Phyton-International Journal of Experimental Botany
DOI:10.32604/phyton.2020.09334
Article

Genome-Wide Identification and Expression Profiling Suggest that Invertase Genes Function in Silique Development and the Response to *Sclerotinia sclerotiorum* in *Brassica napus*

Jingsen Liu[1,2], Jinqi Ma[1,2], Ai Lin[1,2], Chao Zhang[1,2], Bo Yang[1,2], Liyuan Zhang[1,2], Lin Huang[1,2] and Jiana Li[1,2,*]

[1]Chongqing Rapeseed Engineering Research Center, College of Agronomy and Biotechnology, Chongqing, 400715, China
[2]Academy of Agricultural Sciences, Southwest University, Chongqing, 400715, China
*Corresponding Author: Jiana Li. Email: ljn1950@swu.edu.cn
Received: 03 December 2019; Accepted: 04 February 2020

Abstract: Invertase (INV), a key enzyme in sucrose metabolism, irreversibly catalyzes the hydrolysis of sucrose to glucose and fructose, thus playing important roles in plant growth, development, and biotic and abiotic stress responses. In this study, we identified 27 members of the *BnaINV* family in *Brassica napus*. We constructed a phylogenetic tree of the family and predicted the gene structures, conserved motifs, *cis*-acting elements in promoters, physicochemical properties of encoded proteins, and chromosomal distribution of the *BnaINV*s. We also analyzed the expression of the *BnaINV*s in different tissues and developmental stages in the *B. napus* cultivar Zhongshuang 11 using qRT-PCR. In addition, we analyzed RNA-sequencing data to explore the expression patterns of the *BnaINV*s in four cultivars with different harvest indices and in plants inoculated with the pathogenic fungus *Sclerotinia sclerotiorum*. We used WGCNA (weighted co-expression network analysis) to uncover *BnaINV* regulatory networks. Finally, we explored the expression patterns of several *BnaINV* genes in cultivars with long (Zhongshuang 4) and short (Ningyou 12) siliques. Our results suggest that BnaINVs play important roles in the growth and development of rapeseed siliques and the defense response against pathogens. Our findings could facilitate the breeding of high-yielding *B. napus* cultivars with strong disease resistance.

Keywords: *Brassica napus*; invertase; *Sclerotinia*; WGCNA; RNA-seq

1 Introduction

Sucrose, the main form of assimilated carbon produced during photosynthesis, is transported over long distances in most land plants [1]. Sucrose can be used as a source of both carbon and energy for plants after it is unloaded from the phloem into sink tissues (such as flowers, fruits, seeds, and roots) and converted into hexoses [2], a process largely catalyzed by sucrose synthase and invertase (INV) [3,4]. INV irreversibly catalyzes the decomposition of sucrose into glucose and fructose, thus playing a key role in sugar metabolism in higher plants [5,6]. INVs are divided into three types based on their subcellular

localization, solubility, optimum pH, and isoelectric point: vacuolar invertase (VINV), cell wall invertase (CWINV), and cytoplasmic invertase (CINV) [7,8].

CWINVs and VINVs are located in different organelles, but they both require an optimum pH of 4.5 to 5.0 to catalyze sucrose hydrolysis and are therefore collectively referred to as acid INVs [9]. Acid INVs contain two characteristic conserved amino acid motifs: an N-terminal β-fructofuran glycosidase motif and a C-terminal cysteine catalytic domain [10]. Since acid INVs share similar enzymatic properties and can hydrolyze sucrose and other β-fructose oligosaccharides, they are also referred to as β-fructofuranosidases [11]. CINVs are non-glycosylated proteins in glycoside hydrolase family 100 (GH100) [12]. CINVs exist as hexamers, with each subunit comprising an 6-barrel core structure and an insertion of three helices [13]. Asp188 and Glu414 are the putative catalytic residues of the CINVs, which are believed to maintain their stringent substrate specificity towards the 1,2-glycosidic bond of sucrose [13].

Increasing evidence suggests that INVs play pivotal roles in plant growth and development [14]. CWINVs affect embryo and endosperm growth in seeds by controlling their sugar composition (sucrose/hexose ratio), which determines their overall carbohydrate composition [14]. Increased CWINV activity in tomato (*Solanum lycopersicum*) led to increased fruit weight and hexose levels [15], while the specific overexpression of a CWINV gene in the *Arabidopsis thaliana* shoot meristem resulted in earlier flowering, more inflorescence branches, and more siliques than the wild type [16]. VINVs play major roles in cell expansion and sugar accumulation [17]. For example, VINV activity in cotton (*Gossypium hirsutum*) is closely related to fibroblast elongation, with higher VINV activity resulting in faster cell elongation [18]. VINV activity in tomato increases during fruit ripening, which is also associated with increased hexose levels [19]; however, when VINV activity was inhibited, the fruits were smaller, their hexose contents decreased, and their sucrose level increased [19]. VINVs regulate spikelet sizes in rice (*Oryza sativa*) by adjusting the ratio of hexose to sucrose [20].

INVs also play important roles in biotic and abiotic stress responses [14,21]. For example, under drought stress, the inadequate hydrolysis of sucrose to glucose limits the development of maize (*Zea mays*) ovaries; both CWINV and VINV play important roles in this process [17]. High CWINV activity promoted heat tolerance in young tomato fruits [22]. The overexpression of *CWINV* in tobacco (*Nicotiana* sp.) and rice resulted in constitutively high expression of Pathogenesis-related (*PR*) genes [23], encoding PR proteins typically produced in response to pathogen attack [24]. CINVs have not been as extensively studied as CWINVs and VINVs, but they are known to be essential in plants [25]. In *Lotus japonicus* and *Arabidopsis*, the deletion or mutation of CINV genes (*LjINV1*, *AtCINV1*, and *AtCINV2*) resulted in developmental defects and reduced growth rates [25,26]. CINVs are also thought to be essential for root development [27] and for maintaining reactive oxygen species (ROS) homeostasis [28]. Recent evidence suggests that CINVs also play central roles in cellulose biosynthesis and carbon distribution in *Arabidopsis* [26].

The *INV* gene family has been explored in rice [29], maize [30], pepper (*Capsicum annuum*) [31], *Populus* [32], sugarcane (*Saccharum officinarum*) [33], and many other plants, but little is known about this gene family in *Brassica napus* (rapeseed). INVs play important roles in plant growth and development and in biotic and abiotic stress responses [14,21], all of which are vital for rapeseed production [34].

B. napus is the second largest oilseed crop worldwide [35]. The identification and analysis of *INV* genes in rapeseed could facilitate the development of better rapeseed varieties. Therefore, in the current study, we performed detailed analysis of the *INV* gene family in *B. napus* (*BnaINV*).

2 Materials and Methods

2.1 Identification and Phylogenetic Analysis of INVs in B. napus, B. rapa, and B. oleracea

The sequences of eight *Arabidopsis* AtINVs were obtained from the TAIR database [36], and the protein sequences of BnaINVs were obtained from the Brassica database (BRAD, http://brassicadb.org/brad/index.php) [37]. The putative *BnaINV*, *BrINV*, and *BoINV* homologs were identified by BLAST analysis ($E < e^{20}$) with ncbi-blast-2.2.30 + software (ftp://ftp.ncbi.nlm.nih.gov/blast/executes/blast+/LATEST/) [38]. If the corresponding *Arabidopsis* homolog was identified in this search, the *BnaINV*, *BrINV*, or *BoINV* candidate gene was considered to be a putative homolog of the gene. The genes were aligned, and conserved blocks were identified using Gblocks software [39]. Substitution saturation tests of conserved blocks were performed using DAMBE software [40]. Finally, a maximum likelihood [41] evolutionary tree was constructed using MEGA [42] with default parameters (1000 replications).

2.2 Analysis of the Gene Structure, Conserved Motifs, Protein Characteristics, and Chromosomal Location of the BnaINVs

The isoelectric points and molecular weights of the BnaINV proteins were predicted using the ExPASy website (http://www.expasy.org/tools/) [43]. The chromosomal locations of the *BnaINV*s were obtained from the Genoscope database and visualized using MapChart 2.2 [44]. The gene structures of the *BnaINV*s were analyzed using the Gene Structure Display Server (GSDS; http://gsds.cbi.pku.edu.cn/) [45]. MEME (http://meme-suite.org/tools/meme) was used to identify the conserved motifs in the BnaINV protein sequences [46], with the maximum number of motifs set to 20 and default settings used for the other parameters. InterProScan (http://www.ebi.ac.uk/interpro/) was used to annotate the BnaINV protein motifs [47].

2.3 Prediction of Cis-acting Elements in the BnaINV Promoters

The promoter sequences of the *BnaINV*s (1,500 bp upstream of the initial codon) were extracted from the *B. napus* genome database (http://www.genoscope.cns.fr/brassicanapus/) [48]. PlantCARE (http://bioinformatics.psb.ugent.be/webtools/plantcare/html/) was used to identify the *cis*-regulatory elements predicted in the promoters of the *BnaINV*s [49].

2.4 Expression of the BnaINVs in Different Tissues and after Sclerotinia sclerotiorum Inoculation

To explore the roles of BnaINV in different yield-related traits in *B. napus*, the expression levels of the *BnaINV*s were extracted from transcriptome data downloaded from the GEO database at NCBI (https://www.ncbi.nlm.nih.gov/geo/). The expression levels of the *BnaINV*s were mapped using R-Studio, and SPSS was used to test significant differences [50]. The two transcriptome datasets used in this analysis were the SRP072900 and SRP075294 datasets. The SRP072900 dataset, which was studied, collated, and published by our team, contains the expression patterns of genes in eight different tissues from four groups of plant materials. The tissues included Le (mature leaves), St (stems), BP (buds on the primary branch), BM (buds on the main stem of the inflorescence), SM and SP (seeds on the main stem of the inflorescence and primary branches, respectively), and SPM and SPP (silique pericarps on the main stem of the inflorescence and primary branches, respectively). The four groups of plant cultivars were as follows: 1) cultivars with low yields and high harvest indices, 2) cultivars with high yields and low harvest indices, 3) cultivars with low yields and low harvest indices, and 4) cultivars with high yields and high harvest indices [51]. The SRP075294 dataset contains transcriptome data from the leaves of *B. napus* cultivars Zhongyou 821 and Westar extracted 1 day after inoculation with *S. sclerotiorum* [52].

2.5 RNA Isolation and qRT-PCR Analysis

The tissue-specific expression patterns of the *BnaINV*s were verified by RT-PCR using total RNA extracted from *B. napus* cultivar Zhongshuang 11. The differences in expression of selected *BnaINV*

genes between plants with long and short siliques were explored using total RNA extracted from *B. napus* cultivars Zhongshuang 4 and Ningyou 12, respectively. Total RNA was extracted from roots, stems, buds, stalks, and leaves during the full-bloom stage and from seeds and silique pericarps at 7, 14, 21, 30, and 40 days after flowering using an RNeasy Extraction kit (Thermo Fisher Scientific, Waltham, MA, USA). Reverse transcription was performed using an RNA PCR kit, version 3.3 (Takara Biomedical Technology, Beijing, China). Specific primers were designed based on the sequences of nine randomly selected genes using Primer Premier 5 software [53]. SsoAdvanced Universal SYBR Green Supermix was used for fluorescence quantitative PCR performed on a Bio-Rad CFX96 Real-time Quantitative System (Bio-Rad Laboratories, Hercules, CA, USA). The internal reference gene was *BnUBC21* (EV086936). Each condition had three biological and three technical repeats. The expression levels were calculated using the $2^{-\Delta\Delta Ct}$ method described by Livak et al. [54].

2.6 Weighted Gene Co-expression Network Analysis of the BnaINVs

Weighted Gene Co-expression Network Analysis (WGCNA) is an analytical method for exploring "supergenes" (gene modules with similar expression patterns) from high-throughput RNA-sequencing (RNA-seq) data. Here, 44 RNA-seq samples (including 12 samples from leaves inoculated with *S. sclerotiorum* and 32 samples from different tissues of four different cultivars) were subjected to WGCNA. A Topological Overlap Measure (TOM) was used to cluster genes with shared expression patterns into discrete modules based on the pairwise correlation between genes. After sample clustering and soft threshold screening, co-expression networks and module visualizations were constructed, and the *BnaINV*s were grouped into the identified modules. Gene Ontology (GO) analysis of genes in the modules was carried out using the R package topGO [55].

3 Results

3.1 Identification and Phylogenetic Analysis of the INV Family in B. napus

We obtained the sequences of eight AtINVs from the TAIR database and used them as query sequences [56] for a BLASTP search of the Brassica database. 27 BoINVs, 31 BrINVs, and 54 BnaINVs were identified as candidate INVs in *B. oleracea*, *B. rapa*, and *B. napus*, respectively (Fig. 1). We aligned the candidate protein sequences with the AtINV protein sequences to exclude those that did not belong to this family. We ultimately identified 14 BoINVs, 18 BrINVs, and 27 BnaINVs. By comparing the *Brassica* INVs with those of *Arabidopsis* in a phylogenetic tree, we classified the BnaINVs into 11 BnaCWINVs, 10 BnaVINVs, and 6 BnaCINVs.

3.2 Predicted Protein Characteristics and Chromosome Locations of BnaINV Family Genes

The 27 *BnaINV*s encode proteins ranging from 163 (BnaCINV1-1) to 664 (BnaVINV2-1) amino acid residues (aa) long, with an average length of 572 aa. The corresponding molecular weights of the BnaINVs range from 18.43 kDa (BnaCINV1-1) to 75.56 kDa (BnaVINV2-4), with an average molecular weight of 64.18 kDa. The BnaINVs have pI values ranging from 5.84 (BnaVINV2-2) to 9.25 (BnaCWINV2-2); the pI values of the BnaCWINVs are > 8.0, while those of BnaVINV and BnaCINV are < 7.0 and tend to be ~6.0 in almost all cases (Tab. 1).

As shown in Fig. 2, 15 *BnaINV*s are located on chromosomes in the A genome and 12 are located on chromosomes in the C genome. The three types of *BnaINV*s are concentrated in different chromosome regions. Some *BnaINV*s are located very close to each other on A and C chromosomes; for example, *BnaVINV1-1*, *-2*, and *-3* are located on the upper part of chromosome A09, while *BnaVINV1-4*, *-5*, and *-6* are located on the upper sections of chromosome C09. These findings suggest that these genes might have evolved from the *BrINV*s and *BoINV*s and have been conserved throughout the process of evolution.

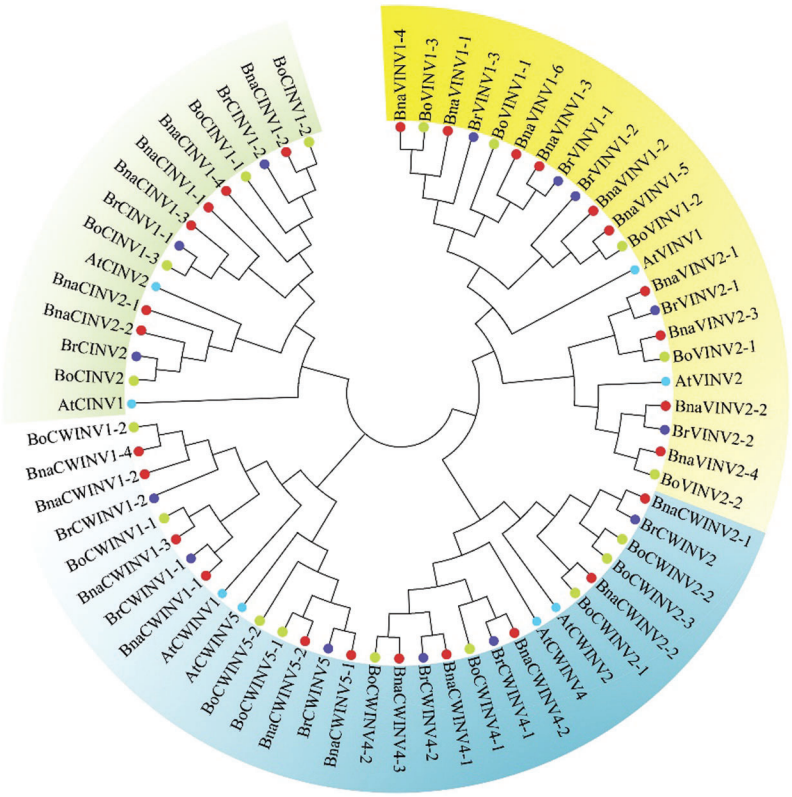

Figure 1: Phylogenetic tree generated by aligning 67 INV protein sequences from *Arabidopsis*, *B. rapa*, *B. oleracea*, and *B. napus*. The tree was generated using the maximum likelihood method

3.3 Analysis of Gene Structures and Conserved Motifs in BnaINV Family Genes

The genetic structures of the putative homologous genes reflect their levels of conservation. The diversity of the *BnaINV*s is reflected in their different exon numbers (Fig. 3). The exon number of the *BnaINV*s ranges from two (*BnaINV1-1*) to eight, with the majority of *BnaINV*s containing six to eight exons. The exon distributions and lengths differ among the three types of *BnaINVs*, although *BnaINV*s in the same category (CWINV, CINV, or VINV) share a higher degree of similarity. In general, the closer the relationship between genes, the more similar their structures. This notion was confirmed in the gene structure map; for example, the *BnaCWINV1* genes showed significant similarities in terms of exon/intron number, distribution, and length. Although *BnaCWINV5-1* has one less intron than *BnaCWINV5-2*, the locations and lengths of both of their introns are very similar. All *BnaCWINV*4 genes contain six exons, and all but one (*BnaCWINV4-1*) contains an untranslated region. All *BnaVINV2* genes possess seven exons except *BnaVINV2-4* (8 exons), but this group appeared to be divided into two groups (*BnaVINV2-1* and *BnaVINV2-3*, and *BnaVINV2-2* and *BnaVINV2-4*) based on their structural similarities. *BnaCINV1-1* is the shortest *BnaINV* gene, containing only two exons.

Twenty conserved motifs were identified in the BnaINV proteins using the MEME tool. The types and distributions of the motifs were conserved between BnaCWINV and BnaVINV proteins, while the distributions of conserved motifs among BnaCINV proteins were more diverse (Fig. 4). All BnaCWINV proteins contain Motifs 1-8, 10, 11, 13, and 14. Among the BnaCWINVs, only BnaCWINV5-1 and BnaCWINV4-2 lack Motif 9, while no BnaCWINV2 or BnaCWINV4 protein contains Motif 12. The conserved motifs in the BnaVINVs are similar to those of the BnaCWINVs, although the BnaCWINVs

Table 1: Complete list of the 27 *BnaINV* genes identified in this study

Gene Name	Locus Name	Arabidopsis Ortholog	Chr.	Location	CDS Length	Protein		
						pI	MW (Da)	Length (aa)
BnaCWINV1-1	BnaA01g37220D	AT3G13790.1	chrA01_random	2553748…2556634	1758	9.21	66,449.52	586
BnaCWINV1-3	BnaC01g37450D	AT3G13790.1	chrC01	36604664…36607569	1755	8.78	66,292.26	586
BnaCWINV1-2	BnaA05g25600D	AT3G13790.1	chrA05	18912001…18914938	1755	8.89	66,165.97	586
BnaCWINV1-4	BnaC05g39500D	AT3G13790.1	chrC05	38006404…38009436	1818	8.98	68,650.78	607
BnaCWINV5-1	BnaA05g25610D	AT3G13784.1	chrA05	18916401…18918499	1719	9.11	64,334.77	574
BnaCWINV5-2	BnaC05g39510D	AT3G13784.1	chrC05	38011430…38013543	1593	9.21	59,028.17	532
BnaCWINV2-1	BnaA04g05320D	AT3G52600.1	chrA04	4106722…4109307	1779	9.16	66,529.55	594
BnaCWINV2-2	BnaA09g47290D	AT3G52600.1	chrC04	29471181…29473081	1374	9.25	51,685.00	459
BnaCWINV4-2	BnaAnng07590D	AT2G36190.1	chrAnn_random	7358877…7361995	1779	8.61	67,352.41	594
BnaCWINV4-1	BnaA03g16550D	AT2G36190.1	chrA03	7727563…7730663	1773	8.51	67,350.32	592
BnaCWINV4-3	BnaC03g19970D	AT2G36190.1	chrC03	10446381…10449802	1773	8.48	67,261.28	592
BnaVINV2-1	BnaA06g08310D	AT1G12240.1	chrA06	4439391…4443430	1989	6.05	73,424.80	664
BnaVINV2-3	BnaC05g09640D	AT1G12240.1	chrC05	5237962…5241747	1980	6.00	73,119.49	661
BnaVINV2-2	BnaA09g47290D	AT1G12240.1	chrA09	31881446…31885131	1992	5.84	73,755.12	665
BnaVINV2-4	BnaC08g41390D	AT1G12240.1	chrC08	36051660…36055357	2043	6.05	75,457.03	682
BnaVINV1-2	BnaA09g10570D	AT1G62660.1	chrA09	5403407…5407214	1821	5.95	68,056.91	608
BnaVINV1-5	BnaC09g10840D	AT1G62660.1	chrC09	7407937…7411434	1821	5.91	68,203.25	608
BnaVINV1-1	BnaA09g10560D	AT1G62660.1	chrA09	5395224…5399630	1896	6.84	70,753.98	633
BnaVINV1-4	BnaC09g10830D	AT1G62660.1	chrC09	7370678…7376296	1896	6.86	70,802.11	633
BnaVINV1-3	BnaA09g12880D	AT1G62660.1	chrA09	6886365…6889885	1965	5.96	73,175.53	656
BnaVINV1-6	BnaC09g13230D	AT1G62660.1	chrC09	9903977…9907695	1965	6.04	73,261.61	656

Table 1 (continued).

Gene Name	Locus Name	Arabidopsis Ortholog	Chr.	Location	CDS Length	Protein		
						pI	MW (Da)	Length (aa)
BnaCINV1-1	BnaA05g34980D	AT1G35580.1	chrA05_random	444033…444598	486	6.43	18,426.01	163
BnaCINV2-1	BnaA03g24030D	AT4G09510.1	chrA03	11533628…11535979	1587	6.02	60,951.27	553
BnaCINV2-2	BnaC03g28560D	AT4G09510.1	chrC03	16861682…16864101	1587	6.11	60,983.32	530
BnaCINV1-4	BnaC08g06790D	AT1G35580.1	chrC08	9356365…9358411	1515	6.38	56,939.70	506
BnaCINV1-2	BnaA08g06300D	AT1G35580.1	chrA08	6285510…6287554	1656	6.51	62,935.52	553
BnaCINV1-3	BnaAnng38240D	AT1G35580.1	chrAnn_random	43282963…43284331	1089	6.55	41,577.16	364

do not contain Motif 15 or 19. The BnaVINV1s also lack Motif 19, but they contain two copies of Motif 13. By contrast, the conserved motif composition in the BnaCINVs was quite different. Motifs 16, 17, 18, and 20 were only found in these shorter proteins. Among the BnaCINVs, BnaCINV1-1 is particularly unusual because it only contains a single motif, Motif 11. All BnaCINVs contain Motifs 7, 14, 16, 17, and 18, except for BnaCINV1-1.

Of the 20 conserved motifs identified, all BnaINVs except BnaCINV1-1 contain Motifs 7 and 14, and all BnaINVs except BnaCINV1-3 contain Motif 11. With the exception BnaCINV1-1 and BnaCINV1-3, all BnaINVs contain Motif 10. We used the InterProScan program to annotate the motifs. Motifs 1-7 and 9 are predicted to encode glycosyl hydrolase domains and belong to the glycosyl hydrolase 43 (GH43) family. Motifs 16, 18, and 20 belong to the glycoside hydrolase 100 (GH100) family, while Motif 17 is associated with glycosyltransferase and glycoside bond formation. The other motifs could not be annotated.

3.4 Prediction and Analysis of cis-Acting Elements in the BnaINV Promoters

Cis-acting elements in promoter function as binding sites for transcription factors, which play important roles in regulating gene transcription. In addition to common *cis*-regulators such as TATA boxes, CAAT boxes, and light response elements, we identified 14 other *cis*-regulators in the *BnaINV* promoters. Among these 14 *cis*-acting elements, the MEJA (CGTCA motif and TGACG motif)-reacting elements were the most common and were detected 103 times among the 27 *BnaINV* genes. *Cis*-acting elements for hormone and stress responses were also fairly common, including abscisic acid response elements (ABREs), gibberellin response elements (TATC boxes, P boxes, and GARE motifs), salicylic acid response elements (TCA elements), auxin response elements (TGA elements and AuxRR cores), low-temperature response elements (LTRs), and anaerobic induction response elements (AREs). AREs and ABREs were present in most genes, with 97 and 89, respectively, identified across the 27 *BnaINV* genes. Circadian and RY-elements appeared in only a few genes.

3.5 BnaINV Expression Patterns in Different Tissues and Organs

As shown in Fig. 5, the expression patterns of several *BnaINV* genes significantly differed among cultivars. For example, the expression levels of *BnaVINV2-1* in seeds were higher in high harvest index materials (materials 1 and 4) than in low harvest index materials (materials 2 and 3). The expression

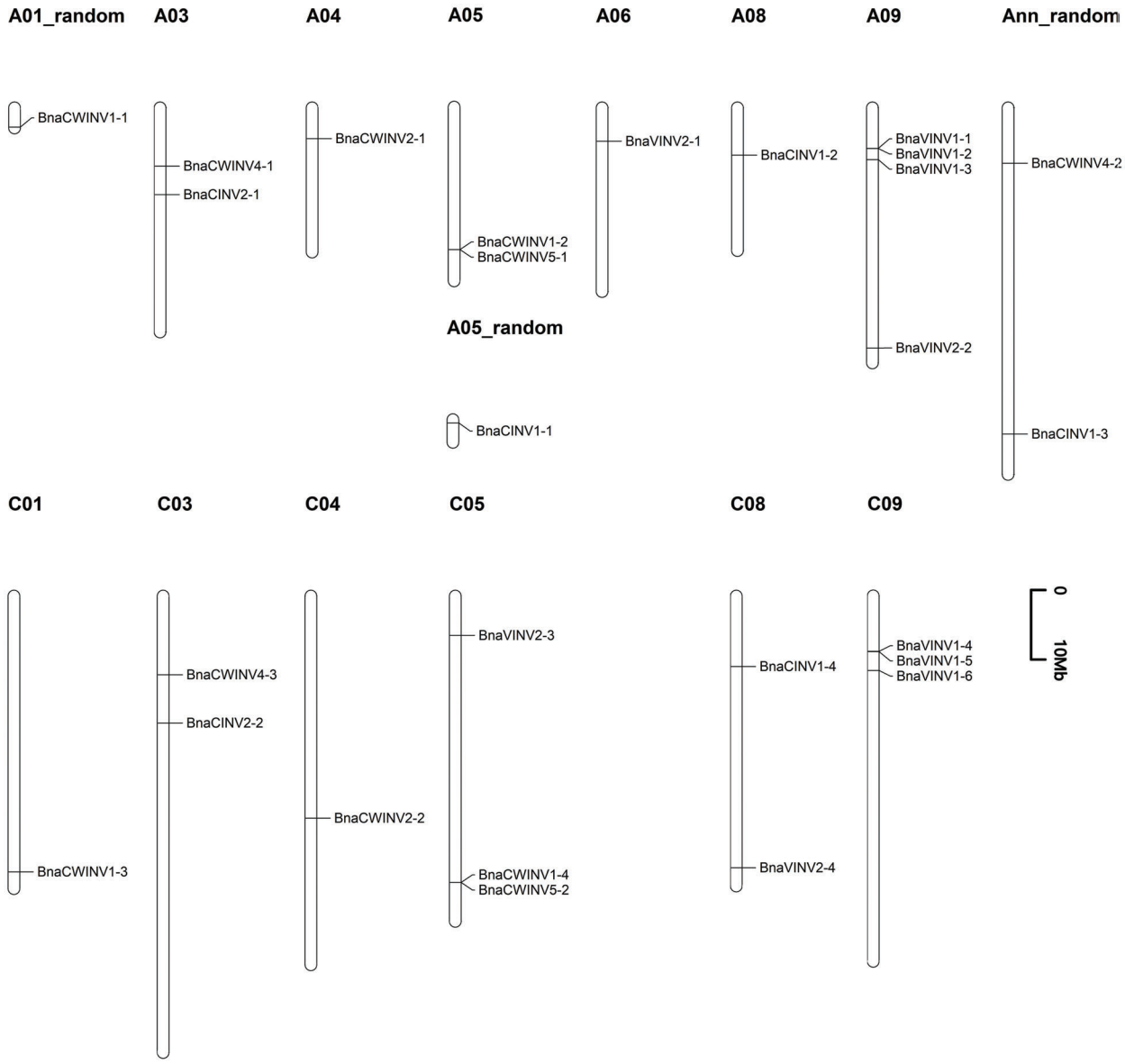

Figure 2: Distribution of *BnaINV* genes on the chromosomes of the *B. napus* genome. The chromosome number is indicated above each chromosome. The scale is in megabases (Mb)

levels of *BnaCWINV1-1* and *BnaCWINV1-3* in leaves were higher high-yield materials (materials 1 and 3) than in low-yield materials (materials 2 and 4). The 27 *BnaINV* genes showed a variety of expression patterns in different tissues, with similar expression patterns among members with similar genetic relationships. The *BnaCWINV*s showed high levels of tissue specificity: *BnaCWINV1* was specifically expressed in the main inflorescence and collateral pericarp, *BnaCWINV1-1* and *BnaCWINV1-3* were highly expressed in the pericarp; and *BnaCWINV*5 and *BnaCWINV*4 were specifically expressed in the main inflorescence bud and lateral branch buds. *BnaCWINV4-2* was expressed at extremely high levels, while the overall expression level of *BnaCWINV2* was low. *BnaVINV1-3* and *BnaVINV1-6* were specifically expressed in the main inflorescence bud and lateral branch bud, whereas *BnaVINV1-2* was not expressed in most tissues. The *BnaCINV*s were expressed in all tissues, and the expression patterns of *BnaCINV2-1* and

Figure 3: Exon-intron structures of the *BnaINV* genes according to their phylogenetic relationships. The unrooted phylogenetic tree was constructed with 1,000 bootstraps based on the full-length sequence of *BnaINV*. Exon-intron structure analysis of the *BnaINV* genes was performed using the online tool GSDS. The lengths of the exons (yellow boxes) and introns (black lines) of each *BnaINV* gene are drawn to scale

BnaCINV2-2 were similar. *BnaCINV1-1* was expressed at moderate levels in all tissues, while *BnaCINV1-4* showed the lowest expression levels in all tissues. The *BnaINV*s were widely expressed in the pericarp but at low levels in seeds.

3.6 Dynamic Expression of the BnaINVs after Inoculation with S. sclerotiorum

S. sclerotiorum is a major pathogen that causes yield reductions in *B. napus*. We therefore used transcriptome data to produce heatmaps to explore the expression pattern of *BnaINV* in *B. napus* following inoculation with *S. sclerotiorum*. As shown in Fig. 6, one day after *S. sclerotiorum* infection, the expression patterns of the *BnaINV*s were similar between cultivars Westar and Zhongyou 821. The *BnaCWINV1*s were significantly upregulated following *S. sclerotiorum* infection, while the expression levels of the *BnaCWINV2*s and *BnaCWINV5*s were not significantly affected by this pathogen. The expression level of *BnaCWINV4-3* decreased significantly following *S. sclerotiorum* infection. *BnaVINV1*s were expressed at very low levels both before and after treatment, whereas the overall expression levels of the *BnaCINV*s were relatively high both before and after treatment. *BnaVINV2-1* and *BnaVINV2-3* were significantly upregulated in infected plants, whereas *BnaVINV2-4* was upregulated only in Zhongyou 821. The expression of *BnaCINV1-1* increased significantly after infection, whereas *BnaCINV1-3* expression increased significantly only in Zhongyou 821. By contrast, the expression of *BnaCINV1-2* decreased significantly in Westar plants infected with *S. sclerotiorum*. Overall, the responsiveness of the *BnaINV*s to *S. sclerotiorum* infection appeared to be lower in Westar than in

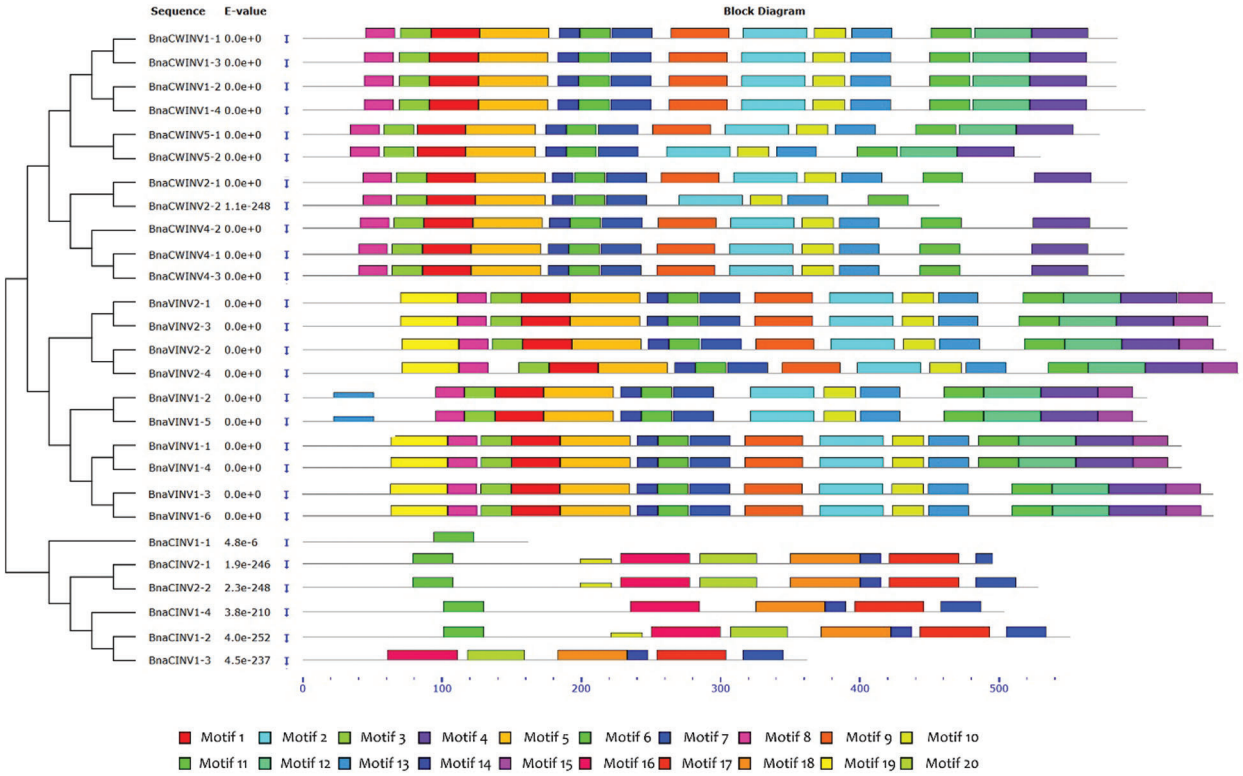

Figure 4: The conserved motifs of the BnaINV proteins according to their phylogenetic relationships. The conserved motifs of the BnaINV proteins were identified using MEME. Gray lines represent non-conserved sequences, and each motif is denoted by a colored box. The lengths of the motifs in each protein are drawn to scale

Zhongyou 821. This finding is consistent with the observation that Westar is a pathogen-sensitive cultivar, while Zhongyou 821 is a disease-resistant line [57].

3.7 qRT-PCR Analysis of BnaINV Expression in Specific Tissues and at Different Developmental Stages

To elucidate the expression patterns of the *BnaINV*s in *B. napus*, we selected several *BnaINV* genes (including members of the three INV categories) with high expression levels or specific expression patterns based on the transcriptome data. We used qRT-PCR to detect the expression levels of these genes in the flowers, leaves, roots, buds, stems, pericarp, and seeds of Zhongshuang 11. As shown in Fig. 7, the expression patterns of the same types of *BnaINV* were more similar to each other than to those of the other categories; for example, four *BnaCINV*s were expressed in all tissues to varying degrees, especially in stems and silique pericarps collected 7 and 14 days after flowering. Three of the four *BnaCINV*s (except *BnaCINV2-1*) were expressed in roots, although all four of these genes were expressed at relatively low levels in flowers and buds.

All four *BnaVINV*s were highly expressed in silique pericarps at 40 days after flowering but were expressed at low levels in roots. *BnaVINV2-1* and *BnaVINV2-3* were expressed in seeds at various developmental stages at significantly higher levels compared to *BnaVINV2-2* and *BnaVINV2-4*. *BnaCWINV1-2*, *BnaCWINV1-3*, *BnaCWINV4-2*, and *BnaCWINV4-3* showed strong tissue specificity in their expression patterns. *BnaCWINV1-2* was highly expressed in leaves, buds, and silique pericarps at 7 and 14 days after flowering, but not in other tissues. *BnaCWINV1-3* was highly expressed in silique

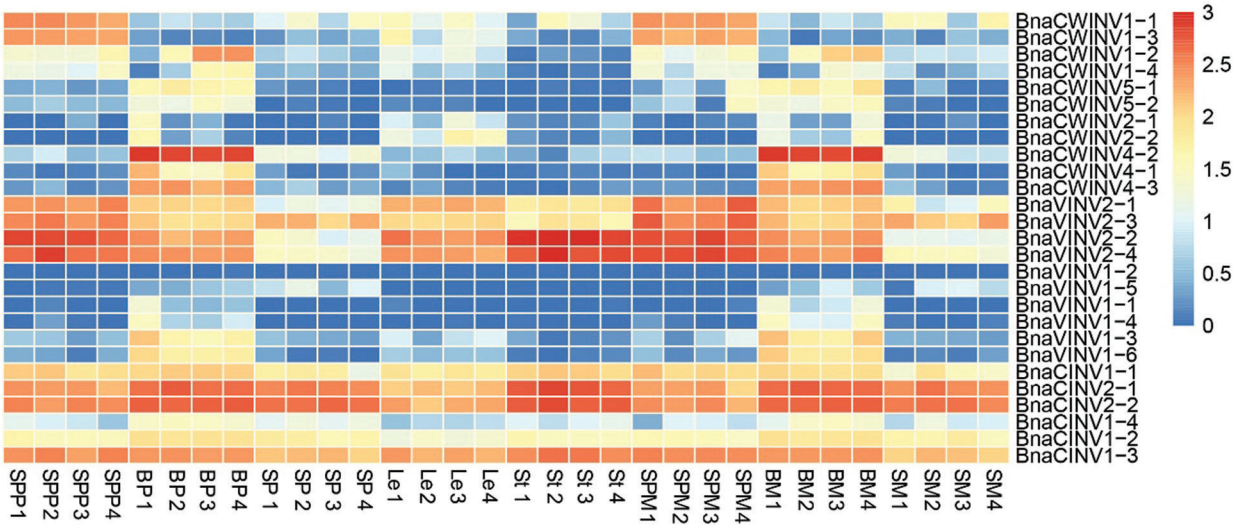

Figure 5: Heatmap of the expression patterns of the *BnaINV*s in eight organs of *B. napus* (Chongqing cultivar). The expression data were obtained from publicly available RNA-seq datasets (SRP072900). Le, mature leaves; S, stems; BP, buds on the primary branch; BM, buds on the main stem of the inflorescence; SM and SP, seeds on the main stem of the inflorescence and primary branch, respectively; SPM and SPP, silique pericarps on the main stem of the inflorescence and primary branch, respectively; 1, cultivars with low yields and high harvest indices; 2, cultivars with high yields and low harvest indices; 3, cultivars with low yields and low harvest indices; 4, cultivars with high yields and high harvest indices

pericarps at 40 days after flowering, like *BnaCWINV1-2*, but was also moderately expressed in leaves, roots, and silique pericarps at 14 days after flowering and in seeds collected 30 days after flowering, but it showed limited expression in all other tissues. *BnaCWINV4-2 a*nd *BnaCWINV1-3* were almost exclusively expressed in flowers and buds, which is consistent with the transcriptome data. All three types of *BnaINV* genes were highly expressed in different stages of silique development, suggesting they play important roles in various stages of pericarp development.

3.8 WGCNA and GO Enrichment Analysis of the BnaINVs

We used WGCNA to further clarify the relationship between *BnaINV* genes, as shown in Fig. 8. The sample consisted of 44 sets of transcriptome data (Fig. 8A). After sample clustering and soft threshold screening (where the screening value was 1), nine modules were ultimately identified: the black (829 genes), blue (4,982 genes), brown (4,272 genes), green (2,553 genes), pink (234 genes), red (1,273 genes), turquoise (5,367 genes), yellow (3,320 genes), and gray modules. The gray module contained all genes of the entire *B. napus* genome that could not be classified into one of the other modules. Of the 27 members of the *BnaINV* family, 13 genes were categorized in the gray module, while the 14 remaining genes were grouped into the yellow, brown, blue, green, and turquoise modules.

Genes in the latter modules are more likely to play important roles than genes in the gray model. Therefore, we analyzed the GO terms associated with all the genes in each module and found that genes in different modules have different functions. The yellow module contained seven *BnaINV* genes, including *BnaCWINV1-1*, *-2*, *-3*, and *-4*, *BnaCINV1-1* and *-3*, and *BnaVINV2-3*. These genes are involved in the response to chitin and are highly correlated with the response to external biological stimuli and stresses (Fig. 9). The functions of *BnaCWINV4-2* and *BnaCWINV4-3* in the brown module are mainly

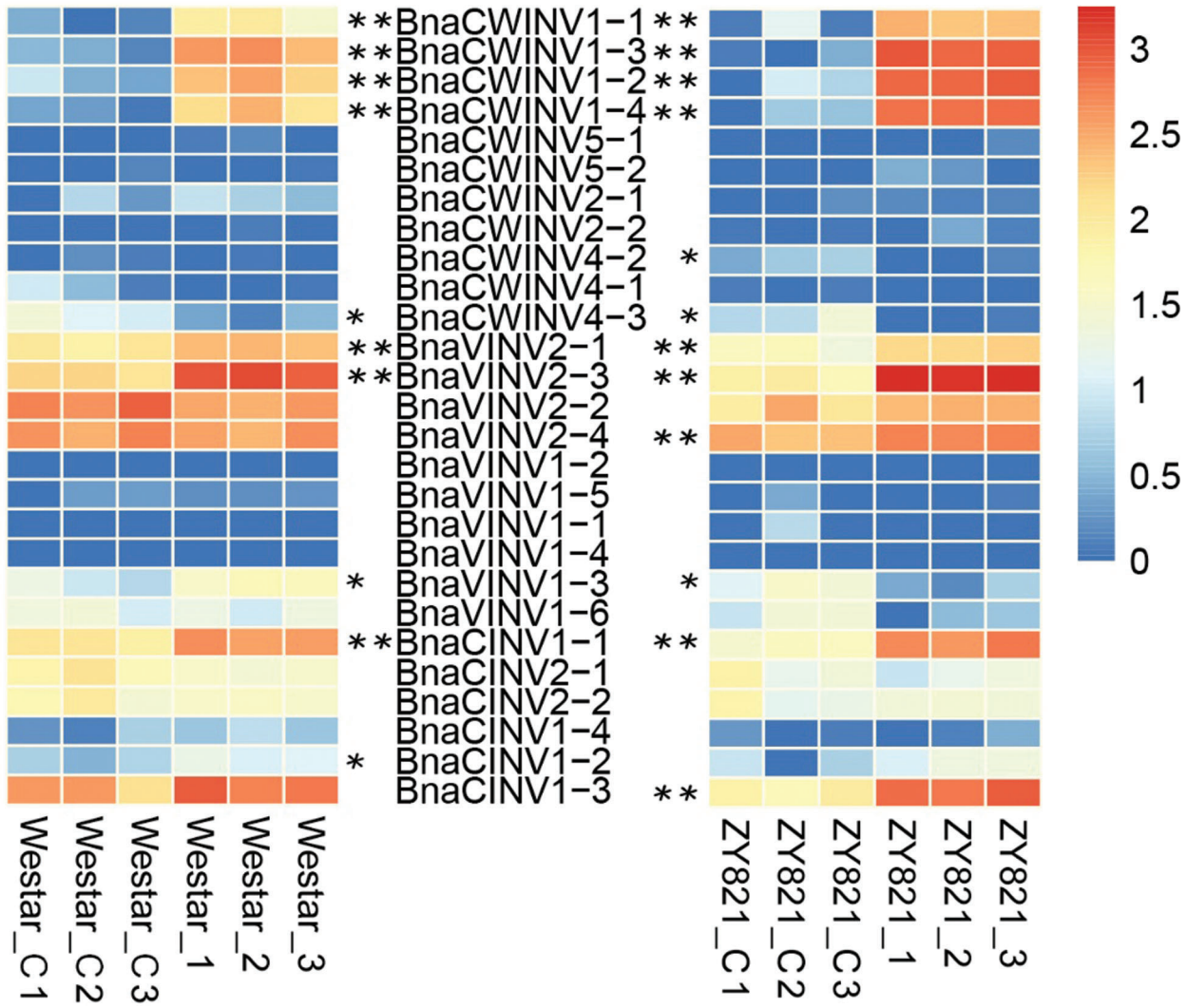

Figure 6: Heatmap of *BnaINV* expression in leaves at 1 day after inoculation with *S. sclerotiorum*. C, control; 1, 2, and 3 represent the three repetitions

related to lipid localization and the formation of cell walls and pollen walls. *BnaCINV2-1* and *-2* in the blue module are functionally associated with the metabolism of amides and peptides, protein translation, and ribosome composition. *BnaVINV2-2* and *-4* in the green module are involved in seed maturation, seed oil body biogenesis, and lipid localization. *BnaVINV2-1* in the turquoise module is involved in the photoreaction in photosynthesis.

