



## Article

# Transcriptome Profiling Reveals Candidate Key Genes Involved in Sinigrin Biosynthesis in *Brassica nigra*

Yang Li <sup>1,2,†</sup>, Youjian Yu <sup>1,†</sup>, Liai Xu <sup>1</sup> , Erbiao Guo <sup>1</sup>, Yunxiang Zang <sup>1</sup>, Yong He <sup>1,\*</sup> and Zhujun Zhu <sup>1,\*</sup>

<sup>1</sup> Key Laboratory for Quality Improvement of Agricultural Products of Zhejiang Province, Collaborative Innovation Center for Efficient and Green Production of Agriculture in Mountainous Areas of Zhejiang Province, College of Horticulture Science, Zhejiang A&F University, Wusu Street 666, Lin'an, Hangzhou 311300, China; lylily0327@163.com (Y.L.); yjyu@zafu.edu.cn (Y.Y.); 11416052@zju.edu.cn (L.X.); geb1632021@163.com (E.G.); yxzang78@163.com (Y.Z.)

<sup>2</sup> College of Forestry and Biotechnology, Zhejiang A&F University, Wusu Street 666, Lin'an, Hangzhou 311300, China

\* Correspondence: heyong@zafu.edu.cn (Y.H.); zhujun@zafu.edu.cn (Z.Z.); Tel.: +86-571-6374-3001 (Z.Z.)

† These authors have contributed equally to this work.

**Abstract:** Glucosinolates (GSLs) are important secondary metabolites in Brassicales related to insect and disease resistance, flavor formation, and human health. Here, we determined the GSL profile with sinigrin as the predominant GSL in *Brassica nigra*. A total of 184 GSL biosynthetic genes (*BniGSLs*) were identified in *B. nigra* by a genome-wide search for orthologs of 82 of the 95 known GSL genes in *Arabidopsis thaliana*. Transcriptome data demonstrated that at least one *BniGSL* was highly expressed in stems and leaves at each step of the sinigrin synthesis pathway, which ensured the synthesis of a large amount of sinigrin in *B. nigra*. Among these key candidates of *BniGSLs*, the high expression of *BniMAM1-2*, *BniCYP79F1*, and *BniAOP2-1/2*, and the absence of *MAM3* and *AOP3*, may contribute remarkably to the synthesis and accumulation of sinigrin. In addition, the low expression of some key *BniGSLs* partially explains the low content of indolic and aromatic GSLs in *B. nigra*. This study provided a genetic explanation for the formation of the unique GSL profile with sinigrin as the main GSL in *B. nigra*. The results of this study will be valuable for further functional analysis of *BniGSLs* and genetic improvement of GSLs in *B. nigra* and other *Brassica* species.

**Keywords:** glucosinolates; sinigrin; biosynthetic genes; gene expression; *Brassica nigra*



**Citation:** Li, Y.; Yu, Y.; Xu, L.; Guo, E.; Zang, Y.; He, Y.; Zhu, Z.

Transcriptome Profiling Reveals Candidate Key Genes Involved in Sinigrin Biosynthesis in *Brassica nigra*. *Horticulturae* **2021**, *7*, 173. <https://doi.org/10.3390/horticulturae7070173>

Academic Editors: Xiaowu Wang and Jian Wu

Received: 26 May 2021

Accepted: 27 June 2021

Published: 2 July 2021

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Glucosinolates (GSLs) are a group of sulfur-rich and nitrogen-containing secondary metabolites that are synthesized from amino acids and sugars in plants. Currently, over 100 different GSLs have been identified [1], most of which are exclusively found in the order Brassicales [2]. GSL is composed of three common moieties, including a  $\beta$ -D-thioglucose group, a sulfonated aldoxime moiety, and a variable side chain derived from a precursor amino acid [3,4]. Upon hydrolysis by the myrosinase enzyme, GSLs are degraded into different bioactive products, mainly isothiocyanates [5]. These broken down products exhibit a variety of biological activities, which not only endow *Brassica* vegetables with characteristic flavor [6] and help defend against pathogens and insect herbivores [7], but also function in preventing carcinogenesis in animals by stimulating apoptosis and regulating the cell cycle [8]. For instance, mounting studies have shown that the enzymatic hydrolysate of sinigrin has anti-cancer, anti-inflammatory, anti-oxidant, anti-bacterial, antifungal, wound healing properties, and biofumigation applications [9]. The pharmacological and therapeutic properties of GSLs that are beneficial to human health have made the *Brassica* species attract the considerable interest of many plant breeders and geneticists in the past 30 years [3,10,11].

On the basis of the chemical structure of different side chains, GSLs are classified into aliphatic, indolic, and aromatic GSLs, with methionine (or alanine, valine, leucine,

and isoleucine), tryptophan, and phenylalanine (or tyrosine) as the basic amino acid precursors, respectively [3,12]. The GSL biosynthesis is a tripartite pathway involving three independent phases in at least two different locations: (i) side-chain elongation of selected precursors (only methionine and phenylalanine) in the chloroplast, (ii) formation of the core GSL structure in the cytosol interface of the endoplasmic reticulum, and (iii) side-chain modification in the cytosol [3,10,11]. To date, GSL biosynthesis has been well elucidated mainly in *Arabidopsis*, and the inventory of related genes in this process is close to completion [11]. Most GSLs in *Arabidopsis* and *Brassica* crops are synthesized from methionine, beginning with side-chain elongation and condensation, which involves branched-chain aminotransferase (BCAT), bile acid transporter 5 (BAT5), and methylthioalkylmalate synthase (MAM) [13–15]. Subsequently, core structures are formed via a five-step process that includes the conversion of 2-oxo-methylthio acid homologs into aldoximes by cytochrome P450 of the CYP79s family [16], oxidation of aldoximes by the CYP83s family [17,18], followed by C-S cleavage (SUR1) [19], and the formation of desulfoglucosinolate (UGT74s) [20] and the basic GSL structure (SOTs) [21,22]. Finally, the side chains of these basic GSLs are modified by oxygenation, alkenylation, hydroxylation, and benzoylation. Genes involving these modification events, including those encoding GS-ELONG, GS-OX, GS-AOP, and GS-OH, have been well studied in recent years [10,11,23,24]. Several transcription factors (TFs) have also been characterized to participate in the regulation of GSL biosynthesis [25–29].

*Brassica nigra*, commonly known as black mustard, is a member of the Brassicaceae family that has been cultivated for thousands of years over a wide range of climates [30]. *B. nigra* has a strong, pungent flavor and has been used as a major condiment crop in many countries. Its seed has also been used as a traditional herbal medicine [31,32]. *B. nigra* is one of the three important diploid species (*B. rapa* (AA), *B. nigra* (BB), and *B. oleracea* (CC)) in the so-called “Triangle of U” theory that explains how the three allotetraploid *Brassica* species (*B. juncea* (AABB), *B. napus* (AACC), and *B. carinata* (BBCC)) were evolved [33,34]. Although the current genomes of these three progenitor *Brassica* species were derived from a common hexaploid ancestor, the evolutionary process of gene loss and neo- or sub-functionalization of genes made the genome of *B. nigra* evolve separately from *B. rapa* and *B. oleracea* [35]. In terms of GSL synthesis, the differences in the diversity and content of GSLs in these three *Brassica* species have been confirmed. *B. oleracea* mainly synthesizes 4C and 3C GSLs [36,37], *B. rapa* accumulates 4C and 5C GSLs [38]. However, with sinigrin as the predominant GSL, the GSL profile in *B. nigra* is much different from those in *B. rapa* and *B. oleracea* [30,39–41]. Excitingly, genome-wide characterization and expression analysis of genes involved in GSL synthesis in *B. rapa* and *B. oleracea* can more or less explain the differences [42,43]. The genetic and genomic information of *B. nigra* have been published through whole-genome sequencing recently [44]. However, no systematic investigation on GSL genes have been reported in *B. nigra* to date, and the expression patterns of genes related to GSL metabolism in different organs of *B. nigra* are still limited and remain to be further explored.

In this study, we first investigated the GSL profile of black mustard and found that sinigrin was the most predominant GSL in *B. nigra*. Next, we conducted a genome-wide in silico search to determine the GSL biosynthesis genes by using the assembled genome sequence data of *B. nigra* [44], and provided detailed information of *B. nigra* GSL genes (*BniGSLs*). Combining the expression profile data and phylogenetic tree analysis of key GSL genes, we screened out the key candidate *BniGSLs* involved in the massive synthesis of sinigrin in *B. nigra*. The knowledge gained in this study will be useful for further studies on the biological functions of *BniGSLs* and the genetic improvement of GSLs in *B. nigra* and other *Brassica* species.

## 2. Materials and Methods

### 2.1. Plant Materials and Growth Conditions

The black mustard genotype used in this experiment was *B. nigra* cv. 1511-01 (an inbred line kept in our laboratory). The seeds of the mustard were planted in a controlled climate chamber at 26 °C/20 °C day/night temperature, 14/10 h light/dark photoperiod, 600  $\mu\text{mol}\cdot\text{m}^{-2}\cdot\text{s}^{-1}$  light intensity, and 60–70% relative humidity. The stems, rosette leaves, cauline leaves, inflorescences, and siliques (about 3 cm in length) used for the GSL profiling were collected during the flowering period. Cauline leaves, stems, and roots were used for RNA sequencing. All the materials were mixed and frozen in liquid nitrogen immediately and stored at  $-80$  °C. All samples were collected from at least three plants, and both glucosinolate analysis and RNA sequencing were carried out in three biological replicates.

### 2.2. GSL Extraction and Analysis

GSL extraction and analysis were performed as previously described with only slight modification [45]. First, 0.25 g sample powder was boiled with 10 mL of 70% methanol after adding 200  $\mu\text{L}$  of 5 mM glucotropaeolin (CAS, 5115-71-9; Code No., A5300,0050; Applichem, Darmstadt, Germany) as internal standard. Then, the supernatant was loaded onto a 1 mL mini-column to desulfate overnight with 200  $\mu\text{L}$  arylsulfatase (Sigma-Aldrich Co., St. Louis, MO, USA). The mini-column containing 250  $\mu\text{L}$  activated DEAE Sephadex A-25 was equilibrated at room temperature for at least 2 h prior to use. Resultant desulfoglucosinolates were eluted with ultrapure water and stored at  $-20$  °C until analysis. Samples were analyzed by high-performance liquid chromatography (HPLC) in an Agilent 1200 HPLC system equipped with a C-18 reversed-phase column (250  $\times$  4  $\mu\text{m}$ , 5  $\mu\text{m}$ , Bischoff, Leonberg, Germany). Elution was performed with ultrapure water (solvent A) and acetonitrile (solvent B) in a linear gradient from 0% to 20% B for 45 min and then constant 20% B for 6 min, followed by 100% A for 5 min prior to the injection of the next sample. The flow rate was 1  $\text{mL}\cdot\text{min}^{-1}$  (injection volume of 20  $\mu\text{L}$ ). The eluent was monitored by diode array detection at 229 nm. The data of GSL concentrations were analyzed using analysis of variance (ANOVA) software. Mean values were compared using the least significant difference at 0.05 significance level.

### 2.3. In Silico Identification of *BniGSL* Genes in *B. nigra*

Based on previous studies [11,42,46], sequences of the GSL biosynthetic genes in *A. thaliana*, *B. rapa*, and *B. oleracea* (Table S1) were acquired from The Arabidopsis Information Resource (TAIR) website (<http://www.arabidopsis.org/>) (accessed on 10 November 2020), The National Center for Biotechnology Information (NCBI) website (<http://www.ncbi.nlm.nih.gov>) (accessed on 10 November 2020), and the *Brassica* database website (<http://brassicadb.cn/#/>, BRAD) (accessed on 10 November 2020). The whole sequences of each *B. nigra* chromosome were downloaded from the BRAD database (<http://brassicadb.cn/#/Download/>) (accessed on 10 November 2020). Using the sequences of GSL genes acquired from the databases mentioned above, BLASTn was performed to search for homologous candidate genes in *B. nigra*. All candidates in the *B. nigra* genome, together with flank regions of 5000 bp upstream and downstream of each candidate, were analyzed and re-annotated using FGENESH (<http://www.softberry.com/>) (accessed on 10 November 2020). By comprehensively considering the candidate's re-annotation results, the collinearity relationship with the known GSL genes, the similarity of amino acid sequence, and whether it contained the corresponding key domains or sequences, it is determined whether the candidate is a GSL gene. Given that the GSL genes of *A. thaliana* have been clearly named, the nomenclature system for the *BniGSL* genes in this study was based on their homology and identities with their counterparts in *Arabidopsis*. The resulting *BniGSL* genes were further used as query sequences to determine the precise locations of each gene on chromosomes through Oligo 6.0 software.

#### 2.4. RNA Extraction, Library Construction, Sequencing, and Gene Expression Analysis

Total RNA of leaves, stems, and roots was isolated using Trizol reagent (Invitrogen, Waltham, MA, USA) following the manufacturer's instructions. Nanodrop, Qubit 2.0, and Agilent 2100 were used separately to measure the purity, concentration, integrity, and other values of RNA to ensure the qualified samples for transcriptome sequencing. The mRNA was purified with 20 mg total RNA by using oligo (dT) magnetic beads. After purification, the mRNA was fragmented by adding a fragmentation buffer. The fragments were used to synthesize the first-strand cDNA by using random hexamer adapters and reverse transcriptase (Code No., RR047A; Takara, Japan). The second-strand cDNA was synthesized using buffer, dNTPs, RNase H, and DNA polymerase I. The cDNA fragments that went through an end-repair process were added with a single 'A' base and had the sequencing joints connected. AMPure XP beads were used to select the fragment size for the ligation of adapter sequences. The products were purified, enriched with PCR, and then used as templates for sequencing. Sequencing and assembly were performed by the Biomarker Biotechnology Corporation (Beijing, China) using the Illumina HiSeq™ 2500 platform. The Log2FoldChange and false discovery rate (FDR) value of *BniGSL* genes between two samples (Root vs Stem, Root vs Leaf, and Stem vs Leaf) were calculated, and the *BniGSL* genes with  $|\text{Log2FoldChange}| > 1$  and  $\text{FDR} < 0.05$  were regarded as differentially expressed genes.

All RNA samples used for RNA-Seq were also used for qRT-PCR analysis (Code NO., RR820A; Takara, Japan). All qRT-PCR experiments included three independent biological repetitions. *Brassica nigra* tonoplast intrinsic protein-41 (*TIPS*) gene was used as a reference gene [47]. The  $2^{-\Delta\Delta C_t}$  method was used to calculate the relative gene expression values. The gene-specific primers were listed in Table S2.

The expression data of *B. rapa* is from Tong et al. [48] (GEO accession number GSE43245). Root, stem, and leaf tissues of *B. rapa* accession Chiifu-401-42 were collected from seven-week-old plants. The expression data of *B. oleracea* is from Liu et al. [43] and Yu et al. [49] (GEO accession number GSE 42891).

### 3. Results

#### 3.1. Analysis of Glucosinolate Profile in *B. nigra*

HPLC analyses revealed the presence of seven different types of GSLs in five organs of *B. nigra* cv. 1511-01 (Table 1, Figure S1), including two aliphatic GSLs (i.e., sinigrin and gluconapin), four indolic GSLs (i.e., 4-hydroxy glucobrassicin, glucobrassicin, 4-methoxyglucobrassicin, and 1-methoxyglucobrassicin), and one aromatic GSL (i.e., gluconasturtiin). Results showed that the stems of black mustard contained all the above-mentioned GSLs. Rosette leaves contained all GSLs except gluconapin. Cauline leaves and siliques contained all GSs except gluconasturtiin. However, only four GSLs were detected in the inflorescences of black mustard, with three indolic GSLs were absent. In terms of total GSL concentration, black mustard stems possessed the highest concentration of total GSL, followed by inflorescences, cauline leaves and siliques, and rosette leaves contained the lowest total GSL content. Notably, although the GSL profiles in different organs of black mustard were different, sinigrin was the predominant GSL in all five tested organs and accounted for 90.7–98.5% of the total GSL. Moreover, as the main GSL, the content of sinigrin in five organs ranked the same as the total GSL. These results indicated that the GSL profiles in *B. nigra* are significantly different from those in *B. rapa* and *B. oleracea*, suggesting that there may be a major GSL biosynthesis pathway in *B. nigra*, which directs the synthesis of sinigrin.

**Table 1.** Glucosinolate content ( $\mu\text{mol}\cdot\text{g}^{-1}$  DW) in different organs of *Brassica nigra*.

Organ	Aliphatic			Indolic			Aromatic
	SIN	GNA	4OHI3M	I3M	4MOI3M	1MOI3M	2PHET
St	82.35 ± 2.12d	0.52 ± 0.3c	0.25 ± 0.02a	0.12 ± 0.02b	0.05 ± 0.02ab	0.09 ± 0.02b	0.25 ± 0.01a
RL	13.20 ± 1.08a	-	0.50 ± 0.14b	0.04 ± 0.01a	0.04 ± 0.01a	0.03 ± 0.01a	0.78 ± 0.10b
CL	53.80 ± 2.52b	0.32 ± 0.05b	0.32 ± 0.01ab	0.13 ± 0.02b	0.05 ± 0.00ab	0.08 ± 0.01b	-
In	59.14 ± 2.39c	0.32 ± 0.25b	1.50 ± 0.17c	-	-	-	0.34 ± 0.01a
Si	51.31 ± 0.25b	0.35 ± 0.11a	0.37 ± 0.07ab	0.13 ± 0.03b	0.06 ± 0.01b	0.08 ± 0.01b	-

Different letters indicate significant difference. SIN, Sinigrin; GNA, Gluconapin; 4OHI3M, 4-OH-Glucobrassicin; I3M, Glucobrassicin; 4MOI3M, 4-Methoxy-Glucobrassicin; 1MOI3M, 1-Methoxy-Glucobrassicin; 2PHET, Gluconasturtiin. St, stems; RL, rosette leaves; CL, cauline leaves; In, inflorescences; Si, siliques; DW, dry weight.

### 3.2. Identification and Annotation of *BniGSL* Genes from *B. nigra* Genome

We first searched for *BniGSL* genes in the whole genome of *B. nigra* before looking for clues as to why black mustard preferred to synthesize sinigrin. Preliminary BLAST searches for *BniGSL* genes in the whole-genome sequences of *B. nigra* were performed using *GSL* genes of *A. thaliana*, *B. rapa*, and *B. oleracea*. Using these pre-screened *BniGSL* genes, re-annotation, and BLASTP search against known *GSL* gene sequences of *A. thaliana*, *B. rapa*, and *B. oleracea* resulted in 184 *BniGSL* genes as orthologs of 82 of the 95 known *AtGSL* genes, with 13 *AtGSL* genes having no *B. nigra* ortholog. The number of *BniGSL* genes in *B. nigra* has expanded, with an average of 2.24 copies per gene. It is worth noting that there are thirteen copies of SOT18 in *B. nigra*, which is far more than the average.

All the identified *BniGSL* genes are listed in Tables 2 and 3, and the sequences of DNA, CDS, and amino acid are listed in Table S3. A total of 124, 48, and 12 *BniGSLs* encode enzymes involved in glucosinolate biosynthesis (Table 2), TFs with regulatory functions, and transporters involved in glucosinolate transport (Table 3), respectively. More specifically, there are 16, 55, 21, and 32 genes involved in side-chain elongation, core structure synthesis, side-chain modification, and co-substrate pathways, respectively (Table 2). Among the 48 TFs, 33, 8, and 4 TFs act as activators, repressors, and mediators, respectively. HY5, which has both activating and suppressing functions during glucosinolate biosynthesis, consists of three copies in *B. nigra* (Table 3). In addition, there are 12 *BniGSL* genes encoding five transporters in black mustard (Table 3).

**Table 2.** The inventory of glucosinolate biosynthetic genes (*BniGSLs*) in *Brassica nigra*.

Name1	Name2	Gene ID	Chromosome Location		AA	Identity/%	AGI ID
<b>Side-chain elongation</b>							
BCAT4	<i>BniBCAT4-1</i>	BniB01g044210	B1 (+)	48,590,123–48,592,361	356	82.3	AT3G19710
	<i>BniBCAT4-2</i>	BniB07g030730	B7 (–)	41,353,700–41,356,224	306	62.7	
BCAT6	<i>BniBCAT6</i>	BniB04g060340	B4 (+)	51,535,934–51,537,864	359	86.7	AT1G50110
MAM1	<i>BniMAM1-1</i>	BniB04g003060	B4 (+)	1,498,531–1,501,807	505	79.8	AT5G23010
	<i>BniMAM1-2</i>	BniB02g076300	B2 (+)	64,723,274–64,726,274	504	79.1	
	<i>BniMAM1-3</i>	BniB08g036280	B8 (–)	21,370,428–21,376,553	495	74.8	
	<i>BniMAM1-4</i>	BniB02g076310	B2 (+)	64,728,011–64,730,517	397	59.6	
	<i>BniMAM1-5</i>	BniB05g055400	B5 (–)	49,610,184–49,616,319	388	56.6	
MAM3	*						AT5G23020
IPMI SSU1	<i>BniIPMI SSU1-1</i>	BniB08g028010	B8 (+)	14,699,124–14,699,873	249	87.7	AT2G43090
	<i>BniIPMI SSU1-2</i>	BniB06g001990	B6 (+)	917,687–917,463	258	79.6	
IPMI SSU2	<i>BniIPMI SSU2</i>	BniB06g002000	B6 (+)	919,426–920,271	281	72.7	AT2G43100

Table 2. Cont.

Name1	Name2	Gene ID	Chromosome	Location	AA	Identity/%	AGI ID
IPMI SSU3	*						AT3G58990
ILL1	<i>BniILL1-1</i>	BniB07g011950	B7 (+)	15,245,075–15,248,497	505	94.5	AT4G13430
	<i>BniILL1-2</i>	BniB07g011910	B7 (+)	15,114,671–15,118,112	505	94.3	
IMD1	<i>BniIMD1</i>	BniB05g024330	B5 (–)	11,915,893–11,917,777	408	91.7	AT5G14200
BCAT3	<i>BniBCAT3-1</i>	BniB08g063340	B8 (+)	59,567,292–59,569,478	419	87.1	AT3G49680
	<i>BniBCAT3-2</i>	BniB05g048020	B5 (–)	27,787,734–27,790,073	419	84.5	
<b>Core structure formation</b>							
CYP79F1	<i>BniCYP79F1</i>	BniB06g041700	B6 (+)	46,406,179–46,408,341	541	82.1	AT1G16410
CYP79F2	*						AT1G16400
CYP79A2	<i>BniCYP79A2-1</i>	BniB02g055340	B2 (+)	53,932,413–53,934,250	532	77.6	AT5G05260
	<i>BniCYP79A2-2</i>	BniB02g055360	B2 (+)	53,941,026–53,942,864	532	77.6	
	<i>BniCYP79A2-3</i>	BniB02g055420	B2 (+)	53,958,847–53,960,685	532	76.3	
	<i>BniCYP79A2-4</i>	BniB05g029220	B5 (+)	14,358,148–14,360,431	562	73.1	
	<i>BniCYP79A2-5</i>	BniB08g051390	B8 (+)	34,728,808–34,732,349	439	56.6	
CYP79B2	<i>BniCYP79B2-1</i>	BniB02g088940	B2 (–)	70,487,544–70,489,564	541	94.5	AT4G39950
	<i>BniCYP79B2-2</i>	BniB05g000130	B5 (–)	64,703–66,466	542	93.2	
	<i>BniCYP79B2-3</i>	BniB03g014770	B3 (–)	6,330,659–6,332,937	518	89.8	
CYP79B3	<i>BniCYP79B3</i>	BniB01g021760	B1 (–)	13,001,786–13,004,048	564	89.0	AT2G22330
CYP79C1	<i>BniCYP79C1-1</i>	BniB08g040770	B8 (–)	25,176,913–25,179,607	531	80.3	AT1G79370
	<i>BniCYP79C1-2</i>	BniB08g040860	B8 (–)	25,245,858–25,248,871	531	80	
	<i>BniCYP79C1-3</i>	BniB08g040830	B8 (–)	25,220,662–25,222,785	396	62.1	
	<i>BniCYP79C1-4</i>	BniB08g040890	B8 (–)	25,258,462–25,261,159	382	59.7	
CYP79C2	<i>BniCYP79C2-1</i>	BniB04g007960	B4 (–)	3,767,912–3,772,131	530	78.7	AT1G58260
	<i>BniCYP79C2-2</i>	BniB01g003800	B1 (–)	1,903,991–1,908,135	528	74.9	
	<i>BniCYP79C2-3</i>	BniB06g009140	B6 (–)	4,583,904–4,586,034	527	73.6	
CYP83A1	<i>BniCYP83A1-1</i>	BniB01g003820	B1 (+)	1,927,665–1,929,469	500	86.9	AT4G13770
	<i>BniCYP83A1-2</i>	BniB06g067060	B6 (–)	59,968,418–59,970,020	498	86.3	
CYP83B1	<i>BniCYP83B1</i>	BniB07g013920	B7 (–)	18,805,780–18,807,433	499	95.4	AT4G31500
CYTB5-C	<i>BniCYTB5-C-1</i>	BniB01g000830	B1 (–)	449,656–450,237	135	80.1	AT2G46650
	<i>BniCYTB5-C-2</i>	BniS02554g140	utg2554 +	100,663–101,227	114	57.7	
GSTF9	<i>BniGSTF9-1</i>	BniB01g013320	B1 (–)	6,927,611–6,928,497	215	97.2	AT2G30860
	<i>BniGSTF9-2</i>	BniB08g019200	B8 (+)	9,243,485–9,244,436	215	96.7	
GSTF10	<i>BniGSTF10</i>	BniB08g019210	B8 (+)	9,246,072–9,247,215	215	94.4	AT2G30870
GSTF11	<i>BniGSTF11</i>	BniB01g061730	B1 (+)	57,396,702–57,397,557	214	82.2	AT3G03190
GSTU13	<i>BniGSTU13-1</i>	BniB04g039630	B4 (–)	21,519,677–21,520,647	227	81.1	AT1G27130
	<i>BniGSTU13-2</i>	BniB03g031270	B3 (–)	14,627,558–14,628,949	227	76.2	
GSTU20	<i>BniGSTU20</i>	BniB06g050340	B6 (–)	51,398,415–51,399,277	217	88.9	AT1G78370
GGP1	<i>BniGGP1-1</i>	BniB02g084810	B2 (+)	68,616,575–68,618,095	250	88.4	AT4G30530
	<i>BniGGP1-2</i>	BniB05g007160	B5 (–)	3,545,600–3,546,824	250	86.4	
	<i>BniGGP1-3</i>	BniB03g010950	B3 (–)	4,449,737–4,451,401	228	80	
	<i>BniGGP1-4</i>	BniB05g007170	B5 (–)	3,548,331–3,552,009	251	62.4	
GGP3	<i>BniGGP3</i>	BniB03g010940	B3 (–)	4,446,818–4,448,424	258	84.9	AT4G30550
SUR1	<i>BniSUR1-1</i>	BniB03g070290	B3 (–)	54,745,695–54,747,733	458	88.6	AT2G20610
	<i>BniSUR1-2</i>	BniB02g015530	B2 (–)	9,570,907–9,578,729	445	83.3	
UGT74B1	<i>BniUGT74B1</i>	BniB04g039340	B4 (–)	21,237,953–21,239,435	465	85.2	AT1G24100
UGT74C1	<i>BniUGT74C1-1</i>	BniB06g020830	B6 (+)	11,872,313–11,874,700	456	85.1	AT2G31790
	<i>BniUGT74C1-2</i>	BniB01g012360	B1 (–)	6,333,513–6,335,348	456	83.6	
SOT16	<i>BniSOT16-1</i>	BniB03g052560	B3 (–)	28,106,607–28,107,620	337	92.9	AT1G74100
	<i>BniSOT16-2</i>	BniB05g070980	B5 (–)	64,926,130–64,927,143	337	92.9	

Table 2. Cont.

Name1	Name2	Gene ID	Chromosome	Location	AA	Identity/%	AGI ID
SOT17	<i>BniSOT17</i>	BniB04g033130	B4 (+)	17,622,762–17,623,781	339	86.5	AT1G18590
SOT18	<i>BniSOT18-1</i>	BniB03g052540	B3 (+)	28,101,782–28,102,844	334	80.9	AT1G74090
	<i>BniSOT18-2</i>	BniB04g019280	B4 (–)	9,875,796–9,876,833	345	78.6	
	<i>BniSOT18-3</i>	BniB03g038560-1	B3 (–)	19,264,549–19,265,610	353	73.7	
	<i>BniSOT18-4</i>	BniB03g052550	B3 (+)	28,105,157–28,106,167	319	72.9	
	<i>BniSOT18-5</i>	BniB06g047660	B6 (+)	49,729,557–49,730,615	352	72.7	
	<i>BniSOT18-6</i>	BniB06g047670	B6 (+)	49,740,347–49,741,381	344	72.6	
	<i>BniSOT18-7</i>	BniB06g047650	B6 (+)	49,723,604–49,724,665	353	72.2	
	<i>BniSOT18-8</i>	BniB03g038560-2	B3 (–)	19,273,089–19,274,147	352	71.3	
	<i>BniSOT18-9</i>	BniB04g020570	B4 (–)	10,687,342–10,688,406	354	70.4	
	<i>BniSOT18-10</i>	BniB02g035710	B2 (+)	43,650,924–43,651,937	337	70.1	
	<i>BniSOT18-11</i>	BniB04g020560	B4 (+)	10,684,415–10,685,521	368	69.4	
	<i>BniSOT18-12</i>	BniB06g047680	B6 (+)	49,744,098–49,745,159	333	68.3	
	<i>BniSOT18-13</i>	BniB03g038630	B3 (–)	19,304,714–19,306,727	367	49.6	
<b>Secondary modification</b>							
<i>FMO<sub>GS-OX1</sub></i>	<i>BniFMO<sub>GS-OX1</sub></i>	BniB04g021440	B4 (+)	11,434,174–11,440,389	461	77.2	AT1G65860
<i>FMO<sub>GS-OX2</sub></i>	*						AT1G62540
<i>FMO<sub>GS-OX3</sub></i>	*						AT1G62560
<i>FMO<sub>GS-OX4</sub></i>	*						AT1G62570
<i>FMO<sub>GS-OX5</sub></i>	<i>BniFMO<sub>GS-OX5-1</sub></i>	BniB03g004970	B3 (+)	1,983,575–1,985,457	459	84.6	AT1G12140
	<i>BniFMO<sub>GS-OX5-2</sub></i>	BniB02g006840	B2 (+)	3,630,178–3,632,014	459	82.8	
<i>FMO<sub>GS-OX6</sub></i>	*						AT1G12130
<i>FMO<sub>GS-OX7</sub></i>	*						AT1G12160
CYP81F1	<i>BniCYP81F1-1</i>	BniB05g001500	B5 (–)	845,091–846,941	499	86.2	AT4G37430
	<i>BniCYP81F1-2</i>	BniB03g041020	B3 (+)	20,720,008–20,721,740	496	80.4	
	<i>BniCYP81F1-3</i>	BniB05g001510	B5 (–)	849,140–850,741	504	79.5	
CYP81F2	<i>BniCYP81F2-1</i>	BniB02g036600	B2 (+)	44,402,518–44,404,789	491	90.7	AT5G57220
	<i>BniCYP81F2-2</i>	BniB05g036790	B5 (+)	19,374,525–19,376,334	493	88.4	
	<i>BniCYP81F2-3</i>	BniB08g013930	B8 (–)	6,333,770–6,33,6220	493	87.2	
CYP81F3	<i>BniCYP81F3-1</i>	BniB05g001540	B5 (–)	866,140–869,499	497	89.3	AT4G37400
	<i>BniCYP81F3-2</i>	BniB03g014280	B3 (+)	6,049,124–6,051,489	491	86.2	
CYP81F4	<i>BniCYP81F4-1</i>	BniB05g001530	B5 (–)	859,533–861,269	501	78.7	AT4G37410
	<i>BniCYP81F4-2</i>	BniB03g014290	B3 (+)	6,062,864–6,065,223	519	76.9	
AOP1	<i>BniAOP1-1</i>	BniB08g045800	B8 (+)	28,827,457–28,828,861	320	76.5	AT4G03070
	<i>BniAOP1-2</i>	BniB08g039120	B8 (–)	23,904,179–23,911,127	423	49.2	
AOP2	<i>BniAOP2-1</i>	BniB05g055540	B5 (–)	49,961,828–49,965,927	432	59.2	AT4G03060
	<i>BniAOP2-2</i>	BniB08g045740	B8 (–)	28,785,908–28,787,937	475	51.3	

Table 2. Cont.

