RNA-Seq (BSR-Seq). PLoS One. 2012;7(5):e36406.

46. Yang J, Liu D, Wang X, Ji C, Cheng F, Liu B, Hu Z, Chen S, Pental D, Ju Y, et al. The genome sequence of allopolyploid Brassica juncea and analysis of differential homeolog gene expression influencing selection. Nat Genet. 2016;48(10):1225–32.

47. Langmead B, Trapnell C, Pop M, Salzberg SL. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol. 2009;10(3):R25.

48. Kumar S, Stecher G, Tamura K. MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. Mol Biol Evol. 2016;33(7):1870–7.

Publisher’s Note
Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:
• fast, convenient online submission
• thorough peer review by experienced researchers in your field
• rapid publication on acceptance
• support for research data, including large and complex data types
• gold Open Access which fosters wider collaboration and increased citations
• maximum visibility for your research: over 100M website visits per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions