Article

Molecular Characterization and Expression Analysis of MYB Transcription Factors Involved in the Glucosinolate Pathway in Chinese Cabbage (*Brassica rapa* ssp. *pekinensis*)

Shipra Kumari [1], **Jung Su Jo** [2], **Hyo Seon Choi** [1], **Jun Gu Lee** [2], **Soo In Lee** [1], **Mi-Jeong Jeong** [1] and **Jin A Kim** [1,*]

[1] Department of Agricultural Biotechnology, National Institute of Agricultural Science, Rural Development Administration, 370, Nongsaengmyeong-ro, Wansan-gu, Jeonju-si, Jeollabuk-do 54874, Korea; tunnee.sipra@gmail.com (S.K.); tjiey0323@naver.com (H.S.C.); silee@korea.kr (S.I.L.); center1097@korea.kr (M.-J.J.)

[2] Department of Horticulture, College of Agriculture & Life Sciences, Chonbuk National University, Jeonju 54896, Korea; jjs446@naver.com (J.S.J.); jungu@jbnu.ac.kr (J.G.L.)

* Correspondence: jakim72@korea.kr; Tel.: +82-63-238-4619

check for updates

Abstract: Chinese cabbage (*Brassica rapa*) is a perennial crucifer vegetable that has long been used for forage. Crucifers are rich sources of glucosinolates (GSLs), which are anti-carcinogenic in humans and involved in plant defense responses. Myeloblastosis (MYB) proteins are a large family of transcription factors (TFs) in plants and play major regulatory roles in many biological processes. We identified 14 functional R2R3-*MYB* genes involved in glucosinolate biosynthesis in *B. rapa* ssp. *pekinensis*. Bioinformatic analysis of their phylogeny, protein motifs, gene interaction network, and molecular characteristics showed that Chinese cabbage MYB genes are comparable to those of *Arabidopsis thaliana*. The expression levels of the 14 *BrMYB* genes under fluorescent lamp, blue, and red light were quantitated using qRT-PCR analysis. Almost all of the R2R3-*BrMYBs* were upregulated and expressed more under red light than under fluorescent lamp or blue light, except *BrMYB34s*. We also calculated the total GSLs under each light condition. The total GSL content was higher under red light than under fluorescent lamp or blue light. Furthermore, the individual glucosinolates, comprised of four aliphatic GSLs (progoitrin, sinigrin, gluconapin, and glucobrassicanapin) and one indolic GSL (glucobrassicin), were higher under red light than the other light conditions. The relationships between light quality and glucosinolate biosynthesis require further investigation.

Keywords: aliphatic glucosinolate; bioinformatics; indolic glucosinolate; gene expression; light quality

1. Introduction

The genus *Brassica* contains several important crops that are used for oil and as condiments, vegetables, and sources of vitamin C and dietary fiber [1]. Among *Brassica* species, *Brassica rapa* has many subspecies with marked morphological variation, including non-heading Chinese cabbage (*B. rapa* ssp. *chinensis*) "pak choi", Chinese cabbage (*B. rapa* ssp. *pekinensis*), and turnip (*B. rapa* ssp. *rapifera*). In Asia, Chinese cabbage is an important vegetable crop. It is used as an ingredient for various recipes as a healthy food containing secondary metabolites, such as glucosinolates (GSLs), flavonoids, and anthocyanins [2].

Glucosinolates are plant secondary metabolites that are derivatives of sugars and amino acids and are rich in sulfur. Some GSLs and their degradation products have anti-carcinogenic and anti-oxidative activities in humans [3], and impart special tastes and flavors to *Brassica* vegetables [4]. In *Arabidopsis thaliana*, a transcription factor related to GSL biosynthesis has been reported [5]. GSL biosynthetic pathways and GSL products are regulated by many genes, as well as conditions such as light quality, abiotic stress, and temperature [6].

Myeloblastosis (MYB) proteins function mainly as transcription factors (TFs). The first recognized *myb* gene was derived from avian myeloblastosis virus and is known as *v-myb* [7]. *myb* genes have been identified from diverse fungi, vertebrates, slime molds, and insects [8]. In plants, *myb* genes play roles in processes such as cell shape determination, organ development [9], hormone signal transduction [10], secondary metabolism [11], disease resistance, and abiotic stress tolerance [12]. Comparative analysis with *Arabidopsis* showed that *B. rapa* contained approximately twice as many *MYB* genes, likely resulting from genome triplication via genome evolution and polyploidy [13]. Wang et al. have reported the analysis and annotation of a draft genome sequence of *B. rapa* [14]. In plant genomes, about 7% of all coding sequences are TFs, reflecting the complexity of plant transcriptional regulation [15]. TFs are proteins that act together with other transcription regulators to employ or obstruct RNA polymerases acting on the genome [15].

Most knowledge of plant *MYB* genes comes from *Arabidopsis thaliana* [16]. Recently, a genomic investigation examined *MYB* genes in Chinese cabbage and the response of R2R3-type MYBs to hormone treatments and abiotic stress [17]. As Chinese cabbage has a large (485 Mb) and complex genome, not all of its *MYB* genes have been characterized. In this work, we systematically characterized in silico 14 *MYB* genes linked to GSL biosynthesis in *B. rapa*. We investigated the effects of fluorescent lamp, blue, and red light on the production of GSLs in *B. rapa*.

2. Materials and Methods

2.1. Identification of BrMYB Genes

Using the keyword "MYB", the *Brassica* database (BRAD) [18] was searched using SWISS-PROT and TrEMBL. The predicted coding and protein sequences of the *MYB* genes were obtained from the *Brassica* database (ver. 1.5; tar.gz) [18]. Candidate *MYB* genes encoding full-length protein sequences were analyzed using the program Pfam (http://pfam.sanger.ac.uk/) [19]. The MYB domains of the selected *MYB* genes were confirmed using the Simple Modular Architecture Research (SMART)-EMBL web tool (http://smart.embl.de/smart/set_mode.cgi). The primary structures of the *MYB* genes were analyzed using ProtParam (http://expasy.org/tools/protparam.html).

2.2. Sequence Analysis of BrMYB Genes

The predicted protein sequences of selected candidate *MYB* genes from Chinese cabbage were aligned using ClustalX2 with the default parameters [20]. A phylogenetic tree of the 14 R2R3MYB proteins was generated by MEGA (ver. 7.0) (http://www.megasoftware.net/; [21]) using the neighbor-joining method. Multiple EM for Motif Elicitation (MEME software, ver. 4.10.1; [22]) was used to search for motifs in the MYB protein sequences. The MEME search parameters were as follows: (a) optimum motif width 6 to 65, and (b) maximum number of motifs to find 20 to 65. Manual inspection was performed such that the putative MYB predicted protein sequences contained conserved Trp residues. Chinese cabbage *MYB* genes were compared using NCBI-BLAST analysis (http://www.ncbi.nlm.nih.gov/BLAST/) to identify candidate genes with > 80% similarity to those of *A. thaliana*.

2.3. Growth Condition and Preparation of Plant Materials

Chinese cabbage was grown in soil culture medium under a 16/8 h light/dark cycle in a growth chamber (Hanbaek culture chamber, HB-302S-2, Hanbaek Scientific Co., Bucheon, Korea) for 2 weeks,

under fluorescent lamp (FL; visible ray (400–750 nm), OSRAM FPL36EX-D, OsRAM Korea, Seoul, Korea), blue (B; 455 nm), or red (R; 625 nm) light (150 μ mole m^{-2} s^{-1}), in the Department of Agricultural Biotechnology, National Academy of Agricultural Science, RDA, South Korea [23]. Seedlings were collected and immediately frozen in liquid nitrogen, and then stored at −80 °C. For RNA extraction, three independent seedlings were collected under each light condition. For GSLs, bulk seedlings dried were divided to 3 group for biological replication. Total RNA was isolated with the RNeasy Plant Mini Kit (QIAGEN, Valencia, CA, USA), following the manufacturer's instructions. The quantity and quality of RNA was checked with a 2100 Bioanalyzer (Agilent, Santa Clara, CA, USA).

2.4. Quantitative Real-Time PCR (qRT-PCR) Expression Analysis in B. Rapa

qRT-PCR was conducted with 1 ng of cDNA in a 20-μL reaction volume using iTaq™ SYBR® Green supermix along with ROX dye (Bio-Rad, Hercules, CA, USA). Table S1 lists the gene-specific primers used for qRT-PCR. The qRT-PCR cycles consisted of 95 °C for 10 min, followed by 40 cycles of 95 °C for 20 seconds, 58 °C for 20 s, and 72 °C for 25 seconds. The fluorescence was examined after the last step of every cycle. Amplification, data processing, and detection were performed using the CFX96 Real-Time PCR Detection System (Bio-Rad). Detected quantification cycle (cq) values were examined using the $2^{-\Delta\Delta CT}$ method to reveal changes in gene expression [24]. For validation of MYBs isolated from RNAseq data set, three-week-old plants grown under a 12/12 h light/dark cycle, following growth under a 16/8 h light/dark cycle for 2 weeks, were used as input samples for qRT-PCR (quantitative real-time PCR). Plant tissues were collected at 4 h intervals, for a 24 h period under light/dark conditions with 3 technical replications. For *MYB* expression tests response to light conditions, three independent seedlings were collected under each light condition after being grown by methods indicated in Section 2.3.

2.5. Extraction of Desulfo-Glucosinolates (DS-GSLs) and UPLC Analysis

Sample preparation and GSL analyses were performed according to modified version of the method described by Bhandari et al. [25]. Briefly, freeze-dried powder samples (100 mg) of whole aerial seedling were mixed with 2 mL of boiling methanol (70%) for 20 min and centrifuged at 12,000 rpm for 10 min at 4 °C. Thereafter, the pellet was re-extracted once more time under the same condition and the supernatants were combined. Desulfo-glucosinolates were then prepared using purified sulfatase isolated from *Helix pomatia* and then quantitatively determined by high performance liquid chromatography (HPLC). The crude GSL were loaded onto a Mini Bio-Spin chromatography column (Bio-Rad Laboratories, Hercules, CA, USA) containing 0.5 mL of DEAE-Sephadex A 25 anion exchange resin, which was pre-activated with 0.1 M sodium acetate (pH 4.0) Next, desulfation was conducted by adding 200 μL purified aryl sulfatase (EC 3.1.6.1, type H-1 from *Helix pomatia*; Sigma-Aldrich Co., St. Louis, MO, USA). The column was capped and allowed to stand at room temperature for 24 h, and the desulfo-GSLs were eluted with 1.5 mL distilled water, filtered through a 0.2-μm syringe filter, injected into H-Class UPLC system (H-Class, Waters Co., Milford, MA, USA) using an Acquity UPLC® BEH-C18 column (1.7 μm, 2.1 × 100 mm; Waters Co.), and its absorbance spectrum was measured at 229 nm with a photodiode array (PDA) detector. Solvent A (100% distilled water) and solvent B (20% acetonitrile in water) were used for the elution of compounds at the flow rate of 0.2 mL·min^{-1}. The gradient programs were as follows: a linear step from 1% to 99% of solvent B within 6 min, followed by constant conditions for up to 10 min, and then a quick dropdown to 1% of solvent B at 12 min, and isocratic conditions of 1% of solvent B up to 18 min. Authentic standards of GSLs were desulfated and used for the identification and quantification of the peaks. Concentrations of individual desulfo-GSLs were determined from the experimental peak area by analytical interpolation in a standard calibration curve of each desulfo-GSL across different ranges depending upon the GSLs and were expressed as micromoles per gram (μmol·g^{-1}) of dry weight.

The glucosinolates were isolated and quantified by Ultra-Performance Liquid Chromatography (UPLC)based on comparison of the retention times, response factors, and peak areas with those of

external standard compounds. The GSLs include six aliphatic glucosinolates (AGs; progoitrin, epiprogoitrin, sinigrin, glucoraphanin, gluconapin, and glucobrassicanapin) and one indolic glucosinolate (IG; glucobrassicin).

2.6. Statistical Analysis

Statistical analysis was performed for gene expression and glucosinolate contents under different light conditions using GraphPad ver. 5.0 software (California, USA). Significant differences were calculated using two-way analysis of variance (ANOVA) follow by Bonferroni's multiple comparison tests at $p < 0.05$. The expression of genes and glucosinolate contents were compared under fluorescent lamp, blue, and red light by computing the standard error of mean.

3. Results

3.1. Identification of the MYB Genes in Chinese Cabbage

Kim et al. performed a transcriptome-wide diurnal gene expression analysis in Chinese cabbage [26]. The expression profiles of 323 BrMYB genes were measured six times at 4 h intervals over 24 h (12/12 h light/dark cycle) and expressed as higher and lower fold change values. From the 323 BrMYB proteins, we selected 14 R2R3-MYB proteins involved in GSL biosynthesis for further analysis. A heatmap revealed that the expression levels of these genes differed at different time points (Figure 1). Most MYB members showed changed expression levels at each time point. Some, such as *BrMYB28-1*, *BrMYB29*, *BrMYB29-1*, *BrMYB51-1*, *BrMYB51-3*, *BrMYB34-3*, *BrMYB34-4*, and *BrMYB122-2*, showed a diurnal rhythm with a peak or trough. These genes might be associated with diurnal function or circadian rhythm.
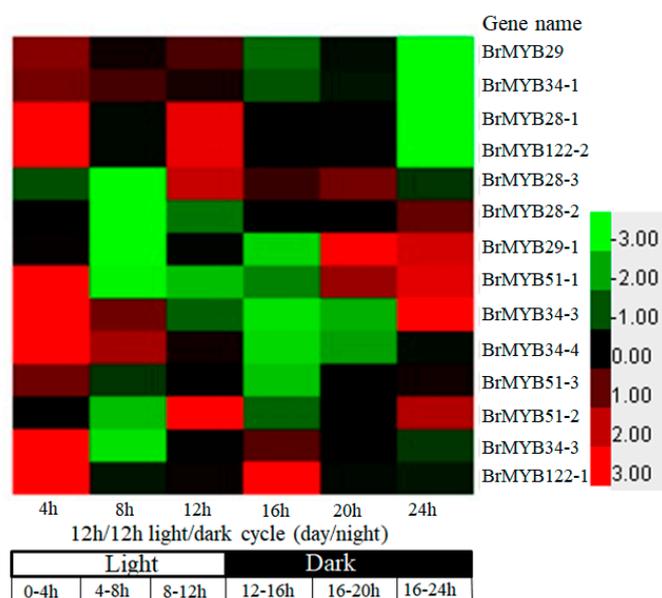
Figure 1. Expression profiles of *B. rapa* R2R3-BrMYB transcription factors (TFs) over time. The heatmap shows the expression patterns of the 14 BrMYB TFs involved in the *B. rapa* glucosinolate pathway, based on transcriptome data. Differentially Expressed Genes (DEGs) values were calculated at 4, 8, 12, 16, 20, and 24 h versus 0 h (light-on time) under a 12/12 h light/dark cycle (bottom). Names of the *B. rapa* genes are shown at right. Color bar represents expression: red indicates upregulation and green indicates downregulation (right).

Using the MYB DNA-binding domain (DBD) model sequence (PF00249) as a query, we searched for MYB domains. An analysis was performed to identify MYB protein sequences in our RNA-Seq

dataset. The TFs were annotated functionally based on domain similarities and sequence homology. The 14 *BrMYB* genes are described in Table 1.

Table 1. Sequence analyses of the BrMYB TFs linked to glucosinolates (GSLs) biosynthesis in *B. rapa*.

Gene ID	Gene Name	Chr	Length (aa)	Length (bp)	pI/Mw	No. of MYB Domain	*A. thaliana* Gene ID & Name	% Identity with *A. thaliana*
Bra012961	*BrMYB28-1*	A03	354	1065	5.69/39870.09	2	AT5G61420 (AtMYB28)	84.1–85.4
Bra035929	*BrMYB28-2*	A09	357	1074	5.59/40509.60	2		
Bra029311	*BrMYB28-3*	A02	372	1119	6.08/42063.90	2		
Bra005949	*BrMYB29*	A03	330	993	4.78/37140.91	2	AT5G07690 (AtMYB29)	83.2–87.0
Bra009245	*BrMYB29-1*	A10	88	267	9.30/10148.72	1		
Bra013000	*BrMYB34-1*	A03	316	951	4.99/34916.87	2	AT5G60890 (AtMYB34)	82.1–94.3
Bra035954	*BrMYB34-2*	A09	302	909	5.45/33544.80	2		
Bra029350	*BrMYB34-3*	A02	309	930	5.38/34468.68	2		
Bra029349	*BrMYB34-4*	A02	280	843	5.12/31179.99	2		
Bra025666	*BrMYB51-1*	A06	341	1026	5.96/38407.18	2	AT1G18570 (AtMYB51)	81.0–89.6
Bra031035	*BrMYB51-2*	A09	320	963	5.60/36371.6	2		
Bra016553	*BrMYB51-3*	A08	333	1002	5.12/37303.49	2		
Bra015939	*BrMYB122-1*	A07	326	981	5.77/6555.59	2	AT1G74080 (AtMYB122)	82.4
Bra008131	*BrMYB122-2*	A02	334	1005	5.76/7635.81	2		

3.2. Multiple Sequence Alignment, Phylogenetic, Motif and Physiochemical Property Analyses of BrMYBs

We performed multiple sequence alignment using five amino acid sequences of MYBs from *Arabidopsis thaliana* (AtMYB28 AtMYB29, AtMYB34, AtMYB51, and AtMYB122) and 14 R2R3-MYB repeats to explore the homologous domains and sequence features of Chinese cabbage MYBs. Conserved tryptophan (Trp; W) residues were identified in the R2R3 MYB encoded proteins. All MYBs contained motifs 1, 2, and 3, which are MYB repeats. In Chinese cabbage, each DBD repeat contained highly conserved tryptophan residues at positions 5 and 25 of the R2 repeat (Figure S1A) and at positions 10, 42, and 61 in the R3 repeat (Figure S1A). The first tryptophan (W10) residue belonged to the R2 repeat in the R3 repeat. The tryptophan residues are highly conserved. Alternative residues E-9, D-10, L-13, and G-21 in the R2 repeats and G-21, E-27, G-39, N-40, and R-52 in the R3 repeats (Figure S1A) were conserved in the R2R3 domains of Chinese cabbage. The conserved motifs of the R2R3-MYB proteins from *B. rapa* were analyzed. According to the sequence logos in Figure S1A, the 14 R2R3-BrMYB proteins possess most of the conserved motifs, including motifs 2, 3, and 1. In plants, the MYB domains are located at the N- and C-terminal ends, and in the middle of the protein.

Amino acid (aa) composition analysis revealed that BrMYB TFs are rich in Leu, Glu, Asn, and Ser. All sequences had two helix-turn-helix (HTH) structures, representing the MYB domain. Alignment analysis showed three tryptophans (Trp) at positions 6, 26, 46, 78, and 97 in the R2 and R3 repeats, characteristic of the MYB DBD (Figure S1C). Most of these Trps were conserved in the BrMYBs. In addition, 90% of the residues in the BrMYBs were highly conserved in MYB-DNA binding domain of R2R3 repeat. The LRPD element (Leu50, Arg51, Pro52, and Asp53 in Figure S1C) at the end of the R2 repeat was also highly conserved. In addition, C-terminal area showed variation in amino acid sequences (Figure S1B). These insertions and deletions indicate the structural divergence of the fourteen MYB proteins.

The physiochemical characteristics of the 14 BrMYBs were predicted based on sequence similarity analysis (Table 1). BrMYB proteins related to GSL biosynthesis had high sequence identity with *Arabidopsis* MYB proteins. The BrMYBs involved in GSL biosynthesis have varying lengths; R2R3-BrMYB29-1 was the smallest at 88 amino acids, whereas R2R3-BrMYB28-3 was the longest at 372 amino acids. We found 13 copies of completely coded orthologous sequences and one partial sequence corresponding to five *A. thaliana* MYB TFs. Only three predicted MYB28 (BrMYB28-1, BrMYB28-2, and BrMYB28-3) TFs identified in *B. rapa* showed greater than 84% sequence similarity with *Arabidopsis* MYBs. Two MYB29 TFs (BrMYB29, and BrMYB29-1) were found in *B. rapa*; they exhibit 83%–87% sequence identity with *Arabidopsis* MYBs. The predicted proteins encoded by four copies of MYB34 had more than 90% sequence similarity with MYBs from *Arabidopsis*. Three BrMYB51 TFs (BrMYB51-1,

BrMYB51-2, and BrMYB51-3) had 80%–89% sequence identity with *Arabidopsis* MYBs. Two MYB122 TFs had 81% and 82% sequence identity with *Arabidopsis* MYBs. The 14 orthologous and paralogous BrMYB TFs have differing overall lengths and sequences. Only one R2R3-BrMYB had a pI greater than 7; the others were acidic. Overall, the pI values of different BrMYBs ranged from basic to acidic. The varying MW and pI range of plant proteins will be useful to determine their biochemical and functional features.

A phylogenetic tree was constructed for the 14 R2R3-BrMYB proteins related to GSL biosynthesis and five related proteins from *Arabidopsis* (Figure S2). The 19 MYBs were classified into two groups: Group 1 contains aliphatic MYB proteins, whereas Group 2 contains indolic MYB proteins with orthologous MYBs in *Arabidopsis*. This analysis revealed the close relationship of orthologous MYBs of *Arabidopsis* and *B. rapa*.

3.3. Expression Pattern of BrMYB Genes

The RNA-Seq data were validated using quantitative real-time PCR (qRT-PCR) performed with 14 selected *BrMYB* genes at varying time points. Expression levels of genes were calculated at 0, 4, 8, 12, 16, 20, and 24 h, where light-on time was 0 h and light-off time was 12 h under a 12 /12 h light/dark cycle. The expression of *BrMYB28-1* was maximal at 4 h after light irradiation began, remained nearly constant until 4 h after light exposure was stopped, and then decreased sharply. The two paralogues of *BrMYB28-1* (*BrMYB28-2* and *BrMYB28-3*) were expressed at much lower levels than *BrMYB28-1* (Figure 2A). *BrMYB29* expression increased slowly until 8 h, increased rapidly until 16 h, and then decreased sharply during the night, whereas *BrMYB29-1* showed low expression throughout (Figure 2B). The expression of *BrMYB34-2* and *BrMYB34-3* increased from the initiation of light irradiation until 4 h later and then decreased sharply, whereas *BrMYB34-1* showed a diurnal rhythm during 1 day (Figure 2C). The expressions of all three *BrMTB51s* showed similar levels throughout the day, with only *BrMTB51-1* spiking in the middle of the night. In two *BrMYB122s*, *BrMYB122-2* expression was higher than that of the other member (Figure 2D,E).
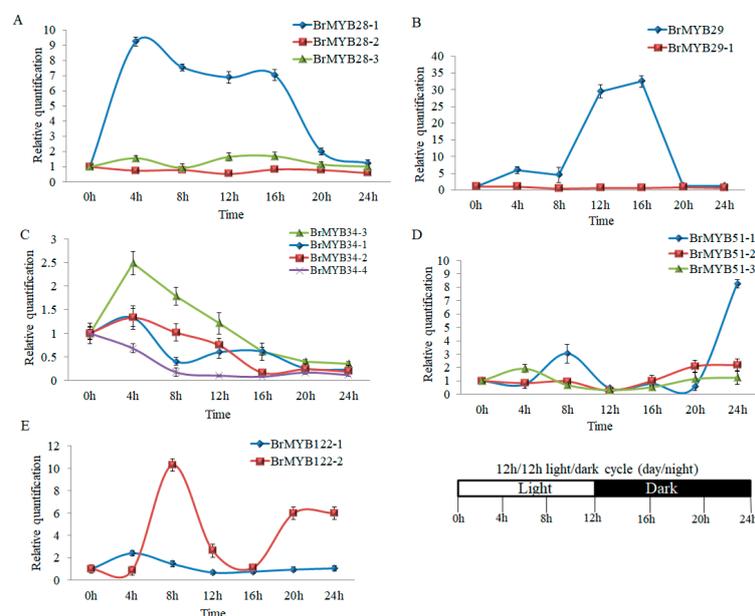
Figure 2. Expression patterns of *BrMYB* genes involved in the GSL pathway. The expression patterns of the (**A**) *BrMYB28-1*, *BrMYB28-2*, and *BrMYB28-3*, (**B**) *BrMYB29* and *BrMYB29-1*, (**C**) *BrMYB34-1*, *BrMYB34-2*, *BrMYB34-3*, and *BrMYB34-4*, (**D**) *BrMYB51-1*, *BrMYB51-2*, and *BrMYB51-3*, (**E**) *BrMYB122-1* and *BrMYB122-2* genes are shown. Under a 12 /12 h light-dark period, seedlings were harvested at seven-time points at 4 h intervals from the light-on time (0 h). Using the same RNAs which were used in previous RNA-Seq data set, qRT-PCR were carried out with 3 technical replicates.

### 3.4. Functional Network Analysis of Candidate BrMYB in Glucosinolate Pathway

Network analyses were performed by submitting the candidate *BrMYBs* with reported interactors into the Search Tool for the Retrieval of Interacting Genes/Proteins (STRING) database [27]. The known GSL biosynthetic genes were found to be highly co-expressed with MYB28, MYB29, MYB34, MYB51, and MYB122 (interconnected with thick lines). All of the proteins were in two major sub-clusters: Clusters I and II. Cluster I consist of all candidate interactors detected in the GSL pathway. The detected interactors of BrMYBs formed a network with 12 nodes (green and magenta spheres) and 29 edges (thick, thin, and broken lines; six edges had expected and interaction enrichment *p*-values of 2.07×10^{-13}) (Figure 3). Cluster II contained proteins that were not characterized in *B. rapa* and were co-expressed in this cluster. Network analyses were performed to identify the interactions of BrMYB proteins with other GSL biosynthesis enzymes, using BrMYB28, BrMYB29, BrMYB34, BrMYB51, and BrMYB122 as bait. In Cluster I, BrMYB28 and BrMYB29 were connected directly with other GSL biosynthesis proteins, such as MAM1, SOT18, CYP83A1, and CYP83A1, which belong to the aliphatic GSL pathway. The BrMYB34, BrMYB51, and BrMYB122 proteins were connected with SOT16, CYP83B1, and UGT74B1-1, which are related to indolic GSL biosynthesis.
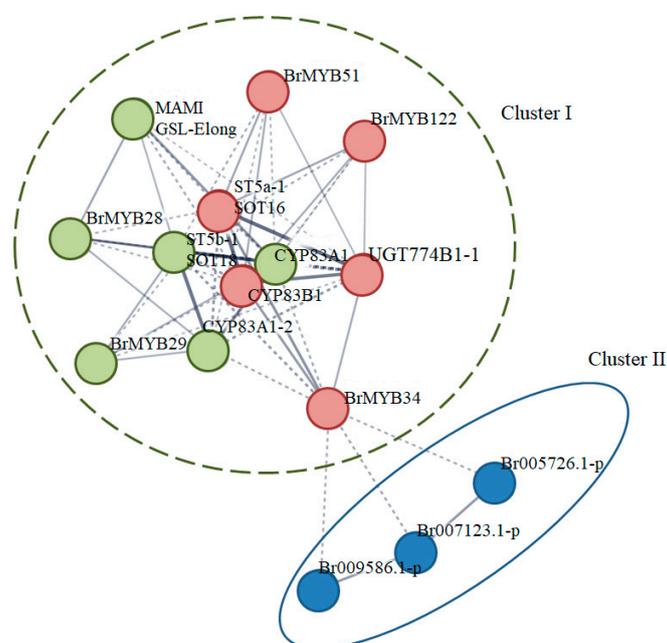
Figure 3. Interaction network analysis of BrMYB proteins in the GSL pathway. Network analysis was used to identify interactions of BrMYB proteins with other glucosinolate biosynthetic enzymes, using BrMYB28, BrMYB29, BrMYB34, BrMYB51, and BrMYB122 as bait. The network was constructed using the Search Tool for the Retrieval of Interacting Genes/Proteins (STRING) database [27]. In Cluster I, green spheres represent aliphatic GSL genes and magenta spheres represent indolic GSL genes. In Cluster II, blue spheres represent uncharacterized genes.

3.5. Expression Analysis of BrMYB Genes under Different Light Qualities

The expression of the 14 *BrMYBs* was analyzed by qRT-PCR under different light conditions. *BrMYB28-1*, *BrMYB28-2*, *BrMYB28-3*, *BrMYB29*, *BrMYB29-1*, *BrMYB51-1*, *BrMYB51-2*, *BrMYB51-3*, *BrMYB122-1*, and *BrMYB122-2* were expressed more strongly under red light than under fluorescent lamp or blue light (Figure 4A–E). The *BrMYB34s* genes (*BrMYB34-1*, *BrMYB34-2*, *BrMYB34-3* and *BrMYB34-4*) showed lower expression under red light than under fluorescent lamp (Figure 4C). Four *BrMYBs* (*BrMYB28-3*, *BrMYB29-1*, *BrMYB51-1*, and *BrMYB122-2*) showed the highest expression under red light. Only *BrMYB34-1* was highly expressed under blue light, and the expression of other *BrMYB34*s was low. Thus, different light qualities affected the expression of most of the *BrMYBs*.
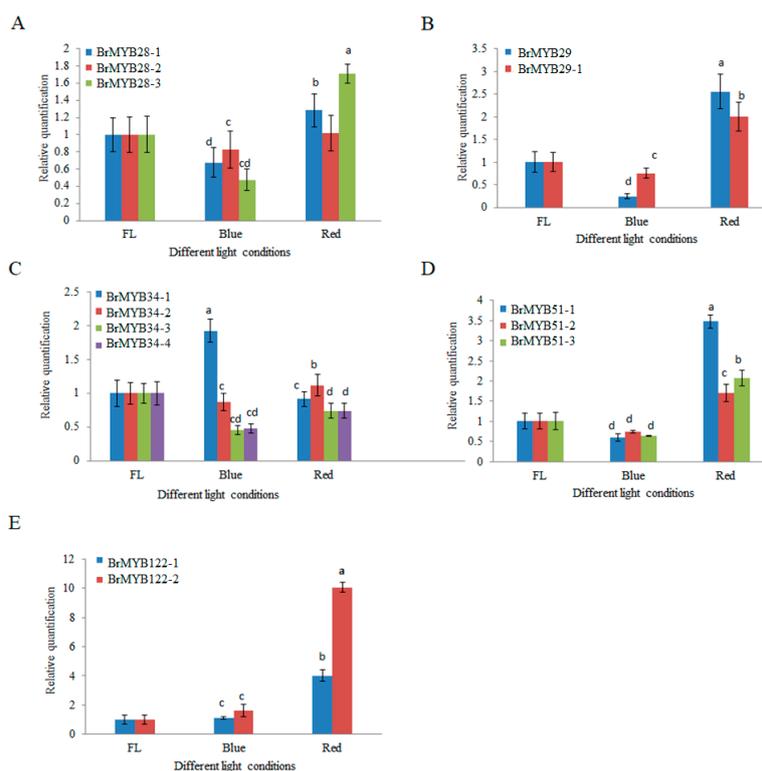
Figure 4. Relative expression levels of candidate *BrMYB* genes exposed to fluorescent lamp (FL), blue (B), or red (R) light. Expression of selected *BrMYBs* in Chinese cabbage is shown with significant differences. Expression profiles of (**A**) *BrMYB28-1*, *BrMYB28-2* and *BrMYB28-3*, (**B**) *BrMYB29* and *BrMYB29-1*, (**C**) *BrMYB34-1*, *BrMYB34-2*, *BrMYB34-3* and *BrMYB34-4*, (**D**) *BrMYB51-1*, *BrMYB51-2*, and *BrMYB51-3* and (**E**) *BrMYB122-1* and *BrMYB122-2* are shown. Error bars indicate the standard error of the mean of 3 biological replicates, and a, b, c, d and cd indicate significant differences ($p < 0.05$).

3.6. Production of Glucosinolates (GSLs) in Different Light Conditions

The GSL contents of *B. rapa* seedlings grown in growth chambers under fluorescent lamp, blue, and red light were analyzed. The plants contained total GSLs at 27.22 ± 0.91, 20.87 ± 1.05 and 31.22 ± 1.85 μmol/g dry wt. of the seedlings after exposure to fluorescent lamp, blue, and red light, respectively (Table 2). The total GSLs were highest under red light. In particular, the glucobrassicin level was two-fold under red light than in the other two light conditions.

Table 2. Total glucosinolates (μmol/g dry wt.) in *B. rapa* seedlings grown under florescence lamp (FL), blue (B), and red (R) light.

Glucosinolate	FL	B	R
Aliphatic glucosinolates			
Progoitrin	4.40 ± 0.80 [zb]	2.65 ± 0.10 [cy]	5.22 ± 0.62 [a]
Epiprogoitrin	0.14 ± 0.01 [a]	0.08 ± 0.0 [c]	0.14 ± 0.02 [a]
Sinigrin	0.00 ± 0.00 [c]	0.12 ± 0.01 [b]	0.13 ± 0.04 [a]
Glucoraphanin	0.02 ± 0.00 [a]	0.00 ± 0.00 [c]	0.01 ± 0.00 [b]
Gluconapin	13.22 ± 0.38 [b]	10.28 ± 0.86 [c]	14.26 ± 0.28 [a]
Glucobrassicanapin	8.11 ± 0.32 [b]	5.37 ± 0.51 [c]	8.22 ± 0.97 [a]
Indolic glucosinolates			
Glucobrassicin	1.34 ± 0.12 [c]	1.62 ± 0.40 [b]	3.24 ± 0.65 [a]
Total GSL	27.22 ± 0.91 [b]	20.87 ± 1.05 [c]	31.21 ± 1.85 [a]

[z] Values are mean ± SD of three replications on a dry weight basis; [y] Different letters in the same row indicate a significant difference ($p < 0.05$).

4. Discussion

MYB TFs regulate GSL biosynthesis in *Brassica*. We have published a transcriptome dataset expressed diurnally in the Chinese cabbage *B. rapa* L. ssp. *pekinensis* inbred line 'Chiifu' [26]. After annotation and analysis of the transcriptome data, we found 323 BrMYB family genes with increased up- and downregulation based during the daily cycle. We identified 14 BrMYBs associated with GSL biosynthesis (Figure 1). All 14 MYB TFs were found to be present in *B. rapa*, *B. juncea*, and *B. oleracea* [28]. Bioinformatics analysis showed that the MYB protein domain is characterized by a highly conserved N-terminal DNA-binding domain (DNA-DBD). The DNA-DBD comprises up to four repeats of an amino acid sequence (R). Each repeat contains around 52 amino acid residues and forms three α-helices. The second and third helices of every repeat have an HTH structure, with three regularly spaced hydrophobic tryptophans forming a hydrophobic core in the three-dimensional (3D) HTH structure [29]. An analysis of amino acid sequence alignment of the 14 BrMYB proteins showed mostly conserved in N-terminal region. In contrast, the amino acid sequence has variations (insertion/deletion) in the C-terminal area (Figure 2B). The C-terminal area functions as an activation domain and varies among MYB proteins that have a broad range of regulatory roles [30,31]. MYB proteins are generally divided into four major groups depending on the number and position of the repeats: 1R-, 2R-, 3R-, and 4R-MYBs [32,33]. All four groups are found in plants and represent the highest variation in MYB proteins. We found 49 alternative splicing among 323 *BrMYBs* based on the report of Tong et al. [34]. Out of 323 *BrMYBs*, four alternative splicing (*BrMYB28-2*, *BrMYB29*, *BrMYB34-1*, and *BrMYB122-1*) have been found in 14 *BrMYBs* related to the GSL biosynthesis pathway. Alternative splicing removes functional domains making transcripts non-functional, thus regulate gene dosage [35]. Alternative splicing also causes sub-functionalization of paralogous genes during whole genome duplication [36]. Gen duplication leads to functional diversity as pseudogenes or novel functionality by polyploidy in plant [37]. *BrMYB28-3* in *B. rapa* has been reported to be non-functional either as pseudogene or due to epigenetic silencing during polyploidy [38]. In our analysis, the duplicated gene *BrMYB28-3* has shown expression in red light (Figure 4A) in spite of insertion in its C-terminal and it might have been controlled by other factors to enhance secondary metabolite production, which is a matter for further study.

Our analysis showed that the 14 MYBs associated with GSL biosynthesis are 2R-MYBs: the R2R3-type. R2R3-MYB TFs also regulate GSL biosynthesis in *Arabidopsis* [39]. Network analysis showed that all 14 MYB TFs identified in *B. rapa* likely regulate aliphatic and indolic GSL biosynthesis (Figure 3).

Light intensity and spectral quality both affect plant life processes, including photosynthesis, differentiation, and flowering [40]. Photoreceptors are found at the roots of complex signaling networks, and control developmental, physiological, and morphological processes. GSL biosynthesis is also controlled by light. In the Brassicaceae, the quantity and composition of GSLs differ across species and among plant organs [41,42]. The metabolic fluctuations in cabbage seedlings are affected by light and temperature [43]. GSL components are regulated and altered by light [44,45]. The long hypocotyl5 (HY5) transcription regulator partly controls the light regulation of GSL biosynthesis genes in *Arabidopsis* [46].

The different GSLs produced in Chinese cabbage seedlings cultivated under different light qualities and their distributions were investigated. Seven components, including aliphatic and indolic GSLs, were identified after comparison with standard compounds. Under fluorescent, blue, and red light, the GSL contents in Chinese cabbage varied significantly. The total GSL contents were lower under fluorescent and blue light than under red light. Gluconapin was produced in large quantities under fluorescent and red light (Table 2). The total GSL produced under red light was approximately 31.21 μmol/g dry wt., compared with 27.22 μmol/g dry wt. for fluorescent light and 20.87 μmol/g dry wt. for blue light. Moon et al. [23] have reported that the expression of some genes (*CYP79F1*, *ST5a*, and *FMOGS-OX1* orthlogs) increased in response to blue light, while others (*MAM1*, *AOP3*, *UGT74B1*, and *BCAT4* orthlogs) responded to red light. In another report, the indolic GSL was greater after

irradiation with by red or green alone rather than other light conditions, including blue [45]. Our data showed red or blue monochromatic light condition could affect glucosinolate contents and some *BrMYB* members (*BrMYB29s*, *BrMYB51s*, and *BrMYB122s*). In particular, the expression levels of two member (*MYB51s* and *MYB122s*) related to the indolic GSL pathway, increased from 3.5- (*BrMYB51-1*) to 10-fold (*BrMYB122-2*) higher under red light conditions than under fluorescent lamp.

Glucobrassicanapin has been reported to be the predominant GSL in Chinese cabbage [47,48]. However, gluconapin, an aliphatic GSL, was produced in large quantities in *B. rapa* exposed to fluorescent lamp, blue, and red light in our study. Under red and far-red light, small RNAs have been shown to regulate gene expression via the RNA-directed DNA methylation pathway and increase the production of cucurbitacins in agarwood [49]. We found that red light promoted the production of progoitrin, sinigrin, gluconapin, and glucobrassicanapin, as well as total GSLs. The glucosinolate-myrosinase system is thought to play a role in plant–pathogen and plant–herbivore interactions in Brassicaceae. The consumption of *Brassica* crops also has significant effects on human health, particularly cancer prevention. Although the physiological mechanisms are not completely understood, GSLs and their hydrolyzed products respond to different environmental conditions, including light and abiotic stresses. Studies of the crosstalk between environmental signal responses and glucosinolate metabolism should help to improve the quality and marketability of *Brassica* crops as functional foods.

5. Conclusions

In total, 14 R2R3-BrMYB TFs related to the glucosinolate pathway were selected based on differentially expressed genes (DEGs) from the Chinese cabbage transcriptome. Phylogenetic and sequence analyses of the R2R3-MYB transcription factors in Chinese cabbage and *Arabidopsis* showed the conserved nature of this family. The DEG expression profiles and protein–protein interaction analysis of the R2R3-MYBs indicate that they have important roles in secondary metabolic pathways. Furthermore, qRT-PCR analysis suggested that these genes function in response to light signals, supporting the functional diversity and importance of the R2R3-*BrMYB* genes in the production of glucosinolates in Chinese cabbage. This study improves the understanding of the R2R3-*BrMYB* genes, which play a vital role in GSL biosynthesis in Chinese cabbage and adds to the knowledge of light response mechanisms in plants.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Talalay, P.; Fahey, W. Phytochemicals from cruciferous plants protect against cancer by modulating carcinogen metabolism. *J. Nutr.* **2001**, *131*, 3027S–3033S. [CrossRef]
2. Hayes, J.D.; Kelleher, M.O.; Eggleston, I.M. The cancer chemo-preventive actions of phytochemicals derived from GSLs. *Eur. J. Nutr.* **2008**, *47*, 73–88. [CrossRef] [PubMed]
3. Kim, S.J.; Ishii, G. Glucosinolate profiles in the seeds, leaves and roots of rocket salad (*Eruca sativa* Mill.) and anti-oxidative activities of intact plant powder and purified 4-methoxyglucobrassicin. *Soil Sci. Plant Nutr.* **2006**, *52*, 394–400. [CrossRef]
4. Padilla, G.; Cartea, M.E.; Velasco, P.; Haro, A.D.; Ordras, A. Variation of glucosinolates in vegetable crops of *Brassica rapa*. *Phytochemistry.* **2007**, *68*, 536–545. [CrossRef] [PubMed]
5. Sønderby, I.E.; Burow, M.; Rowe, H.C.; Kliebenstein, D.J.; Halkier, B.A. A complex interplay of three R2R3 MYB transcription factors determines the profile of aliphatic GSLs in *Arabidopsis*. *Plant Physiol.* **2010**, *153*, 348–363. [CrossRef]
6. Del Carmen Martinez-Ballesta, M.; Moreno, D.A.; Carvajal, M. The physiological importance of glucosinolates on plant response to abiotic stress in *Brassica*. *Int. J. Mol. Sci.* **2013**, *14*, 11607–11625. [CrossRef]
7. Klempnauer, K.H.; Gonda, T.J.; Bishop, J.M. Nucleotide sequence of the retroviral leukemia gene v-myb and its cellular progenitor c-myb: The architecture of a transduced oncogene. *Cell* **1982**, *31*, 453–463. [CrossRef]
8. Lipsick, J.S. One billion years of Myb. *Oncogene* **1996**, *13*, 223–235.
9. Tominaga, R.; Iwata, M.; Okada, K.; Wada, T. Functional analysis of the epidermal-specific *MYB* genes CAPRICE and WEREWOLF in *Arabidopsis*. *Plant Cell* **2007**, *19*, 2264–2277. [CrossRef]
10. Abe, H.; Urao, T.; Ito, T.; Seki, M.; Shinozaki, K.; Yamaguchi-Shinozaki, K. *Arabidopsis* AtMYC2 (bHLH) and AtMYB2 (MYB) function as transcriptional activators in abscisic acid signaling. *Plant Cell* **2003**, *15*, 63–78. [CrossRef]
11. Stracke, R.; Ishihara, H.; Huep, G.; Barsch, A.; Mehrtens, F.; Niehaus, K.; Weisshaar, B. Differential regulation of closely related R2R3-MYB transcription factors controls flavonol accumulation in different parts of the *Arabidopsis thaliana* seedling. *Plant J.* **2007**, *50*, 660–677. [CrossRef] [PubMed]
12. Jung, C.; Seo, J.S.; Han, S.W.; Koo, Y.J.; Kim, C.H.; Song, S.I.; Nahm, B.H.; Choi, Y.D.; Cheong, J.J. Overexpression of AtMYB44 enhances stomatal closure to confer abiotic stress tolerance in transgenic *Arabidopsis*. *Plant Physiol.* **2008**, *146*, 623–635. [CrossRef] [PubMed]
13. Mun, J.H.; Kwon, S.J.; Yang, T.J.; Seol, Y.J.; Jin, M.; Kim, J.A.; Lim, M.H.; Kim, J.S.; Baek, S.; Choi, B.S.; et al. Genome-wide comparative analysis of the *Brassica rapa* gene space reveals genome shrinkage and differential loss of duplicated genes after whole genome triplication. *Genome Biol.* **2009**, *10*, R111. [CrossRef] [PubMed]
14. Wang, Y.; Wang, X.; Tang, H.; Tan, X.; Ficklin, S.P.; Feltus, F.A.; Paterson, A.H. Modes of gene duplication contribute differently to genetic novelty and redundancy but show parallels across divergent angiosperms. *PLoS ONE* **2011**, *6*, e28150. [CrossRef] [PubMed]
15. Udvardi, M.K.; Kakar, K.; Wandrey, M.; Montanri, O.; Murray, J.; Andriankaja, A.; Zhang, J.Y.; Benedito, V.; Hofer, J.M.I.; Chueng, F.; et al. Legume transcription factors: Global regulators of plant development and response to the environment. *Plant Physiol.* **2007**, *144*, 538–549. [CrossRef]
16. Stracke, R. The *R2R3-MYB* gene family in *Arabidopsis thaliana*. *Curr. Opin. Plant Biol.* **2001**, *4*, 447–456. [CrossRef]
17. Wang, Z.; Tang, J.; Hu, R.; Wu, P.; Hou, X.L.; Song, X.M.; Xiong, A.S. Genome-wide analysis of the R2R3-MYB transcription factor genes in Chinese cabbage (*Brassica rapa* ssp. *pekinensis*) reveals their stress and hormone responsive patterns. *BMC Genomics* **2015**, *16*, 17. [CrossRef]
18. Cheng, F.; Liu, S.; Wu, J.; Fang, L.; Sun, S.; Liu, B. BRAD, the genetics and genomics database for *Brassica* plants. *BMC Plant Biol.* **2011**, *11*, 136. [CrossRef]
19. Punta, M.; Coggill, P.C.; Eberhardt, R.Y.; Mistry, J.; Tate, J.; Boursnell, C. The Pfam protein family's database. *Nucl. Acids Res.* **2012**, *40*, D290–D301. [CrossRef]
20. Thompson, J.D.; Gibson, T.J.; Plewniak, F.; Jeanmougin, F.; Higgins, D.G. The CLUSTAL_X Windows interface: Flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucl. Acids Res.* **1997**, *25*, 4876–4882. [CrossRef]

21. Kumar, S.; Stecher, G.; Tamura, K. MEGA7: Molecular Evolutionary Genetics Analysis version 7.0 for Bigger Datasets. *Mol. Biol. Evol.* **2016**, *33*, 1870–1874. [CrossRef] [PubMed]

22. Bailey, T.L.; Williams, N.; Misleh, C.; Li, W.W. MEME: Discovering and analyzing DNA and protein sequence motifs. *Nucl. Acids Res.* **2006**, *34*, 369–373. [CrossRef] [PubMed]

23. Moon, J.; Jeong, M.J.; Lee, S.I.; Lee, J.G.; Hwang, H.; Yu, J.; Kim, Y.R.; Park, S.W.; Kim, J.A. Effect of LED mixed light conditions on the glucosinolate pathway in *Brassica rapa*. *J. Plant Biotechnol.* **2015**, *42*, 245–256. [CrossRef]

24. Livak, K.J.; Schmittgen, T.D. Analysis of relative gene expression data using real-time quantitative PCR and the 2 (-Delta Delta C(T)) Method. *Methods* **2001**, *25*, 402–408. [CrossRef]

25. Bhandari, S.R.; Jo, J.S.; Lee, J.G. Comparison of Glucosinolate Profiles in Different Tissues of Nine *Brassic*a Crops. *Molecules* **2015**, *20*, 15827–15841. [CrossRef]

26. Kim, J.A.; Shim, D.; Kumari, S.; Jung, H.; Jung, K.H.; Jeong, H.; Kim, W.Y.; Lee, S.I.; Jeong, M.J. Transcriptome analysis of diurnal gene expression in Chinese cabbage. *Genes* **2019**, *10*, 130. [CrossRef]

27. Szklarczyk, D.; Franceschini, A.; Wyder, S.; Forslund, K.; Heller, D.; Huerta-Cepas, J.; Simonovic, M.; Roth, A.; Santos, A.; Tsafou, K.P.; et al. STRING v10: Protein–protein interaction networks, integrated over the tree of life. *Nucl. Acids Res.* **2015**, *43*, D447–D452. [CrossRef]

28. Augustine, R.; Majee, M.; Gershenzon, J.; Bisht, N.C. Four genes encoding MYB28, a major transcriptional regulator of the aliphatic glucosinolate pathway, are differentially expressed in the allopolyploid *Brassica juncea*. *J. Exp. Bot.* **2013**, *64*, 4907–4921. [CrossRef]

29. Ogata, K.; Kanei-Ishii, C.; Sasaki, M.; Hatanaka, H.; Nagadoi, A.; Enari, M.; Sarai, A. The cavity in the hydrophobic core of Myb DNA-binding domain is reserved for DNA recognition and trans-activation. *Nat. Struct Biol.* **1996**, *3*, 178–187. [CrossRef]

30. Kranz, H.; Stracke, R.; Weisshaar, B. c-MYB oncogene-like genes encoding three MYB repeats occur in all major plant lineages. *Plant J.* **2000**, *21*, 231–235. [CrossRef]

31. Urzainqui, A.; Bevan, M.; Martin, C.; Smeekens, S.; Tonelli, C.; Paz-Ares, J.; Weisshaar, B. Towards functional characterization of the members of the R2R3-*MYB* gene family from *Arabidopsis thaliana*. *Plant J.* **1998**, *16*, 263–276.

32. Jin, H.; Martin, C. Multi functionality and diversity within the plant MYB-gene family. *Plant Mol. Biol.* **1999**, *41*, 577–585. [CrossRef] [PubMed]

33. Dubos, C.; Stracke, R.; Grotewold, E.; Weisshaar, B.; Martin, C.; Lepiniec, L. MYB transcription factors in *Arabidopsis*. *Trends Plant Sci.* **2010**, *15*, 573–581. [CrossRef] [PubMed]

34. Tong, C.; Wang, X.; Yu, J.; Wu, J.; Li, W.; Huang, J.; Dong, C.; Hua, W.; Liu, S. Comprehensive analysis of RNA-seq data reveals the complexity of the transcriptome in *Brassica rapa*. *BMC Genom.* **2013**, *14*, 689. [CrossRef]

35. Kalsotra, A.; Cooper, T.A. Functional consequences of developmentally regulated alternative splicing. *Nat. Rev. Genet.* **2011**, *12*, 715–729. [CrossRef]

36. Zhou, R.; Moshgabadi, N.; Adams, K.L. Extensive changes to alternative splicing patterns following allopolyploidy in natural and resynthesized polyploids. *Proc. Natl. Acad. Sci USA* **2011**, *108*, 16122–16127. [CrossRef]

37. Adams, K.L. Evolution of duplicate gene expression in polyploidy and hybrid plants. *J. Hered.* **2007**, *98*, 136–141. [CrossRef]

38. Seo, M.-S.; Jin, M.; Chun, J.-H.; Kim, S.-J.; Park, B.-S.; Shon, S.-H.; Kim, J.S. Functional analysis of three BrMYB28 transcription factors controlling the biosynthesis of glucosinolates in Brassica rapa. *Plant Mol. Biol.* **2016**, *90*, 503–516. [CrossRef]

39. Katiyar, A.; Smita, S.; Lenka, S.K.; Rajwanshi, R.; Chinnusamy, V.; Bansal, K.C. Genome-wide classification and expression analysis of MYB transcription factor families in rice and *Arabidopsis*. *BMC Genom.* **2012**, *13*, 544. [CrossRef]

40. Gigolashvili, T.; Engqvist, M.; Yatusevich, R.; Muller, C.; Flugge, U.I. The transcription factor HIG1/MYB51 regulates indolic glucosinolate biosynthesis in *Arabidopsis thaliana*. *Plant J.* **2007**, *50*, 886–901. [CrossRef]

41. Smith, H. Light quality, photoperception, and plant strategy. *Annu. Rev. Plant Physiol.* **1982**, *33*, 481–518. [CrossRef]

42. Sang, J.P.; Minchinton, I.R.; Johnstone, P.K.; Truscott, R.J.W. Glucosinolate profiles in the seed, root and leaf tissue of cabbage, mustard, rapeseed radish and swede. *Can. J. Plant Sci.* **1984**, *64*, 77–93. [CrossRef]

43. Rosa, E.A.S.; Rodrigues, P.M.F. The effect of light and temperature on glucosinolate concentration in the leaves and roots of cabbage seedlings. *J. Sci. Food Agric.* **1998**, *78*, 208–212. [CrossRef]

44. Sofia, D.C.; Kevin, M.F. Sequential light programs shape kale (*Brassica napus*) sprout appearance and alter metabolic and nutrient content. *Hortres* **2014**, *8*. [CrossRef]

45. Abe, K.; Kido, S.; Maeda, T.; Kami, D.; Matsuura, H.; Shimura, H.; Suzuki, T. Glucosinolate profiles in *Cardamine fauriei* and effect of light quality on glucosinolate concentration. *Sci. Hort.* **2015**, *189*, 12–16. [CrossRef]

46. Huseby, S.; Koprivova, A.; Lee, B.-R.; Saha, S.; Mithen, R.; Wold, A.-B.; Bengtsson, G.B.; Kopriva, S. Diurnal and light regulation of sulphur assimilation and glucosinolate biosynthesis in *Arabidopsis*. *J. Exp. Bot.* **2013**, *64*, 1039–1048. [CrossRef]

47. Kim, Y.B.; Li, X.; Kim, S.J.; Kim, H.H.; Lee, J.; Kim, H.; Park, S.U. MYB transcription factors regulate glucosinolate biosynthesis in different organs of Chinese cabbage (*Brassica rapa* ssp. *pekinensis*). *Molecules* **2013**, *18*, 8682–8695. [CrossRef]

48. Lee, M.K.; Arasu, M.V.; Park, S.; Byeon, D.H.; Chung, S.O.; Park, S.U.; Lim, Y.P.; Kim, S.J. LED lights enhance metabolites and antioxidants in Chinese cabbage and kale. *Braz. Arch. Bio. Technol.* **2016**, *59*, e16150546. [CrossRef]

49. Kuo, T.C.Y.; Chen, C.H.; Chen, S.H.; Lu, I.H.; Chu, M.J.; Huang, L.C.; Lin, C.Y.; Chen, C.Y.; Lo, H.F.; Jeng, S.T.; et al. The effect of red light and far-red light conditions on secondary metabolism in Agarwood. *BMC Plant Biol.* **2015**, *15*, 139. [CrossRef]