Name1	Name2	Gene ID	Chromosome Location		AA	Identity/%	AGI ID
AOP3	*						AT4G03050
GSL-OH	*						AT2G25450
IGMT1	<i>BniIGMT1-1</i>	BniB04g035790	B4 (−)	19,012,288–19,013,601	372	90.4	AT1G21100
	<i>BniIGMT1-2</i>	BniB03g008950	B3 (−)	3,546,629–3,547,917	372	89.1	
IGMT2	<i>BniIGMT2</i>	BniB03g008960	B3 (−)	3,550,953–3,552,247	374	92.6	AT1G21120
IGMT5	<i>BniIGMT5</i>	BniB03g055850	B3 (−)	30,178,032–3,0179,487	370	79.4	AT1G76790
<b>Co-substrate pathways</b>							
TSB1	<i>BniTSB1-1</i>	BniB02g033500	B2 (−)	42,216,546–42,218,333	472	92.6	AT5G54810
	<i>BniTSB1-2</i>	BniB05g034690	B5 (−)	17,928,767–17,930,499	309	55.2	
ASA1	<i>BniASA1-1</i>	BniB08g002350	B8 (−)	1,172,677–1,175,600	594	88.9	AT5G05730
	<i>BniASA1-2</i>	BniB02g054940	B2 (+)	53,796,911–53,799,668	604	88.2	
APK1	<i>BniAPK1-1</i>	BniB06g037540	B6 (+)	43,668,583–43,669,978	273	89.5	AT2G14750
	<i>BniAPK1-2</i>	BniB02g011750	B2 (+)	68,87,476–6,888,716	278	88.6	
APK2	<i>BniAPK2-1</i>	BniB02g088950	B2 (−)	70,492,417–70,493,784	293	90.1	AT4G39940
	<i>BniAPK2-2</i>	BniB05g000140	B5 (−)	71,545–73,229	293	89.8	
	<i>BniAPK2-3</i>	BniB03g014780	B3 (−)	6,338,724–6,339,871	227	63.9	
GSH1	<i>BniGSH1-1</i>	BniB05g014010	B5 (+)	6,646,285–6,649,224	514	93.1	AT4G23100
	<i>BniGSH1-2</i>	BniB02g080290-2	B2 (−)	66,603,961–66,606,919	520	92.2	
	<i>BniGSH1-3</i>	BniB02g080290-1	B2 (−)	66,598,228–66,600,769	469	62.6	
	<i>BniGSH1-4</i>	BniB03g021390	B3 (+)	9,550,645–9,553,182	359	61.7	
GSH2	<i>BniGSH2-1</i>	BniB07g038410	B7 (−)	45,708,701–45,711,311	531	79.8	AT5G27380
	<i>BniGSH2-2</i>	BniB02g079220	B2 (+)	66,040,689–66,048,921	478	75.1	
APS1	<i>BniAPS1-1</i>	BniB03g024880	B3 (+)	11,208,027–11,211,252	465	92.9	AT3G22890
	<i>BniAPS1-2</i>	BniB01g039550	B1 (−)	44,897,233–44,899,328	462	92.7	
APS3	<i>BniAPS3</i>	BniB05g021440	B5 (+)	10,445,377–10,447,307	470	89.9	AT4G14680
APR1	<i>BniAPR1-1</i>	BniB08g047710	B8 (−)	30,542,974–30,544,615	463	87	AT4G04610
	<i>BniAPR1-2</i>	BniB05g056120	B5 (−)	50,700,159–50,701,310	223	35.5	
APR2	*						AT1G62180
APR3	<i>BniAPR3-1</i>	BniB02g079650	B2 (−)	66,283,619–66,285,188	464	92.5	AT4G21990
	<i>BniAPR3-2</i>	BniB05g015070	B5 (+)	7,144,150–7,145,729	464	92.2	
OASA1	<i>BniOASA1-1</i>	BniB07g020470	B7 (−)	33,626,673–33,628,609	322	97.5	AT4G14880
	<i>BniOASA1-2</i>	BniB05g021220	B5 (+)	10,294,312–10,296,240	324	94.8	
CHY1	<i>BniCHY1-1</i>	BniB04g000010	B4 (−)	7994–10,682	374	87	AT5G65940
	<i>BniCHY1-2</i>	BniB01g014580	B1 (−)	7,787,511–7,790,188	380	80.7	
	<i>BniCHY1-3</i>	BniB01g014490	B1 (+)	7,718,116–7,721,800	332	73.8	
	<i>BniCHY1-4</i>	BniB06g018370	B6 (+)	10,266,194–10,271,471	370	71.8	
	<i>BniCHY1-5</i>	BniB06g018590	B6 (+)	10,442,521–10,445,218	225	39.4	
AAO4	<i>BniAAO4</i>	BniB06g046290	B6 (+)	49,137,455–49,142,867	1337	86.8	AT1G04580
BZO1	<i>BniBZO1</i>	BniB03g043330	B3 (−)	22,256,842–22,258,819	566	76.4	AT1G65880
SCPL17	<i>BniSCPL17</i>	BniB07g056480	B7 (+)	55,233,985–55,236,333	436	65.3	AT3G12203

\* means that the orthologous gene in *B. nigra* has been missing. In the chromosome location, the positive (+) and negative (−) signs indicate the existence of a gene on the positive and negative strand of that specific chromosome, respectively.

**Table 3.** The inventory of transcription factor genes and transporter genes involved in glucosinolate biosynthesis in *Brassica nigra*.

Name1	Name2	Gene ID	Chromosome	Location	AA	Identity/%	AGI ID
<b>Activator</b>							
MYB28	<i>BniMYB28-1</i>	BniB02g075850	B2 (+)	64,495,904–64,497,224	352	78.6	AT5G61420
	<i>BniMYB28-2</i>	BniB07g041200	B7 (–)	47,224,440–47,225,564	329	68.3	
	<i>BniMYB28-3</i>	BniB04g002670	B4 (–)	1,321,731–1,323,290	329	60.2	
MYB29	<i>BniMYB29-1</i>	BniB08g003370	B8 (+)	1,668,386–1,669,814	339	74.3	AT5G07690
	<i>BniMYB29-2</i>	BniB02g053290	B2 (–)	53,047,760–53,050,344	298	46.2	
MYB76	*						AT5G07700
MYB34	<i>BniMYB34-1</i>	BniB02g076220	B2 (–)	64,684,279–64,685,680	313	75.1	AT5G60890
	<i>BniMYB34-2</i>	BniB04g003010	B4 (–)	1,462,793–1,463,979	276	70.4	
	<i>BniMYB34-3</i>	BniB07g040970	B7 (–)	47,066,742–47,067,688	262	60.7	
MYB51	<i>BniMYB51-1</i>	BniB03g028100	B3 (+)	12,800,136–12,801,361	326	75.4	AT1G18570
	<i>BniMYB51-2</i>	BniB03g007660	B3 (+)	2,995,206–2,996,724	319	72.1	
	<i>BniMYB51-3</i>	BniB04g033080	B4 (+)	17,574,440–17,575,684	323	71.7	
MYB115	*						AT5G40360
MYB118	<i>BniMYB118-1</i>	BniB04g007120	B4 (+)	3,444,507–3,446,768	486	63	AT3G27785
	<i>BniMYB118-2</i>	BniB02g067910	B2 (–)	60,652,891–60,655,308	477	63	
MYB122	<i>BniMYB122-1</i>	BniB03g052520	B3 (+)	28,090,119–28,091,605	335	74.9	AT1G74080
	<i>BniMYB122-2</i>	BniB05g070970	B5 (+)	64,915,053–64,916,920	376	67.8	
MYC2	<i>BniMYC2-1</i>	BniB04g049150	B4 (–)	30,208,579–30,210,393	604	87.2	AT1G32640
	<i>BniMYC2-2</i>	BniB03g035750	B3 (–)	17,597,458–17,599,278	606	86.6	
	<i>BniMYC2-3</i>	BniB03g016310	B3 (+)	7,033,365–7,035,185	606	84.3	
MYC3	<i>BniMYC3-1</i>	BniB02g063410	B2 (+)	58,260,877–58,262,631	584	76.9	AT5G46760
	<i>BniMYC3-2</i>	BniB04g015560	B4 (–)	7,676,518–7,678,221	567	71.5	
MYC4	<i>BniMYC4</i>	BniB05g018490	B5 (+)	8,814,854–8,816,635	593	73.1	AT4G17880
IQD1	<i>BniIQD1-1</i>	BniB07g058100	B7 (+)	55,932,859–55,934,585	444	67.2	AT3G09710
	<i>BniIQD1-2</i>	BniB03g023450	B3 (–)	10,613,021–10,615,049	485	59.9	
SLIM1	<i>BniSLIM1-1</i>	BniB03g052120	B3 (–)	27,906,056–27,908,203	587	85.1	AT1G73730
	<i>BniSLIM1-2</i>	BniB03g038740	B3 (+)	19,356,898–19,358,816	584	82.1	
	<i>BniSLIM1-3</i>	BniB05g070840	B5 (–)	64,849,951–64,851,819	538	77.1	
OBP2	<i>BniOBP2-1</i>	BniB03g002340	B3 (–)	936,743–938,038	340	76.4	AT1G07640
	<i>BniOBP2-2</i>	BniB06g043020	B6 (+)	47,097,586–47,098,898	334	75.8	
	<i>BniOBP2-3</i>	BniB02g003700	B2 (–)	1,819,220–1,820,440	313	69.5	
CAMTA3	<i>BniCAMTA3-1</i>	BniB01g021790	B1 (–)	13,029,617–13,034,577	1034	86.2	AT2G22300
	<i>BniCAMTA3-2</i>	BniB03g062790	B3 (–)	45,548,997–45,550,544	429	34.9	
	<i>BniCAMTA3-3</i>	BniB02g018410	B2 (+)	11,860,712–11,862,740	358	27	
CCA1	<i>BniCCA1</i>	BniS02554g500	utg2554 (–)	328,706–331,283	572	74.4	AT2G46830
<b>Activator/Repressor</b>							
HY5	<i>BniHY5-1</i>	BniB05g026170	B5 (+)	12,836,342–12,837,354	163	91.7	AT5G11260
	<i>BniHY5-2</i>	BniB02g050820	B2 (+)	51,949,284–51,950,286	168	91.1	
	<i>BniHY5-3</i>	BniB06g056490	B6 (–)	54,414,011–54,416,320	208	57.8	
<b>Repressor</b>							
SD1	<i>BniSD1-1</i>	BniB02g069920	B2 (–)	61,609,407–61,611,242	306	92.5	AT5G48850
	<i>BniSD1-2</i>	BniB07g037550	B7 (–)	45,221,658–45,223,201	306	90.5	
	<i>BniSD1-3</i>	BniB04g005760	B4 (+)	2,848,170–2,850,064	299	88.9	
	<i>BniSD1-4</i>	BniB07g037520	B7 (–)	45,208,991–45,210,584	306	88.2	
SD2	<i>BniSD2-1</i>	BniB06g046090	B6 (+)	49,043,247–49,044,406	303	88.2	AT1G04770
	<i>BniSD2-2</i>	BniB03g001070	B3 (–)	446,167–447,323	299	83.8	
FRS7	<i>BniFRS7</i>	BniB01g057400	B1 (+)	55,333,501–55,335,611	534	51.1	AT3G06250
FRS12	<i>BniFRS12</i>	BniB08g009620	B8 (+)	4,493,856–4,496,198	780	86.4	AT5G18960

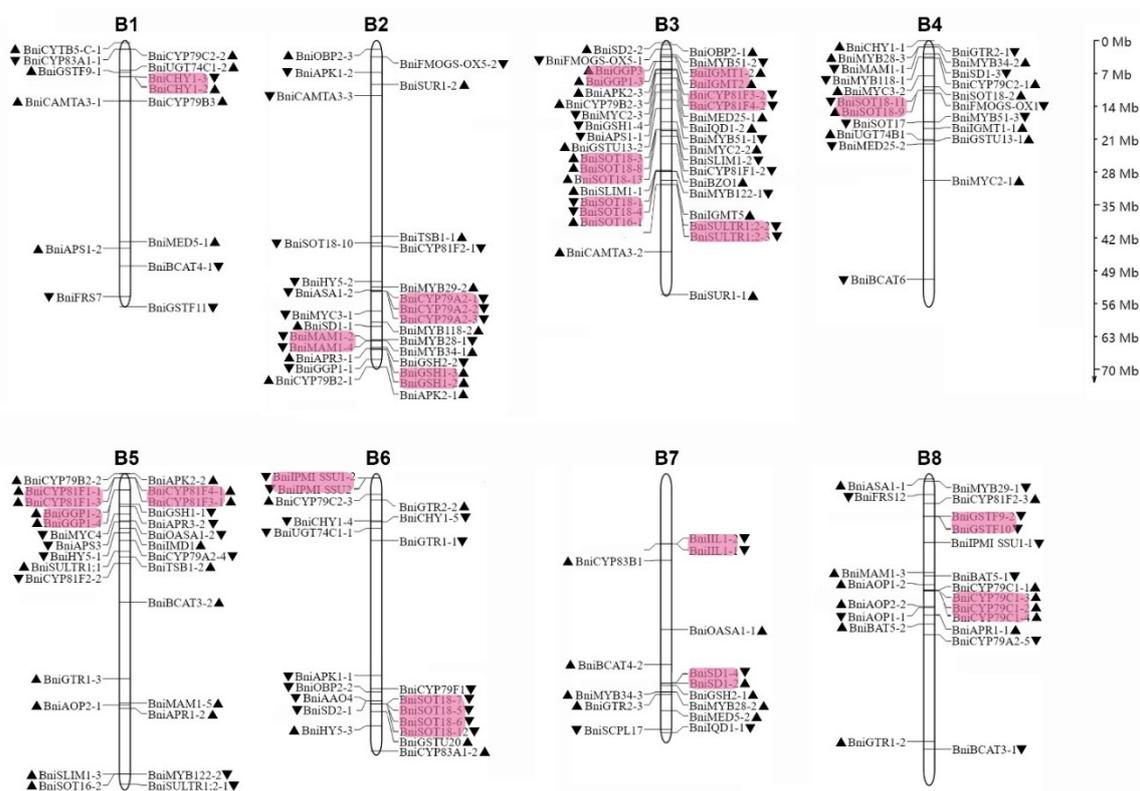
Table 3. Cont.

Name1	Name2	Gene ID	Chromosome Location		AA	Identity/%	AGI ID
<b>Mediator</b>							
MED5	<i>BniMED5-1</i>	BniB01g038210	B1 (–)	43,391,779–43,397,254	1308	89.5	AT3G23590
	<i>BniMED5-2</i>	BniB07g048400	B7 (–)	51,150,764–51,156,016	1297	87.3	
MED25	<i>BniMED25-1</i>	BniB03g018400	B3 (–)	8,065,052–8,070,109	830	87.9	AT1G25540
	<i>BniMED25-2</i>	BniB04g040670	B4 (+)	22,229,344–22,231,613	396	23.7	
<b>Transporter</b>							
BAT5	<i>BniBAT5-1</i>	BniB08g037040	B8 (+)	22,004,327–22,006,322	408	90	AT4G12030
	<i>BniBAT5-2</i>	BniB08g049400	B8 (–)	32,467,906–32,474,047	393	79.1	
SULTR1;1	<i>BniSULTR1;1</i>	BniB05g033180	B5 (–)	16,848,645–16,851,311	519	69.8	AT4G08620
SULTR1;2	<i>BniSULTR1;2-1</i>	BniB05g074030	B5 (+)	67,138,702–67,142,049	652	93.7	AT1G78000
	<i>BniSULTR1;2-2</i>	BniB03g057280	B3 (+)	31,113,533–31,116,846	655	92.7	
	<i>BniSULTR1;2-3</i>	BniB03g057270	B3 (+)	31,102,957–31,106,559	671	91.8	
GTR1	<i>BniGTR1-1</i>	BniB06g024230	B6 (+)	14,337,898–14,340,309	634	83.4	AT3G47960
	<i>BniGTR1-2</i>	BniB08g061180	B8 (–)	57,886,043–57,888,835	615	80.7	
	<i>BniGTR1-3</i>	BniB05g052360	B5 (–)	44,394,313–44,396,700	617	78	
GTR2	<i>BniGTR2-1</i>	BniB04g002130	B4 (+)	1,022,088–1,024,566	612	92	AT5G62680
	<i>BniGTR2-2</i>	BniB06g010940	B6 (–)	5,574,061–5,576,460	612	91.6	
	<i>BniGTR2-3</i>	BniB07g041760	B7 (–)	47,577,010–47,579,359	606	85.2	

\* means that the orthologous gene in *B. nigra* has been missing. In the chromosome location, the positive (+) and negative (–) signs indicate the existence of a gene on the positive and negative strand of that specific chromosome, respectively.

Of the 184 *BniGSL* genes, 182 are unevenly mapped among the eight chromosomes of *B. nigra*, with 14, 28, 36, 21, 28, 20, 14, and 21 *BniGSL* genes anchoring on chromosome B1–B8, respectively (Figure 1). Two other *BniGSL* genes are distributed on a large scaffold, which have not yet been mapped onto chromosomes (Tables 2 and 3). In *A. thaliana*, there are 25 genes that constitute 11 tandem-duplicate gene modules. Here, we found that there are 20 tandem-duplicate gene modules in *B. nigra* involving 46 genes. However, the gene families involved in tandem duplication are not exactly the same between these two species (Table 2 and Figure 1).

By doing pairwise sequence alignment of *GSL* genes between *B. nigra* and *A. thaliana*, we found that a total of 13 homologs corresponding to *A. thaliana* *GSL* genes were absent in *B. nigra*, including two genes involved in side-chain elongation (i.e., *MAM3* and *IPMI-SSU3*), a gene related to core structure synthesis (i.e., *CYP79F2*), seven genes involved in side-chain modification (i.e., *FMO<sub>GS-OX2/3/4/6/7</sub>*, *AOP3*, and *GSL-OH*), a gene related to co-substrate pathways (i.e., *APR2*), and two TFs (i.e., *MYB76* and *MYB115*) (Table 2). In general, 95.7% of the *BniGSL* genes identified in this study shared 51–97% amino acid sequence identity with *AtGSL* genes in *A. thaliana*, with an average of 78.4%.



**Figure 1.** Genomic distribution of glucosinolate genes (*BniGSLs*) on the chromosomes of *Brassica nigra*. The arrowheads next to gene names show the direction of transcription. The chromosome numbers are demonstrated at the top of each chromosome. Tandem repeat genes are marked with a pink background.

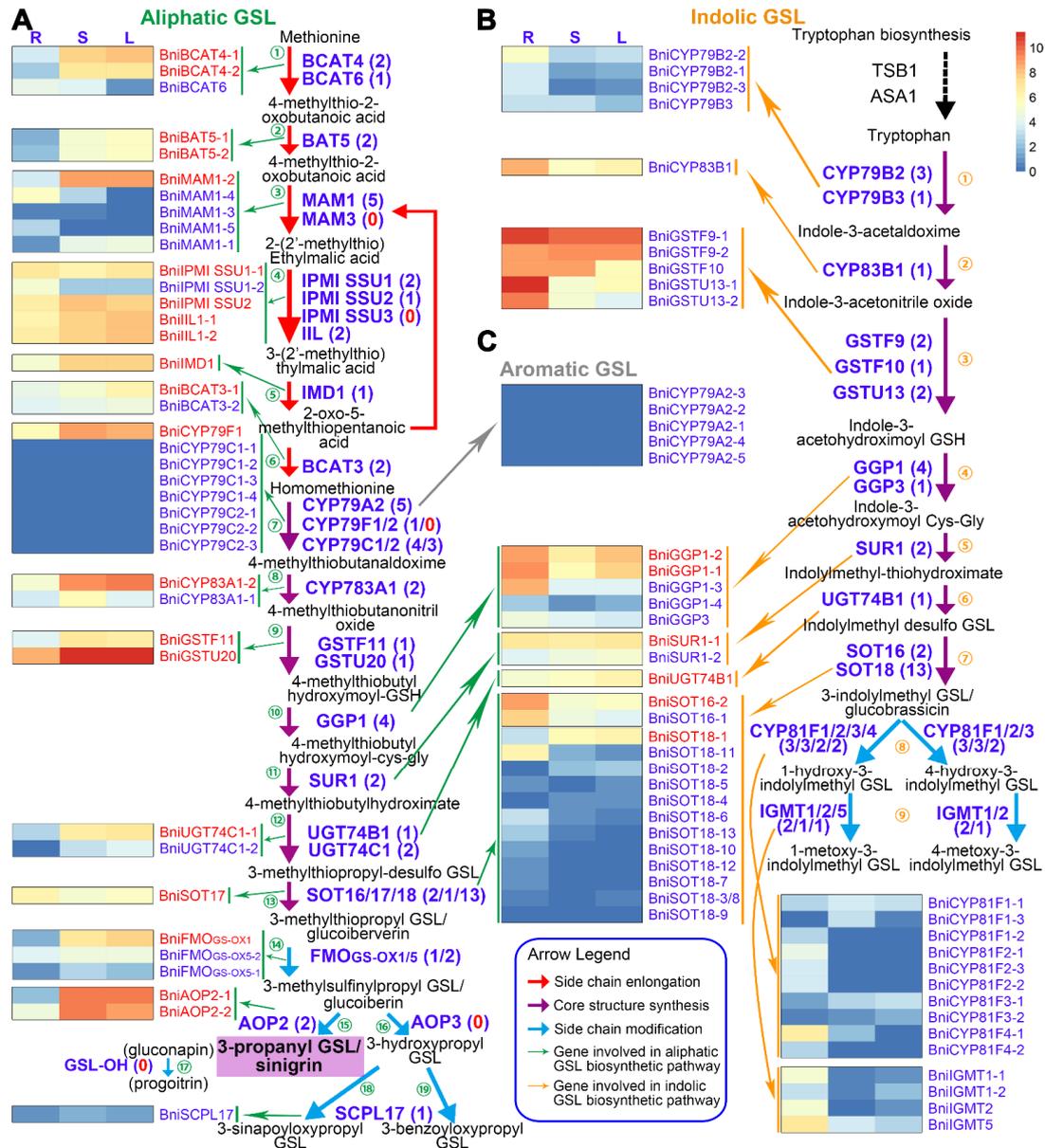
### 3.3. Expression Patterns of *BniGSL* Genes Encoding Enzymes in Three Organs of *B. nigra*

Transcriptome sequencing was conducted to investigate the expression of *BniGSL* genes involved in different processes of GSL biosynthesis, and those that participated in aliphatic, indolic, and aromatic GSLs in three organs (i.e., root, stem, and leaf). First, the correlation analysis between biological replicates was performed, and the results showed that the three biological replicates of root, stem, and leaf all have a good correlation (Figure S2), indicating the high data reliability in this study.

Among 184 *BniGSL* genes determined in this study, 172 were detected in at least one organ, and 12 genes were not expressed in all three organs. Moreover, there were 34 *BniGSL* genes whose expression in roots, stems, and leaves all lower than 1 FPKM (Table S4). The low expression of these genes suggests that they may contribute less to GSL biosynthesis in these three organs.

On the basis of the GSL biosynthetic pathways involving *BniGSL* genes and their biological functions, the expression of 184 *BniGSL* genes in roots, stems, and leaves was further analyzed. The biosynthesis of aliphatic GSLs can be divided into three main phases, and dozens of enzymes have been determined to participate in the corresponding reactions (Figure 2). The side-chain elongation phase of aliphatic GSLs consists of six steps and eleven enzymes, of which MAM3 and IPMI SSU3 have been lost in *B. nigra*. Seven enzymatic steps of the core structure synthesis involved 14 enzymes, with 38 *BniGSL* genes encoding 13 enzymes in *B. nigra* (except CYP79F2). The aliphatic GSL side-chain modification phase mainly includes S-oxygenation, side-chain oxygenation, and further conversion of hydroxyalkyl GSL into benzoylated and sinapoylated GSLs, which are catalyzed by five types of enzymes. However, both AOP3 and *GSL-OH* are absent in *B. nigra*, and there are only two of the seven Arabidopsis *FMOGS-OX* are retained in *B. nigra*. Encouragingly, although there are some aliphatic GSL synthetic *BniGSL* genes that had an extremely low expression, expression data showed that at least one *BniGSL* gene was highly expressed in

every step of the synthesis of sinigrin. Moreover, the expression of these highly expressed *BniGSL* genes in leaves and stems was generally higher than that in roots (Figure 2A), and this finding was consistent with the fact that the side-chain elongation of aliphatic GSLs should be performed in green tissues containing chloroplasts.



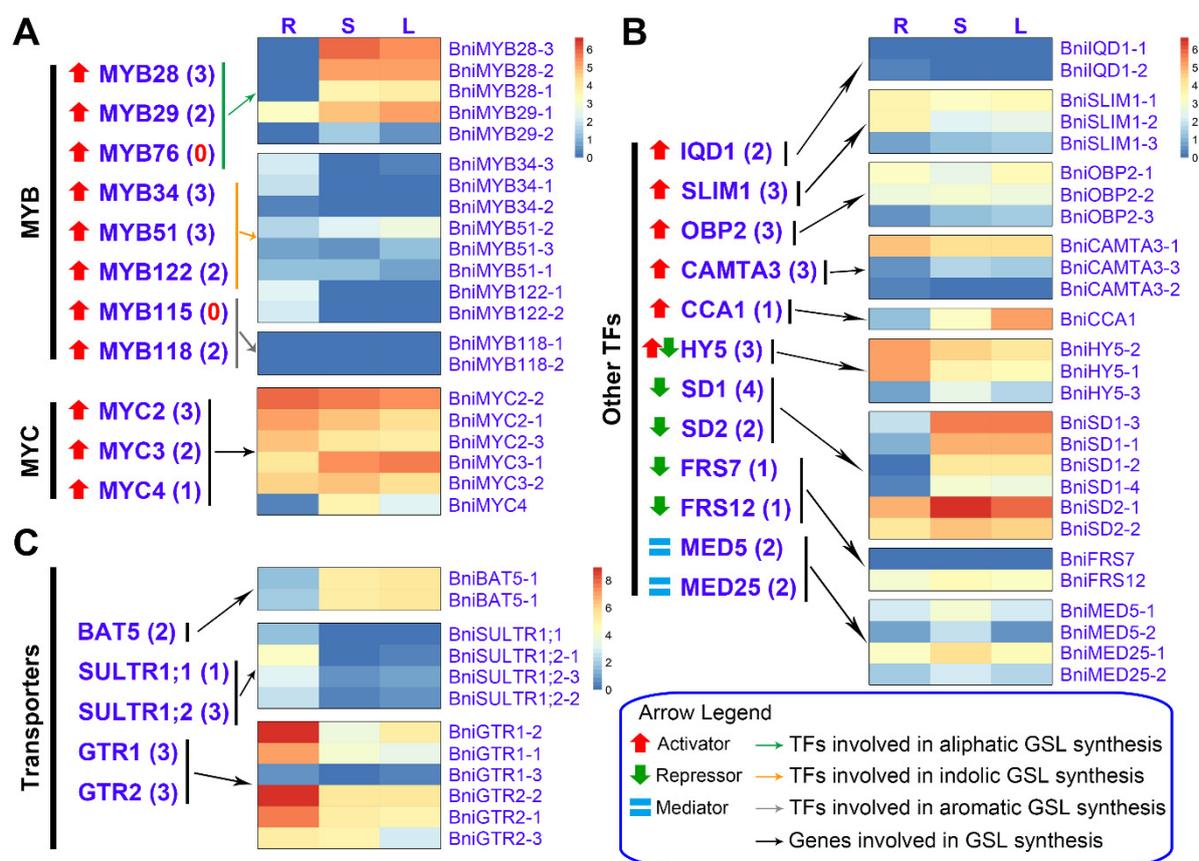
**Figure 2.** Biosynthetic pathways of aliphatic and indolic glucosinolates (GSLs) and the heatmap of related *BniGSLs* in *Brassica nigra*. The pathway contains 3 major phases: side-chain elongation (Steps 1–6), core structure synthesis (steps 7–13 in (A) and steps 1–7 in (B)), and side-chain modification (steps 14–19 in (A) and steps 8 and 9 in (B)). The biosynthetic pathway of aliphatic (A), indolic (B), and aromatic (C) GSLs and the heatmap of related *BniGSLs*. The number in parentheses represents the copy number of the gene. Abbreviations: AOP2, 2-oxoglutarate-dependent dioxygenase; BAT5, probable sodium/metabolite cotransporter BASS5; BCAT, branched-chain amino acid aminotransferase; CYP79A, phenylalanine N-monooxygenase; CYP79B, tryptophan N-monooxygenase; CYP79C/F, cytochrome P450 79C/F; CYP83B, CYP83B monooxygenase; FMO<sub>GS-OX</sub>, flavincontaining monooxygenase; GGP1,  $\gamma$ -glutamyl peptidase 1; GSTF/U, glutathione S-transferase F/U; GTR, GSL transporter; IGMT, indole GSL O-methyltransferase; IIL1, isopropylmalate isomerase large subunit 1; IPMI-SSU, isopropylmalate isomerase small subunit; L, leaf; MAM, methylthioalkylmalate synthase; R, root; S, stem; SOT, sulfotransferase. The framework of the aliphatic and indolic biosynthetic pathways is adapted from ref [11].

The synthesis of indolic GSL does not require side-chain elongation and only consists of core structure synthesis and side-chain modification, which can be divided into seven and two steps, respectively. All enzymes involved in indolic GSL synthesis have corresponding homologs in *B. nigra*, and 33 and 14 *BniGSL* genes encode 12 and 7 enzymes, respectively, to take part in the above two phases. Remarkably, four *BniGSL* genes encoding CYP79B2 and CYP79B3, the indolic cytochrome P450 enzymes that convert tryptophan derivatives into aldoximes, and all indolic *BniGSL* genes responsible for the side-chain modification, showed generally low expression, especially in stems and leaves (Figure 2B). In addition, 12 *BniGSL* genes encoding three cytochrome P450 members (i.e., CYP79A2, CYP79C1 and CYP79C2) were involved in the aromatic core structure GSL pathway but were almost not expressed in roots, stems, and leaves (Figure 2C). This finding might explain the few aromatic GSLs in *B. nigra*, and their content was extremely low.

#### 3.4. Expression Patterns of *BniGSL* Genes Encoding TFs, Transporters and Proteins Involved in Co-Substrate Pathways

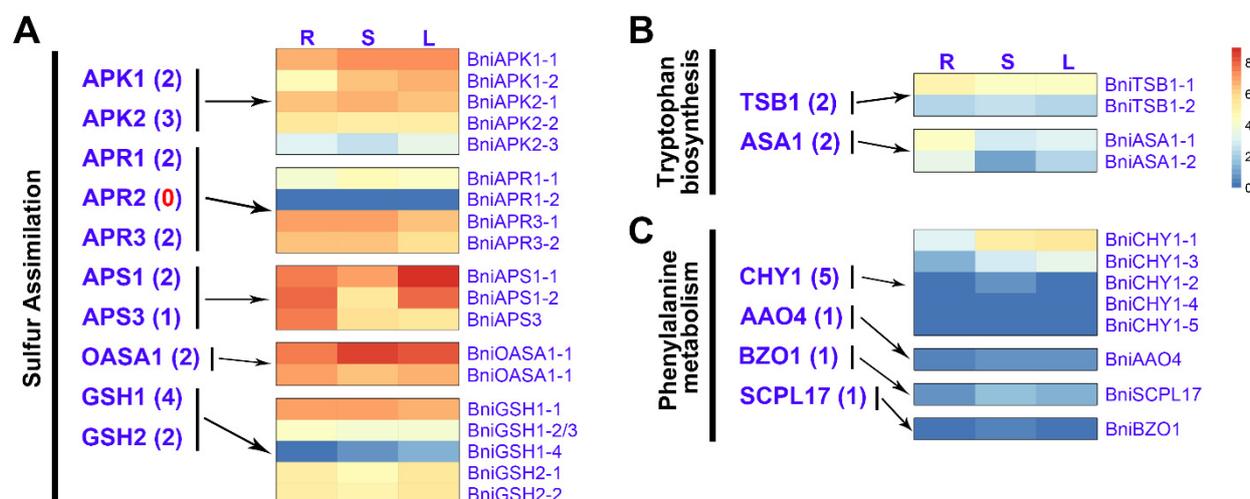
Thus far, a total of 10 TFs of the R2R3 domain MYB family have been characterized as key players in the regulation of GSL genes. MYB28, MYB29, and MYB76 play critical regulatory roles in aliphatic GSL biosynthesis. MYB34, MYB51, and MYB122 are essential regulators of indolic GSL biosynthesis. MYB115 and MYB118 are responsible for modulating the synthesis of aromatic GSLs. Similar to the enzymes involved in GSL biosynthesis, multiple aliphatic GSL-related *BniMYBs* showed high expression in stems and leaves, whereas all indolic and aromatic GSL-related *BniMYBs* had low expressions (Figure 3A). As expected, MYC2, MYC3, and MYC4, the common regulators of aliphatic and indolic GSL synthesis, were highly expressed in all three organs of *B. nigra* (Figure 3A), and had similar expression patterns of some *BniGSL* genes (i.e., *BniGGP1-1/2*, *BniSUR1-1*, *BniUGT74B1*, *BniSOT16-2*, and *BniSOT18-1*) involved in the core structure synthesis of aliphatic and indolic GSLs (Figure 2). Among other TFs that regulated GSL biosynthesis, except *IQD1* and *FRS 7*, at least one copy of a *BniGSL* gene was expressed in at least one organ (Figure 3B).

Five proteins have been experimentally characterized as transporters in GSL biosynthesis. BAT5s, which acts as chloroplast transporter in the side-chain elongation phase, exhibited a similar expression pattern to the GSL genes related to aliphatic GSL biosynthesis in *B. nigra*. SULTR1;1 and SULTR1;2 are two sulfate transporters that function in *Arabidopsis* roots, and their orthologs in *B. nigra* also showed a relatively high expression in roots. GTR1 and GTR2 are responsible for transporting synthesized GSLs from leaves to seeds and roots [50,51]. Expression data also showed that *BniGTR1s* and *BniGTR2s* were predominantly expressed in roots (Figure 3C).



**Figure 3.** Heatmap of transcription factor (TFs) genes and transporter genes involved in GSL biosynthesis. Expression analysis of *BniMYBs* (A) and other TFs (B) involved in GSL biosynthesis. (C) Expression analysis of *BniGSLs* encoding transporters involved in GSL transportation. The number in parentheses represents the copy number of the gene. Abbreviations: BAT5, probable sodium/metabolite cotransporter BASS5; CAMTA3, calmodulin-binding transcription activator 3; CCA1, circadian clock-associated 1; FRS, Far1 related sequence; GTR, GSL transporter; HY5, long hypocotyl 5; IQD1, IQ domain 1; L, leaf; MED, mediator subunit; OBP2, UAS-tagged root patterning 3; R, root; S, stem; SD, protein sulfur deficiency-induced; SLIM1, sulfur limitation 1; SULTR, sulfate transporter.

The gene expression analysis showed that the expression patterns of the *BniGSL* genes encoding proteins involved in the co-substrate pathway were also different. For *BniGSL* genes that function in the sulfur assimilation process, except *BniAPK2-3*, *BniAPR1-2*, and *BniGSH1-4*, the remaining 17 genes were highly expressed in all three organs (Figure 4A), which were similar to the expression patterns of some core structure synthesis related *BniGSL* genes that play roles in aliphatic and indolic GSL synthesis (Figure 2). However, *BniGSL* genes involved in the synthesis of tryptophan and phenylalanine metabolism were all generally low-expressed in the three organs (Figure 4B,C), which indicated that the supplies of tryptophan for the synthesis of indolic GSL and benzoyl-coenzyme A (BzCoA) for the synthesis of benzoylated GSLs (BzGSLs) might not be sufficient. The low expression characteristics of these genes may be one of the reasons for the low indolic GSL content in *B. nigra* and the absence of BzGSLs and sinapoylated GSLs (SnGSLs).

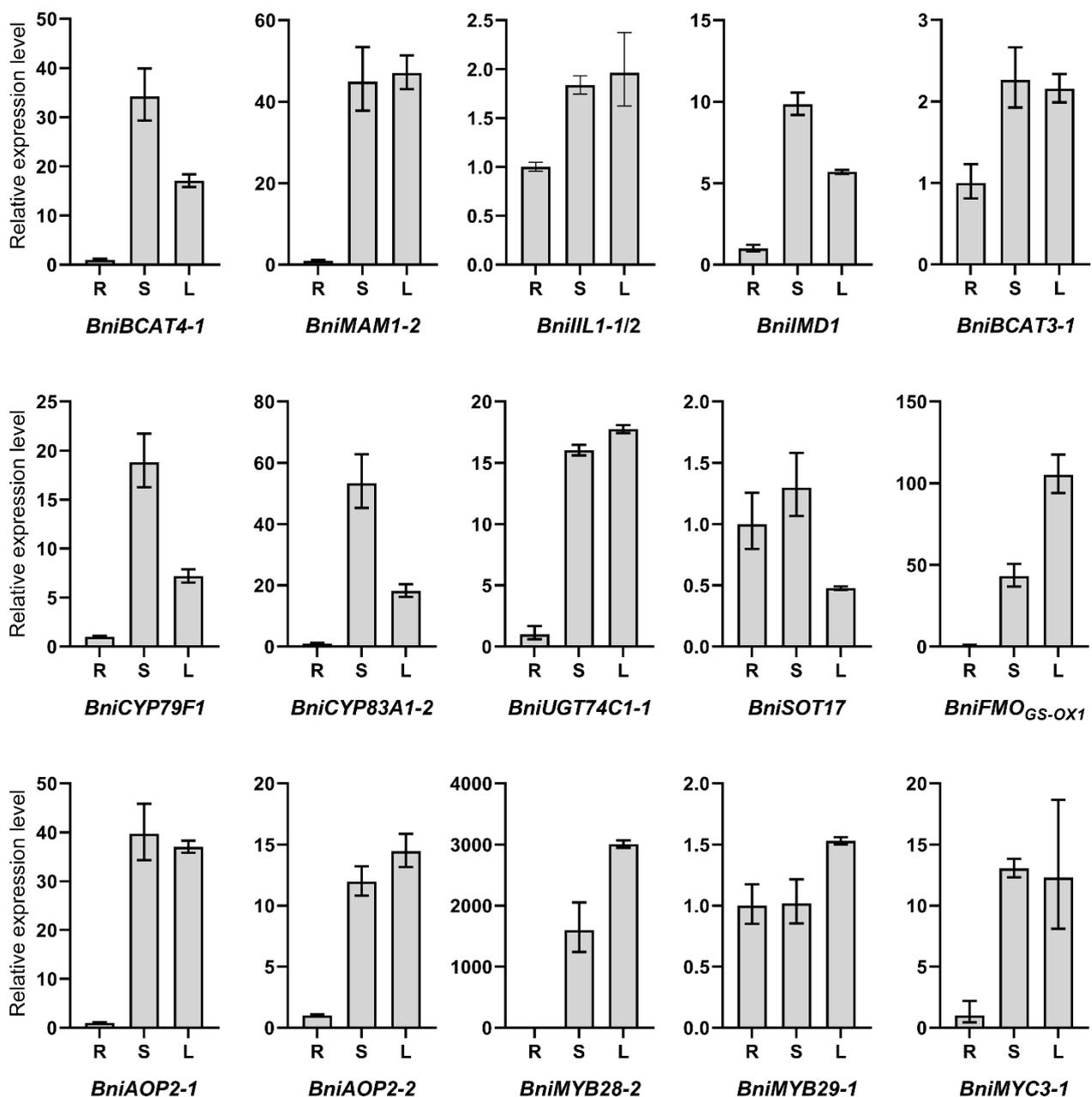


**Figure 4.** Heatmap of genes involved in the co-substrate pathways of GSL biosynthesis. Expression analysis of *BniGSLs* involved in sulfur assimilation (A), tryptophan biosynthesis (B), and phenylalanine metabolism (C). The number in parentheses represents the copy number of the gene. Abbreviations: AAO4, aldehyde oxidase 4; APK, adenylyl-sulfate kinase; APR, 5'-adenylylsulfate reductase; APS, ATP sulfurylase; BZO1, benzoyloxy GSL 1; ASA1, anthranilate synthase  $\alpha$  subunit 1; CHY1, 3-hydroxyisobutyryl-CoA hydrolase 1; GSH1, glutamate–cysteine ligase; GSH2, glutathione synthetase; L, leaf; R, root; S, stem; SCPL17, serine carboxypeptidase-like 17; TSB1, tryptophan synthase  $\beta$  chain 1.

### 3.5. Expression Patterns of Candidate Key Genes Involved in the Synthesis of Aliphatic GSLs by qRT-PCR

qRT-PCR analysis was performed to reconfirm the expression patterns of 15 candidate key genes (including 12 structural genes and three TFs) involved in the biosynthesis of aliphatic GSLs. As shown in Figure 5, the results of qRT-PCR indicated that the expression levels of nine genes (i.e., *BniBCAT4-1*, *BniMAM1-2*, *BniIMD1*, *BniCYP79F1*, *BniCYP83A1-2*, *BniUGT74C1-1*, *BniFMO<sub>GSL-OX1</sub>*, *BniAOP2-1*, and *BniAOP2-2*) out of 12 structural genes in stems and leaves were five times or more than that in roots. The expression levels of the other three structural genes in three organs did not differ by more than 2.5 times. Nevertheless, in addition to *BniSOT17*, the expression levels of *BniILL1-1/2* and *BniBCAT3-1* in stems and leaves were still higher than those in roots. For TFs, notably, as an orthologous gene of *MYB28*, which is the main regulatory gene of aliphatic GSL synthesis, *BniMYB28-2* was extremely highly expressed in stems and leaves. The expression of *BniMYC3-1* in stems and leaves was also higher than that in roots. In contrast, the expression of *BniMYB29-1* in three organs was not much different.

In short, the results of qRT-PCR were consistent with the results of RNA-Seq (Figures 3 and 4, Table S4), which together indicated that most of the candidate key genes involved in the synthesis of aliphatic GSLs in *B. nigra* were mainly highly expressed in stems and leaves, while relatively low in roots.

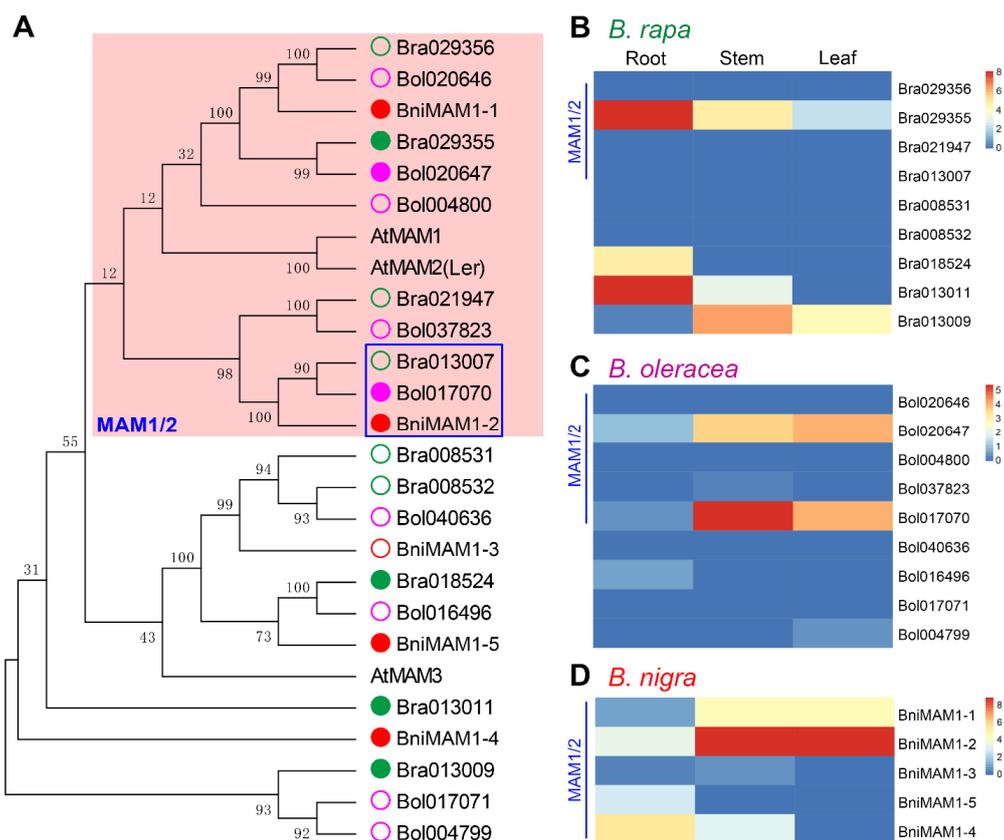


**Figure 5.** The qRT-PCR analysis of relative expression levels of 15 *BniGSL* genes in different organs of flowering plants. R, root; S, stem; L, leaf. *TIPS* was used as a reference gene, and the expression of *BniGSL* genes in root was set as 1. Error bars represent positive and negative deviations from three independent biological replicates.

### 3.6. Key MAM Genes Controlling Side-Chain Elongation in *B. nigra*

The aliphatic GSL biosynthesis initiated by methionine first needs to undergo a six-step side-chain elongation involving five types of enzymes. Among these enzymes, the MAM family remarkably contributes to the diversity of synthesized GSLs because different MAM members have different preferences for catalyzing the side-chain elongation process. In *Arabidopsis*, three tandemly duplicated MAM genes were identified, named *MAM1*, *MAM2* (absent in ecotype Columbia), and *MAM3*. The functional analysis revealed that *MAM2* and *MAM1* were correlated with the accumulation of 3 and 4 carbon (C) side-chain GSLs that had undergone the first and first two rounds of chain elongation, respectively [52,53], whereas *MAM3* was able to catalyze the condensations in the first six elongation cycles [13].

To clarify the evolutionary relationships among MAM homologs, we performed a detailed phylogenetic analysis of predicted amino acid sequences of MAM members from *Arabidopsis* and three basic diploid species of *Brassica*. On the basis of syntenic and sequence similarity analysis, we identified nine, nine, and five MAM members from *B. rapa*, *B. oleracea*, and *B. nigra*, respectively. The resulting phylogenetic tree indicated that four, five, and two MAM members of the three basic diploid species of *Brassica* were phylogenetically close to AtMAM1/2, and three, two, and two MAM members seemed to be closely related to AtMAM3. In addition, there were two BraMAM, two BolMAM, and a BniMAM that were phylogenetically distant from AtMAM1/2 and AtMAM3 (Figure 6A).



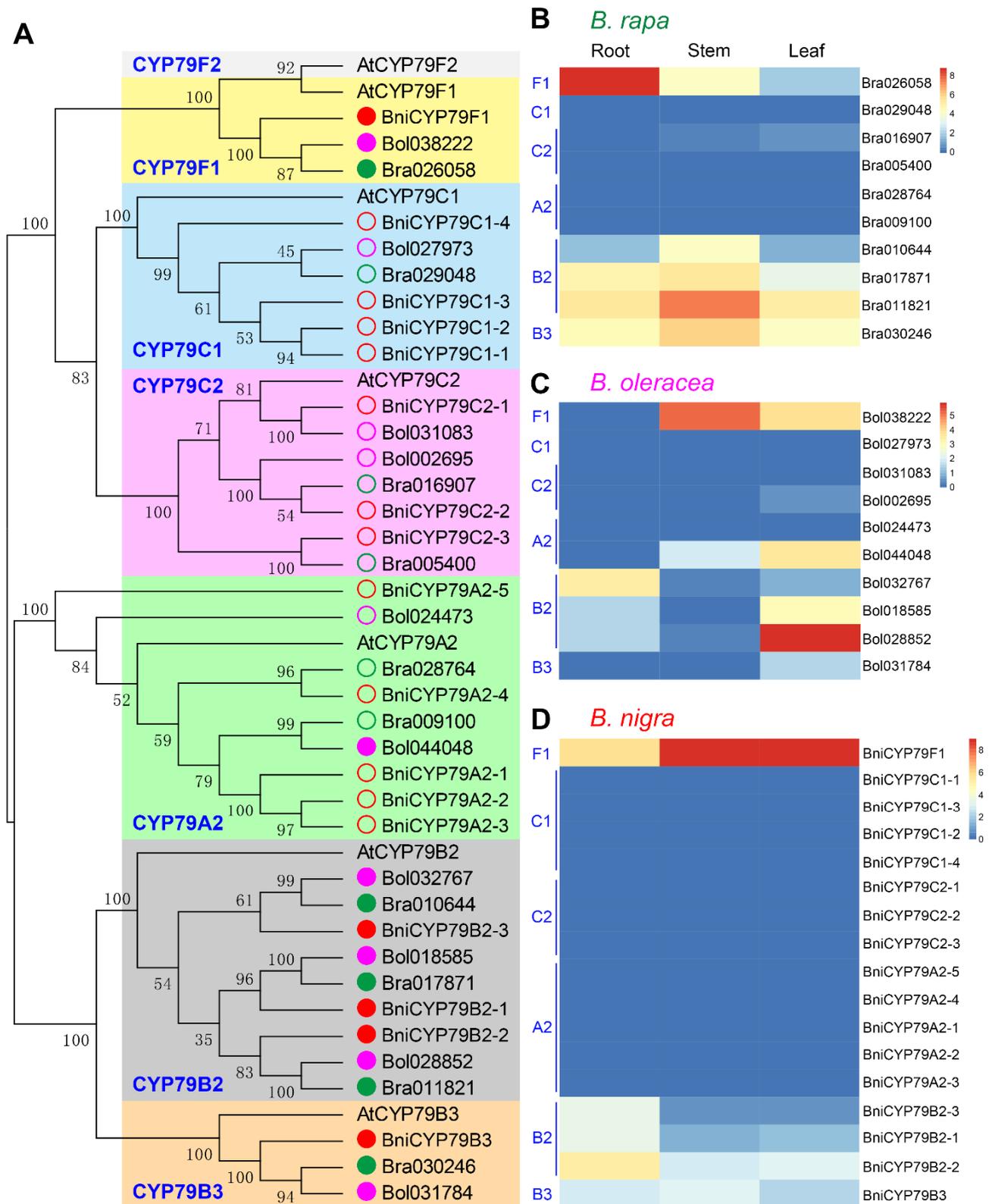
**Figure 6.** The phylogenetic tree and heatmap of MAM genes in *Brassica rapa*, *Brassica oleracea*, *Brassica nigra*, and *Arabidopsis thaliana*. (A) The phylogenetic tree of MAM genes. The MAM genes identified from *Brassica rapa*, *Brassica oleracea*, and *Brassica nigra* are indicated by green, purple, and red hollow circles or solid discs. The silenced genes are represented by a hollow circle, expressed functional genes are indicated by a solid disc. The heatmap of MAM genes in *Brassica rapa* (B), *Brassica oleracea* (C), and *Brassica nigra* (D). The expression data of *B. rapa* and *B. oleracea* were obtained from ref [48,49].

On the basis of expression data of the *BniGSL* genes obtained by RNA sequencing in this study, as well as the reported expression information of *GSL* genes in *B. rapa* and *B. oleracea* [48,49], we found that only four, two, and four MAM genes were expressed in *B. rapa*, *B. oleracea*, and *B. nigra*, respectively. Interestingly, a group of genes comprising Bra013007, Bol017070, and BniMAM1-2, was the only group of orthologs that showed high expression in *B. oleracea* and *B. nigra* but silenced in *B. rapa* (Figure 6B–D). This expression difference most likely explains why 3C and/or 4C *GSLs* can be synthesized and accumulated in large quantities in *B. nigra* (e.g., sinigrin) and *B. oleracea* (e.g., progoitrin, gluconapin, glucoraphanin, and sinigrin), but not in *B. rapa*. In addition, Bol020647 and BniMAM1-1 may also contribute to the synthesis of 3C *GSLs* (Figure 6C,D). Meanwhile, it is likely the loss of MAM3 leads to the low contents of 4C, 5C and long-chained aliphatic *GSLs* in *B. nigra* (Figure 6D).

### 3.7. *BniCYP79F1* Was Extremely Highly Expressed in *B. nigra*

The substrate-specific cytochrome P450s of the CYP79 family act as the entry point in GSL core structure synthesis by catalyzing the conversion of amino acid derivatives into the corresponding aldoximes. In *Arabidopsis*, seven CYP79s have been functionally characterized, and different members showed different preferences for amino acid-derived substrates. To reveal the possible connection between the expression of CYP79s and the GSL profiles, CYP79s involved in GSL synthesis in *B. rapa*, *B. oleracea*, and *B. nigra* were further analyzed (Figure 7). CYP79F1 and CYP79F2 are responsible for the biosynthesis of methionine-derived GSLs, and CYP79F2 only converts long-chained methionine derivatives into aldoximes [54–56]. Results showed that although there is only one ortholog of CYP79F1 in each *Brassica* species, they were all highly expressed in at least one organ. The only difference was that Bra026058 was predominantly expressed in roots, Bol038222 was mainly expressed in stems, followed by leaves, while *BniCYP79F1* was extremely highly expressed in stems and leaves, while slightly lower in roots (Figure 7B–D). Surprisingly, there is no ortholog of CYP79F2 in *Brassica* (Figure 7A). This fact may also connect to the low content of long-chained aliphatic GSLs in the GSL profiles of *Brassica*.

CYP79B2 and CYP79B3 take part in indolic GSL synthesis [57–59]. Genomic analysis showed that CYP79B2 retained three copies in each of the three *Brassica* species, while CYP79B3 retained only one copy (Figure 7). Moreover, these 12 CYP79s have been detected to be expressed in at least one organ. However, the expression patterns of CYP79B2 and CYP79B3 in the three species were different. In general, CYP79B2 and CYP79B3 were highly expressed in all three organs of *B. rapa*, while those in *B. oleracea* were predominantly expressed in leaves, followed by in roots. In *B. nigra*, their expression can be detected in roots but lower than those in stems and leaves (Figure 7B–D). These expression features may partly explain that the content of indolic GSL in *B. nigra* is lower than those in *B. rapa* and *B. oleracea*. In addition, homologous genes of CYP79A2, CYP79C1, and CYP79C2 in *B. nigra* are more than those in *B. rapa* and *B. oleracea*. However, except for one copy of CYP79A2 in *B. oleracea*, the expression of all homologs of these three genes in all three species was extremely low.

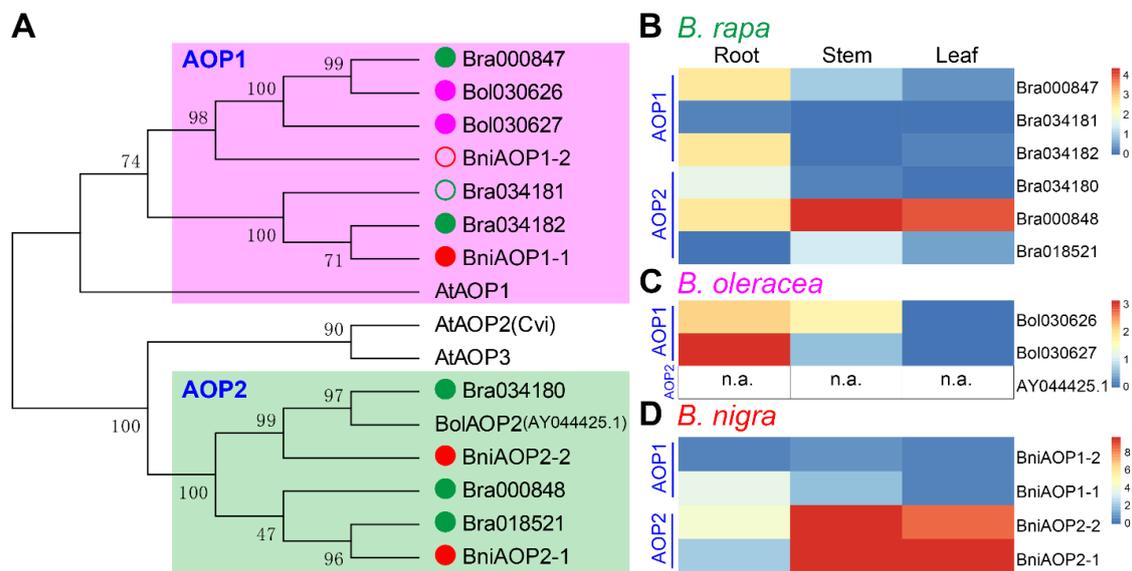


**Figure 7.** The phylogenetic tree and heatmap of CYP79 genes in *Brassica rapa*, *Brassica oleracea*, *Brassica nigra*, and *Arabidopsis thaliana*. (A), The phylogenetic tree of CYP79 genes. The CYP79 genes identified from *Brassica rapa*, *Brassica oleracea*, and *Brassica nigra* are indicated by green, purple, and red hollow circles or solid discs. The silenced genes are represented by a hollow circle, expressed functional genes are indicated by a solid disc. The heatmap of CYP79 genes in *Brassica rapa* (B), *Brassica oleracea* (C), and *Brassica nigra* (D). The expression data of *B. rapa* and *B. oleracea* were obtained from ref [48,49].

### 3.8. The Difference in AOP2 Genes Greatly Affect the Diversity of GSLs in Brassica

Side-chain modification is another pathway to enrich aliphatic GSL species in addition to the side-chain elongation. The *GS-AOP* locus is responsible for side-chain oxygenation and contains three genes encoding 2-oxoglutarate-dependent dioxygenases in Arabidopsis, namely *AOP1*, *AOP2*, and *AOP3* [60]. *AOP2* and *AOP3* are located within the *GSL-ALK* and *GSL-OHP* loci, respectively, and can convert methylsulfinylalkyl GSL into alkenyl and hydroxyalkyl GSL, respectively [23,61]. Although the functionality of *AOP1* is unknown, it is considered to be the ancestral gene of *AOP2* and *AOP3* by gene duplication events, suggesting that *AOP1* may also function in GSL biosynthesis [60].

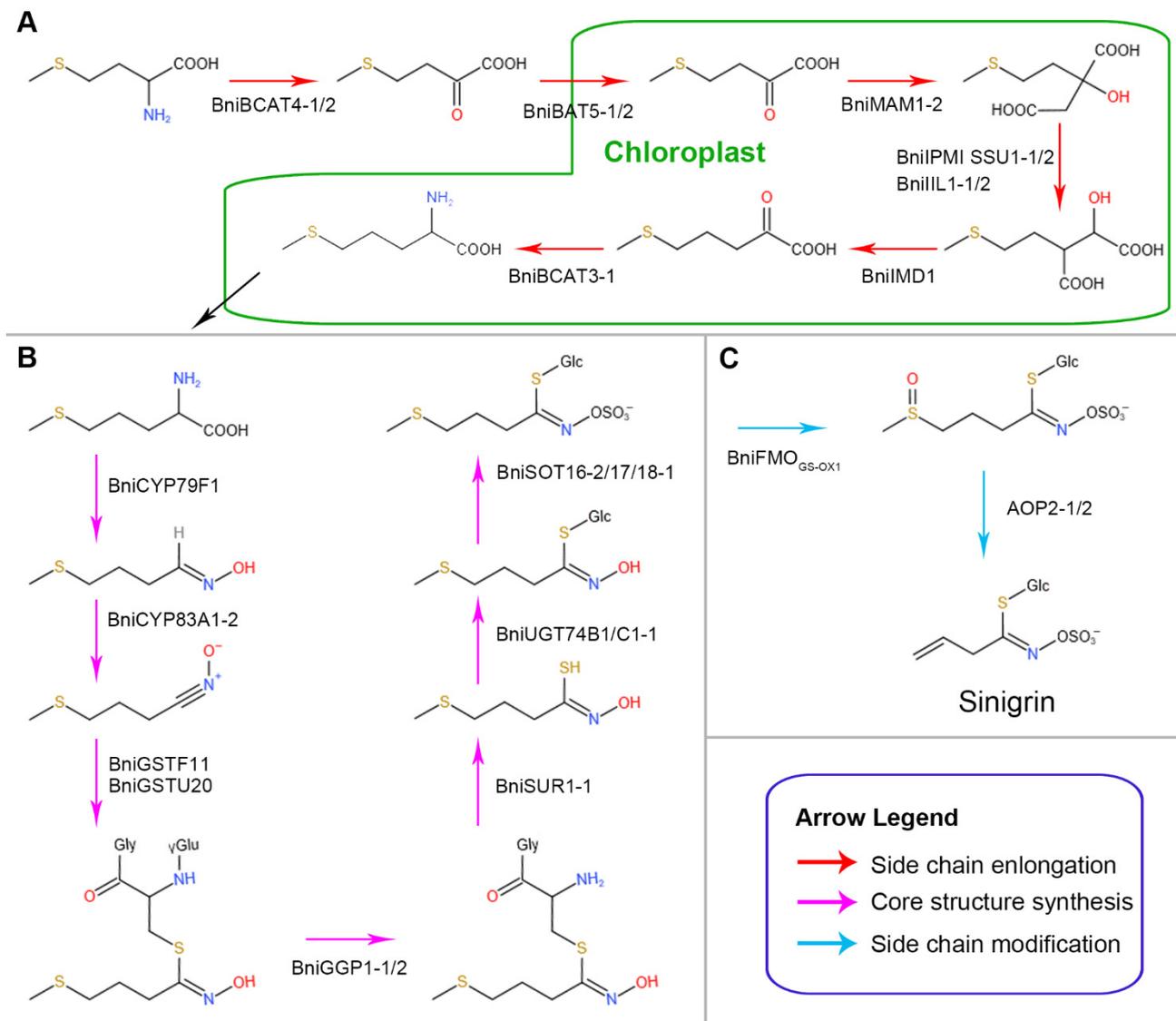
*AOP* genes in *B. rapa*, *B. oleracea*, and *B. nigra* were identified and subjected to expression analysis to assess the contribution of *AOP* genes on aliphatic GSL diversity in *B. nigra*, as well as the difference in *AOP* genes in different *Brassica* species. Results showed that *B. rapa*, *B. oleracea*, and *B. nigra* contained three, two, and two *AOP1* genes, respectively; three, one, and two functional *AOP2* genes, respectively; and no *AOP3* homolog (Figure 8A). The absence of *AOP3* in these three *Brassica* species may be the key reason why they rarely contain hydroxyalkyl GSLs. The *AOP1* genes in *Brassica* were mainly expressed in roots (Figure 8B–D), suggesting that they may function in the side-chain modification of aliphatic GSLs in roots. Most notably, significant differences were observed in the gene function and expression of *AOP2* genes in *Brassica*. Although three *AOP2* copies in *B. oleracea* were identified, the presence of a premature stop codon resulted in two of them being nonfunctional [43], while the other functional *BolAOP2* had no expression data (Figure 8C), which might be caused by low expression. In contrast, all three *AOP2* genes in *B. rapa* were functional, and Bra00848 was highly expressed in both stems and leaves (Figure 8B). Similarly, two copies of *AOP2* in *B. nigra*, i.e., *BniAOP2-1* and *BniAOP2-1*, were also extremely highly expressed in stems and leaves (Figure 8D). The difference in *AOP2* genes in *Brassica* supports their unique GSL profiles and partially explains why *B. oleracea* is rich in glucoraphanin, but not in *B. rapa*, and why sinigrin is abundant in *B. nigra*.



**Figure 8.** The phylogenetic tree and heatmap of *AOP* genes in *Brassica rapa*, *Brassica oleracea*, *Brassica nigra*, and *Arabidopsis thaliana*. (A) The phylogenetic tree of *AOP* genes. The *AOP* genes identified from *Brassica rapa*, *Brassica oleracea*, and *Brassica nigra* are indicated by green, purple, and red hollow circles or solid discs. The silenced genes are represented by a hollow circle, expressed functional genes are indicated by a solid disc. The heatmap of *AOP* genes in *Brassica rapa* (B), *Brassica oleracea* (C), and *Brassica nigra* (D). n.a., not available. The expression data of *B. rapa* and *B. oleracea* were obtained from ref [48,49].

#### 4. Discussion

Despite being one of the three ancestral *Brassica* species in U's triangle model, studies on *B. nigra* always lag behind that on *B. rapa* and *B. oleracea*, including the study on the identification of GSL biosynthesis genes at the genome-wide level. Although the absolute content of various GSLs in *Brassica* can be strongly influenced by environmental factors, the patterns of GSL are mainly controlled genetically [62,63]. Moreover, it was reported that genetic factors were dominant in controlling the synthesis of aliphatic GSLs [64]. Therefore, the genome-wide and expression analyses of GSL genes can help delineate the dominant GSL synthesis pathway. Here, through these analyses, we identified the genes involved in the biosynthesis of GSLs in *B. nigra*, and proposed a sinigrin biosynthesis pathway involving multiple candidate key genes in *B. nigra* (Table S3, Figure 9).



**Figure 9.** A pathway diagram of candidate key genes involved in sinigrin biosynthesis in *B. nigra*. (A) Side-chain elongation machinery. (B) Biosynthesis of core glucosinolate structure. (C) Secondary modification (alkenylations).

In this study, the GSL content survey once again confirmed that sinigrin was the most dominant GSL in *B. nigra* (Table 1), accounting for more than 90% of the total GSLs. In order to explore why *B. nigra* mainly synthesized and accumulated sinigrin from the genetic perspective. We searched out all GSL genes as much as possible on the basis of the latest

whole-genome data of *B. nigra*. *BniGSL* genes were identified by comparing their related homologs in *A. thaliana*, *B. rapa*, and *B. oleracea*. A total of 184 *BniGSL* genes were identified, of which 182 could be labeled on eight chromosomes of *B. nigra* (Tables 2 and 3, Figure 1). Compared with the identified GSL genes in *B. rapa* and *B. oleracea*, more members of *BniGSL* genes were determined in this study, which might be due to the incomplete genome data of *B. rapa* and *B. oleracea* previously [42,43,65]. Nevertheless, the GSL genes in these three *Brassica* species were all amplified considerably in their process of evolution. Six modes of gene duplication have been demonstrated in previous research, including whole-genome duplication (WGD) and tandem duplication (TD). The *Brassica* genome is proven to have triplicated soon after its divergence from *Arabidopsis* [44,66,67]. A functional bias is observed in genes retained after WGD and TD, which may show positive or negative correlations in the expansion of different members in a certain gene family. Duplicated genes may enhance the potential for the quantitative variation of a particular trait [68–70]. Similar to the GSL genes in *B. rapa* and *B. oleracea*, WGD and TD are the main mechanisms accounting for the expansion of *BniGS* genes, and most of them are present in multiple copies (Figure 1, Tables 2 and 3). For instance, previous research showed that *SOT18* is a multigene subfamily with tandem arrays of genes in *B. rapa* (10 members) and *R. sativus* (11 members) [42,71]. Thirteen *BniSOT18s* were identified in the current *B. nigra* genome and there were four groups of members expanded through TD (Figure 1, Table 2), indicating that TD was another factor responsible for the expansion of GSL genes during the evolution of *B. nigra* genome similar to those of *B. rapa* and *R. sativus*.

In general, there is no substantial difference among the inventory of GSL genes in *B. nigra* and those in *B. rapa* and *B. oleracea*, since most GSL genes have successfully retained at least one copy during the evolution of the *Brassica*, despite the difference in copy number of some GSL genes (Tables 2 and 3). Moreover, the absence of some GSL genes exists in all three *Brassica* species. For example, no ortholog of *CYP79F2* and *AOP3* has been identified in *Brassica* [42,43,46], which basically explains the lack of long-chain aliphatic and hydroxyalkyl GSLs [36,38,72]. However, there are significant differences in the GSL profiles of the three *Brassica* species. For example, sinigrin is the main GSL in *B. nigra*, and the sinigrin content in *B. nigra* is much higher than those in *B. rapa* and *B. oleracea*.

By transcriptome sequencing, we believe that the specific expression patterns of *BniGSL* genes in *B. nigra* largely determine its unique GSL profile. In the aliphatic GSL biosynthesis pathway, except the oxidative decarboxylation that involves only one gene (i.e., *BniIMD1*), each step (until the alkenylation of basic GSL) involves two or more genes, and at least one gene is highly expressed in stems and leaves (Figure 2A), ensuring the synthesis of a large amount of aliphatic GSL. Furthermore, the loss of *MAM3* orthologs results in failure to synthesize aliphatic GSL with long side chain, while the absence of *AOP3* and *GSL-OH* prevented the hydroxyalkylation of the basic GSL and the conversion of alkenyl GSL into hydroxylated alkenyl GSL. Therefore, we speculated that the expression characteristics of the aliphatic *BniGSL* genes and the above-mentioned limitations are possibly the main reason for the synthesis of a large amount of sinigrin in *B. nigra*. Specifically, we concluded that the candidate key genes involved in the sinigrin synthesis pathway are *BniBCAT4-1/2*, *BniBAT5-1/2*, *BniMAM1-2*, *BniIPMI SSU1-1/BniIPMI SSU2/BniIIL1-1/2*, *BniIMD1*, *BniBCAT3-1*, *BniCYP79F1*, *BniCYP83A1-2*, *BniGSTF11/BniGSTU20*, *BniGGP1-1/2*, *BniSUR1-1*, *BniUGT74B1/BniUGT74C1-1*, *BniSOT16-2/BniSOT17/BniSOT18-1*, *BniFMO<sub>CG-OX1</sub>*, and *BniAOP2-1/2* (Figure 9). Moreover, the results of expression analysis of *MAMs*, *CYP79s*, and *AOPs* in three *Brassica* species further indicate that the differences in gene expression have significant effects on GSL patterns and contents, and *BniMAM1-2*, *BniCYP79F1*, and *BniAOP2-1/2* may control the key nodes of the sinigrin synthesis pathway in *B. nigra*. In addition, the low expression of some key indolic and aromatic *BniGSL* genes in *B. nigra* is partially responsible for low indolic and aromatic GSL, respectively. These findings enriched our understanding of GSL biosynthesis patterns in *B. nigra* and provided guidance for changing the GSL profile in *B. nigra*, e.g., by regulating the expression of key *BniGSL* genes (e.g., overexpression/knockout) to change the GSL synthesis pathway in *B. nigra*,

thereby constructing new varieties with different GSL profiles (e.g., high/low sinigrin). Moreover, the identification of GSL biosynthesis pathways in *B. nigra* also provided a reference for the regulation of GSL synthesis in other *Brassica* species.

In this study, we determined the GSL profile of *B. nigra*, made an inventory of *BniGSL* genes and characterized their expression patterns. This research contributes to the functional analysis of *BniGSL* genes and improvement of GSLs in *B. nigra*, and provides a new perspective for future research on GSL synthesis in other species.

**Supplementary Materials:** The following are available online at <https://www.mdpi.com/article/10.3390/horticulturae7070173/s1>; Figure S1: HPLC chromatograms of glucosinolates isolated from different organs of *Brassica nigra*. Figure S2: Correlation analysis of samples. Table S1: GSL genes reported in *Brassica rapa* and *Brassica oleracea*. Table S2: List of primers used for qRT-PCR. Table S3: The DNA, CDS and amino acid sequences of glucosinolate biosynthetic related genes (*BniGSLs*) in *Brassica nigra*. Table S4: Expression data of *BniGSL* genes in different organs of *Brassica nigra* (FPKM).

**Author Contributions:** Conceptualization, Y.Y., Y.H. and Z.Z.; methodology, Y.L., L.X. and E.G.; software, Y.L.; validation, Y.Y. and Y.H.; formal analysis, Y.L. and L.X.; investigation, Y.L. and E.G.; resources, Y.Y. and Y.Z.; data curation, Y.Y.; writing—original draft preparation, Y.Y. and L.X.; writing—review and editing, Y.Y., Y.H., and Z.Z.; visualization, L.X.; supervision, Y.Y. and Y.H.; project administration, Y.H. and Z.Z.; funding acquisition, Z.Z. All authors have read and agreed to the published version of the manuscript.

**Funding:** This work was supported by the National Natural Science Foundation of China (No. 32072557 and No. 31572115) and the Zhejiang Provincial Natural Science Foundation of China (No. LZ14C150001).

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** The data used for the analysis in this study are available within the article and supplementary materials.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

- Blazevic, I.; Montaut, S.; Burcul, F.; Olsen, C.E.; Burow, M.; Rollin, P.; Agerbirk, N. Glucosinolate structural diversity, identification, chemical synthesis and metabolism in plants. *Phytochemistry* **2019**, *169*, 112100. [CrossRef]
- Agerbirk, N.; Hansen, C.C.; Kiefer, C.; Hauser, T.P.; Ørgaard, M.; Asmussen Lange, C.B.; Cipollini, D.; Koch, M.A. Comparison of glucosinolate diversity in the crucifer tribe Cardamineae and the remaining order Brassicales highlights repetitive evolutionary loss and gain of biosynthetic steps. *Phytochemistry* **2021**, *185*, 112668. [CrossRef]
- Halkier, B.A.; Gershenzon, J. Biology and Biochemistry of Glucosinolates. *Annu. Rev. Plant Biol.* **2006**, *57*, 303–333. [CrossRef] [PubMed]
- Agerbirk, N.; Olsen, C.E. Glucosinolate structures in evolution. *Phytochemistry* **2012**, *77*, 16–45. [CrossRef]
- Bones, A.M.; Rossiter, J.T. The enzymic and chemically induced decomposition of glucosinolates. *Phytochemistry* **2006**, *67*, 1053–1067. [CrossRef]
- Padilla, G.; Cartea, M.E.; Velasco, P.; De Haro, A.; Ordás, A. Variation of glucosinolates in vegetable crops of *Brassica rapa*. *Phytochemistry* **2007**, *68*, 536–545. [CrossRef] [PubMed]
- Hopkins, R.J.; van Dam, N.M.; van Loon, J.J.A. Role of glucosinolates in insect-plant relationships and multitrophic interactions. *Annu. Rev. Entomol.* **2009**, *54*, 57–83. [CrossRef]
- Arumugam, A.; Abdull Razis, A.F. Apoptosis as a Mechanism of the Cancer Chemopreventive Activity of Glucosinolates: A Review. *Asian Pac. J. Cancer Prev.* **2018**, *19*, 1439–1448. [PubMed]
- Mazumder, A.; Dwivedi, A.; Du Plessis, J. Sinigrin and Its Therapeutic Benefits. *Molecules* **2016**, *21*, 416. [CrossRef] [PubMed]
- Sønderby, I.E.; Geu-Flores, F.; Halkier, B.A. Biosynthesis of glucosinolates—Gene discovery and beyond. *Trends Plant Sci.* **2010**, *15*, 283–290. [CrossRef]
- Harun, S.; Abdullah-Zawawi, M.-R.; Goh, H.-H.; Mohamed-Hussein, Z.-A. A Comprehensive Gene Inventory for Glucosinolate Biosynthetic Pathway in *Arabidopsis thaliana*. *J. Agric. Food Chem.* **2020**, *68*, 7281–7297. [CrossRef] [PubMed]
- Grubb, C.D.; Abel, S. Glucosinolate metabolism and its control. *Trends Plant Sci.* **2006**, *11*, 89–100. [CrossRef]
- Textor, S.; Kraker, J.-W.; De Hause, B.; Gershenzon, J.; Tokuhisa, J.G. MAM3 catalyzes the formation of all aliphatic glucosinolate chain lengths in *Arabidopsis*. *Plant Physiol.* **2007**, *144*, 60–71. [CrossRef] [PubMed]

14. Gigolashvili, T.; Yatusевич, R.; Rollwitz, I.; Humphry, M.; Gershenzon, J.; Flügge, U.-I. The plastidic bile acid transporter 5 is required for the biosynthesis of methionine-derived glucosinolates in *Arabidopsis thaliana*. *Plant Cell* **2009**, *21*, 1813–1829. [[CrossRef](#)] [[PubMed](#)]
15. Zhang, J.; Wang, H.; Liu, Z.; Liang, J.; Wu, J.; Cheng, F.; Mei, S.; Wang, X. A naturally occurring variation in the *BrMAM-3* gene is associated with aliphatic glucosinolate accumulation in *Brassica rapa* leaves. *Hortic. Res.* **2018**, *5*, 69. [[CrossRef](#)] [[PubMed](#)]
16. Wang, C.; Dissing, M.M.; Agerbirk, N.; Crocoll, C.; Halkier, B.A. Characterization of Arabidopsis CYP79C1 and CYP79C2 by Glucosinolate Pathway Engineering in *Nicotiana benthamiana* Shows Substrate Specificity Toward a Range of Aliphatic and Aromatic Amino Acids. *Front. Recent Dev. Plant Sci.* **2020**, *11*, 57. [[CrossRef](#)] [[PubMed](#)]
17. Bak, S.; Feyereisen, R. The Involvement of Two P450 Enzymes, CYP83B1 and CYP83A1, in Auxin Homeostasis and Glucosinolate Biosynthesis. *Plant Physiol.* **2001**, *127*, 108–118. [[CrossRef](#)]
18. Naur, P.; Petersen, B.L.; Mikkelsen, M.D.; Bak, S.; Rasmussen, H.; Olsen, C.E.; Halkier, B.A. CYP83A1 and CYP83B1, two nonredundant cytochrome P450 enzymes metabolizing oximes in the biosynthesis of glucosinolates in Arabidopsis. *Plant Physiol.* **2003**, *133*, 63–72. [[CrossRef](#)]
19. Mikkelsen, M.D.; Naur, P.; Halkier, B.A. Arabidopsis mutants in the C-5 lyase of glucosinolate biosynthesis establish a critical role for indole-3-acetaldoxime in auxin homeostasis. *Plant J.* **2004**, *37*, 770–777. [[CrossRef](#)]
20. Grubb, C.D.; Zipp, B.J.; Kopycki, J.; Schubert, M.; Quint, M.; Lim, E.-K.; Bowles, D.J.; Pedras, M.S.C.; Abel, S. Comparative analysis of Arabidopsis UGT74 glucosyltransferases reveals a special role of UGT74C1 in glucosinolate biosynthesis. *Plant J.* **2014**, *79*, 92–105. [[CrossRef](#)]
21. Piotrowski, M.; Schemenewitz, A.; Lopukhina, A.; Müller, A.; Janowitz, T.; Weiler, E.W.; Oecking, C. Desulfoglucosinolate sulfotransferases from *Arabidopsis thaliana* catalyze the final step in the biosynthesis of the glucosinolate core structure. *J. Biol. Chem.* **2004**, *279*, 50717–50725. [[CrossRef](#)]
22. Hirschmann, F.; Krause, F.; Baruch, P.; Chizhov, I.; Mueller, J.W.; Manstein, D.J.; Papenbrock, J.; Fedorov, R. Structural and biochemical studies of sulphotransferase 18 from *Arabidopsis thaliana* explain its substrate specificity and reaction mechanism. *Sci. Rep.* **2017**, *7*, 4160. [[CrossRef](#)] [[PubMed](#)]
23. Burow, M.; Atwell, S.; Francisco, M.; Kerwin, R.E.; Halkier, B.A.; Kliebenstein, D.J. The Glucosinolate Biosynthetic Gene AOP2 Mediates Feed-back Regulation of Jasmonic Acid Signaling in Arabidopsis. *Mol. Plant* **2015**, *8*, 1201–1212. [[CrossRef](#)] [[PubMed](#)]
24. Kong, W.; Li, J.; Yu, Q.; Cang, W.; Xu, R.; Wang, Y.; Ji, W. Two Novel Flavin-Containing Monooxygenases Involved in Biosynthesis of Aliphatic Glucosinolates. *Front. Plant Sci.* **2016**, *7*, 1292. [[CrossRef](#)] [[PubMed](#)]
25. Gigolashvili, T.; Engqvist, M.; Yatusевич, R.; Müller, C.; Flügge, U.-I. HAG2/MYB76 and HAG3/MYB29 exert a specific and coordinated control on the regulation of aliphatic glucosinolate biosynthesis in *Arabidopsis thaliana*. *New Phytol.* **2008**, *177*, 627–642. [[CrossRef](#)]
26. 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 glucosinolates in Arabidopsis. *Plant Physiol.* **2010**, *153*, 348–363. [[CrossRef](#)]
27. Frerigmann, H.; Gigolashvili, T. MYB34, MYB51, and MYB122 distinctly regulate indolic glucosinolate biosynthesis in *Arabidopsis thaliana*. *Mol. Plant* **2014**, *7*, 814–828. [[CrossRef](#)]
28. Mitreiter, S.; Gigolashvili, T. Regulation of glucosinolate biosynthesis. *J. Exp. Bot.* **2021**, *72*, 70–91. [[CrossRef](#)]
29. Zhang, Y.; Li, B.; Huai, D.; Zhou, Y.; Kliebenstein, D.J. The conserved transcription factors, MYB115 and MYB118, control expression of the newly evolved benzoyloxy glucosinolate pathway in *Arabidopsis thaliana*. *Front. Plant Sci.* **2015**, *6*, 343. [[CrossRef](#)]
30. Bischoff, A.; Trémulot, S. Differentiation and adaptation in *Brassica nigra* populations: Interactions with related herbivores. *Oecologia* **2011**, *165*, 971–981. [[CrossRef](#)] [[PubMed](#)]
31. Obi, R.K.; Nwanebu, F.C.; Ndubuisi, U.U.; Orji, N.M. Antibacterial qualities and phytochemical screening of the oils of *Curcubita pepo* and *Brassica nigra*. *J. Med. Plants Res.* **2009**, *3*, 429–432.
32. Alam, M.B.; Hossain, M.S.; Haque, M.E. Antioxidant and anti-inflammatory activities of the leaf extract of *Brassica nigra*. *Int. J. Pharmaceut. Sci. Res.* **2011**, *2*, 303–310.
33. Song, X.; Wei, Y.; Xiao, D.; Gong, K.; Sun, P.; Ren, Y.; Yuan, J.; Wu, T.; Yang, Q.; Li, X.; et al. *Brassica carinata* genome characterization clarifies U's triangle model of evolution and polyploidy in *Brassica*. *Plant Physiol.* **2021**. [[CrossRef](#)] [[PubMed](#)]
34. Cheng, F.; Liang, J.; Cai, C.; Cai, X.; Wu, J.; Wang, X. Genome sequencing supports a multi-vertex model for Brassicaceae species. *Curr. Opin. Plant Biol.* **2017**, *36*, 79–87. [[CrossRef](#)] [[PubMed](#)]
35. Sharma, S.; Padmaja, K.L.; Gupta, V.; Paritosh, K.; Pradhan, A.K.; Pental, D. Two plastid DNA lineages—*Rapa/Oleracea* and *Nigra*—within the tribe Brassicaceae can be best explained by reciprocal crosses at hexaploidy: Evidence from divergence times of the plastid genomes and R-block genes of the A and B genomes of *Brassica juncea*. *PLoS ONE* **2014**, *9*, e93260. [[CrossRef](#)] [[PubMed](#)]
36. Volden, J.; Borge, G.I.A.; Hansen, M.; Wicklund, T.; Bengtsson, G.B. Processing (blanching, boiling, steaming) effects on the content of glucosinolates and antioxidant-related parameters in cauliflower (*Brassica oleracea* L. ssp. *botrytis*). *LWT Food Sci. Technol.* **2009**, *42*, 63–73. [[CrossRef](#)]
37. Wang, J.; Yu, H.; Zhao, Z.; Sheng, X.; Shen, Y.; Gu, H. Natural Variation of Glucosinolates and Their Breakdown Products in Broccoli (*Brassica oleracea* var. *italica*) Seeds. *J. Agric. Food Chem.* **2019**, *67*, 12528–12537. [[CrossRef](#)]
38. Yang, B.; Quiros, C.F. Survey of glucosinolate variation in leaves of *Brassica rapa* crops. *Genet. Resour. Crop Evol.* **2010**, *57*, 1079–1089. [[CrossRef](#)]

39. Mnzava, N.A.; Olson, K. Studies on tropical vegetables. Part 1: Seed amino, fatty acid and glucosinolate profile of Ethiopian mustards (*Brassica carinata* Braun). *Food Chem.* **1990**, *35*, 229–235. [[CrossRef](#)]
40. Rangkadilok, N.; Nicolas, M.E.; Richard, B.N.; Premier, R.R.; Eagling, D.R.; Taylor, P.W.J. Developmental changes of sinigrin and glucoraphanin in three *Brassica* species (*Brassica nigra*, *Brassica juncea* and *Brassica oleracea* var. *italica*). *Sci. Hortic.* **2002**, *96*, 11–26. [[CrossRef](#)]
41. Bellostas, N.; Sørensen, J.C.; Sørensen, H. Profiling glucosinolates in vegetative and reproductive tissues of four *Brassica* species of the U-triangle for their biofumigation potential. *J. Sci. Food Agric.* **2007**, *87*, 1586–1594. [[CrossRef](#)]
42. Wang, H.; Wu, J.; Sun, S.; Liu, B.; Cheng, F.; Sun, R.; Wang, X. Glucosinolate biosynthetic genes in *Brassica rapa*. *Gene* **2011**, *487*, 135–142. [[CrossRef](#)]
43. Liu, S.; Liu, Y.; Yang, X.; Tong, C.; Edwards, D.; Parkin, I.A.P.; Zhao, M.; Ma, J.; Yu, J.; Huang, S.; et al. The *Brassica oleracea* genome reveals the asymmetrical evolution of polyploid genomes. *Nat. Commun.* **2014**, *5*, 3930. [[CrossRef](#)] [[PubMed](#)]
44. 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 homoeolog gene expression influencing selection. *Nat. Genet.* **2016**, *48*, 1225–1232. [[CrossRef](#)]
45. Zhu, B.; Yang, J.; He, Y.; Zang, Y.; Zhu, Z. Glucosinolate Accumulation and Related Gene Expression in Pak Choi (*Brassica rapa* L. ssp. *chinensis* var. *communis* N. Tsen & S.H. Lee Hanelt) in Response to Insecticide Application. *J. Agric. Food Chem.* **2015**, *63*, 9683–9689.
46. Yi, G.-E.; Robin, A.H.K.; Yang, K.; Park, J.-I.; Kang, J.-G.; Yang, T.-J.; Nou, I.-S. Identification and expression analysis of glucosinolate biosynthetic genes and estimation of glucosinolate contents in edible organs of *Brassica oleracea* subspecies. *Molecules* **2015**, *20*, 13089–13111. [[CrossRef](#)]
47. Roshan, K.; Arya, G.C.; Bisht, N.C. Differential expression and interaction specificity of the heterotrimeric G-protein family in *Brassica nigra* reveal their development- and condition-specific role. *Plant Cell Physiol.* **2014**, *11*, 1954–1968.
48. 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 Genomics* **2013**, *14*, 689. [[CrossRef](#)]
49. Yu, J.; Tehrim, S.; Zhang, F.; Tong, C.; Huang, J.; Cheng, X.; Dong, C.; Zhou, Y.; Qin, R.; Hua, W.; et al. Genome-wide comparative analysis of NBS-encoding genes between *Brassica* species and *Arabidopsis thaliana*. *BMC Genomics* **2014**, *15*, 3. [[CrossRef](#)]
50. Nour-Eldin, H.H.; Andersen, T.G.; Burrow, M.; Madsen, S.R.; Jørgensen, M.E.; Olsen, C.E.; Dreyer, I.; Hedrich, R.; Geiger, D.; Halkier, B.A. NRT/PTR transporters are essential for translocation of glucosinolate defence compounds to seeds. *Nature* **2012**, *488*, 531–534. [[CrossRef](#)]
51. Nambiar, D.M.; Kumari, J.; Augustine, R.; Kumar, P.; Bajpai, P.K.; Bisht, N. GTR1 and GTR2 transporters differentially regulate tissue-specific glucosinolate contents and defence responses in the oilseed crop *Brassica juncea*. *Plant Cell Environ.* **2021**. [[CrossRef](#)] [[PubMed](#)]
52. Benderoth, M.; Textor, S.; Windsor, A.J.; Mitchell-Olds, T.; Gershenzon, J.; Kroymann, J. Positive selection driving diversification in plant secondary metabolism. *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 9118–9123. [[CrossRef](#)] [[PubMed](#)]
53. Benderoth, M.; Pfalz, M.; Kroymann, J. Methylthioalkylmalate synthases: Genetics, ecology and evolution. *Phytochem. Rev.* **2009**, *8*, 255–268. [[CrossRef](#)]
54. Hansen, N.; Ostermeier, A. Completely Derandomized Self-Adaptation in Evolution Strategies. *Evol. Comput.* **2001**, *9*, 159–195. [[CrossRef](#)] [[PubMed](#)]
55. Chen, S.; Glawischnig, E.; Jørgensen, K.; Naur, P.; Jørgensen, B.; Olsen, C.-E.; Hansen, C.H.; Rasmussen, H.; Pickett, J.A.; Halkier, B.A. CYP79F1 and CYP79F2 have distinct functions in the biosynthesis of aliphatic glucosinolates in *Arabidopsis*. *Plant J.* **2003**, *33*, 923–937. [[CrossRef](#)]
56. Miao, H.; Wei, J.; Zhao, Y.; Yan, H.; Sun, B.; Huang, J.; Wang, Q. Glucose signalling positively regulates aliphatic glucosinolate biosynthesis. *J. Exp. Bot.* **2013**, *64*, 1097–1109. [[CrossRef](#)] [[PubMed](#)]
57. Mikkelsen, M.D.; Halkier, B.A. Metabolic engineering of valine- and isoleucine-derived glucosinolates in *Arabidopsis* expressing CYP79D2 from Cassava. *Plant Physiol.* **2003**, *131*, 773–779. [[CrossRef](#)]
58. Zhao, Y.; Hull, A.K.; Gupta, N.R.; Goss, K.A.; Alonso, J.; Ecker, J.R.; Normanly, J.; Chory, J.; Celenza, J.L. Trp-dependent auxin biosynthesis in *Arabidopsis*: Involvement of cytochrome P450s CYP79B2 and CYP79B3. *Genes Dev.* **2003**, *16*, 3100–3112. [[CrossRef](#)]
59. Zang, Y.-X.; Kim, D.-H.; Park, B.-S.; Hong, S.-B. Metabolic engineering of indole glucosinolates in Chinese cabbage hairy roots expressing *Arabidopsis* CYP79B2, CYP79B3, and CYP83B1. *Biotechnol. Bioprocess Eng.* **2009**, *14*, 467–473. [[CrossRef](#)]
60. Kliebenstein, D.J.; Kroymann, J.; Brown, P.; Figuth, A.; Pedersen, D.; Gershenzon, J.; Mitchell-Olds, T. Genetic Control of Natural Variation in *Arabidopsis* Glucosinolate Accumulation. *Plant Physiol.* **2001**, *126*, 811–825. [[CrossRef](#)]
61. Jensen, L.M.; Kliebenstein, D.J.; Burrow, M. Investigation of the multifunctional gene *AOP3* expands the regulatory network fine-tuning glucosinolate production in *Arabidopsis*. *Front. Plant Sci.* **2015**, *6*, 762. [[CrossRef](#)]
62. Vallejo, F.; Tomás-Barberán, F.A.; Benavente-García, A.G.; García-Viguera, C. Total and individual glucosinolate contents in inflorescences of eight broccoli cultivars grown under various climatic and fertilisation conditions. *J. Sci. Food Agric.* **2003**, *83*, 307–313. [[CrossRef](#)]
63. Schonhof, I.; Krumbein, A.; Brückner, B. Genotypic effects on glucosinolates and sensory properties of broccoli and cauliflower. *Nahrung/Food* **2004**, *48*, 25–33. [[CrossRef](#)]

64. Farnham, M.W.; Wilson, P.E.; Stephenson, K.K.; Fahey, J.W. Genetic and environmental effects on glucosinolate content and chemoprotective potency of broccoli. *Plant Breed.* **2004**, *123*, 60–65. [[CrossRef](#)]
65. Zang, Y.-X.; Kim, H.U.; Kim, J.A.; Lim, M.-H.; Jin, M.; Lee, S.C.; Kwon, S.-J.; Lee, S.-I.; Hong, J.K.; Park, T.-H.; et al. Genome-wide identification of glucosinolate synthesis genes in *Brassica rapa*. *FEBS J.* **2010**, *276*, 3559–3574. [[CrossRef](#)]
66. Lysak, M.A.; Koch, M.A.; Pecinka, A.; Schubert, I. Chromosome triplication found across the tribe *Brassicaceae*. *Genome Res.* **2005**, *15*, 516–525. [[CrossRef](#)]
67. Wang, X.; Wang, H.; Wang, J.; Sun, R.; Wu, J.; Liu, S.; Bai, Y.; Mun, J.-H.; Bancroft, I.; Cheng, F.; et al. The genome of the mesopolyploid crop species *Brassica rapa*. *Nat. Genet.* **2011**, *43*, 1035–1039. [[CrossRef](#)] [[PubMed](#)]
68. Li, J.; Hansen, B.G.; Ober, J.A.; Kliebenstein, D.J.; Halkier, B.A. Subclade of flavin-monoxygenases involved in aliphatic glucosinolate biosynthesis. *Plant Physiol.* **2008**, *148*, 1721–1733. [[CrossRef](#)] [[PubMed](#)]
69. Edger, P.P.; Pires, J.C. Gene and genome duplications: The impact of dosage-sensitivity on the fate of nuclear genes. *Chromosome Res.* **2009**, *17*, 699–717. [[CrossRef](#)] [[PubMed](#)]
70. Freeling, M. Bias in plant gene content following different sorts of duplication: Tandem, whole-genome, segmental, or by transposition. *Annu. Rev. Plant Biol.* **2009**, *60*, 433–453. [[CrossRef](#)]
71. Wang, J.; Qiu, Y.; Wang, X.; Yue, Z.; Yang, X.; Chen, X.; Zhang, X.; Shen, D.; Wang, H.; Song, J.; et al. Insights into the species-specific metabolic engineering of glucosinolates in radish (*Raphanus sativus* L.) based on comparative genomic analysis. *Sci. Rep.* **2017**, *7*, 16040. [[CrossRef](#)] [[PubMed](#)]
72. Chen, X. Glucosinolates in Chinese *Brassica campestris* Vegetables: Chinese Cabbage, Purple Cai-tai, Choysum, Pakchoi, and Turnip. *Hortscience* **2008**, *43*, 571–574. [[CrossRef](#)]