3.9 qRT-PCR Analysis of BnaINV Expression in Different B. napus Cultivars

Based on the results of RNA-seq, qRT-PCR of gene expression in Zhongshuang 11, WGCNA, and GO analysis, we selected five *BnaINV* genes (*BnaCWINV4-2, BnaCWINV4-3*, *BnaVINV2-1*, *BnaVINV2-2*, and *BnaVINV2-4*) that might be related to pod growth and development and explored their expression patterns in *B. napus* cultivars with long (Zhongshuang 4) and short (Ningyou 12) siliques via qRT-PCR (Tab. S7). These five genes showed diverse expression patterns in different tissues or developmental stages in Zhongshuang 4 *vs*. Ningyou 12 (Fig. 10). *BnaCWINV4-2* and *BnaCWINV4-3* were specifically

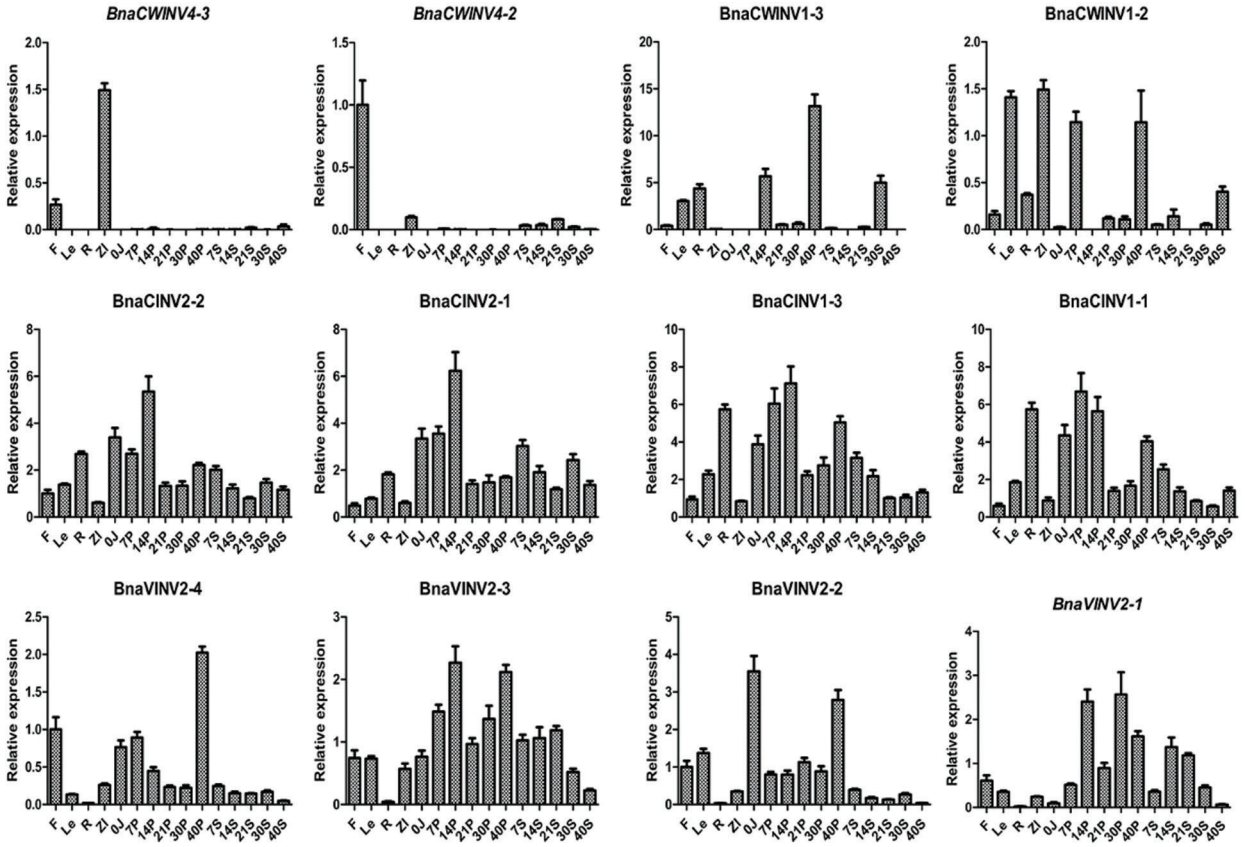

Figure 7: Relative expression levels of selected *BnaINV* genes in different tissues and at different developmental stages in *B. napus* cultivar Zhongshuang 11, as determined by qRT-PCR. F, flowers; Le, leaves; R, roots; Zl, buds; 0J, stems; 7S, 14S, 21S, 30S, and 40S, seeds collected at 7, 14, 21, 30, and 40 days after flowering, respectively; 7P, 14P, 21P, 30P, and 40P, silique pericarps collected at 7, 14, 21, 30, and 40 days after flowering, respectively. The internal reference gene was *BnUBC21* (EV086936)

expressed in flowers and buds, with significantly higher expression levels in Ningyou 12 than in Zhongshuang 4. The expression levels of *BnaVINV2-2* and *BnaVINV2-4* were significantly different in the flowers, buds and pods of the two cultivars. In particular, *BnaVINV2-2* was expressed at extremely high levels in silique pericarps at 40 days after flowering in Zhongshuang 4 but not Ningyou 12. The expression patterns of *BnaVINV2-1* in different organs and tissues significantly differed in Zhongshuang *vs*. Ningyou 12.

4 Discussion

All Brassica species evolved from a common hexaploid ancestor that experienced a whole genome triplication event. *B. napus* is a hybrid of *B. rapa* and *B. oleracea* [48,58]; therefore, the *B. napus* genome contains the *B. rapa* and *B. oleracea* genomes, it has six copies of conserved *Arabidopsis* genome segments [59]. Here, we used eight *AtINV* sequences as queries to search for *Brassica* INV homologs and identified 14 *BoINVs*, 18 *BrINVs*, and 27 *BnaINVs* using forward and reverse BLAST searches. This finding suggests that functional redundancy between these genes led to different degrees of gene contraction between these species. Using phylogenetic analysis, we explored the relationships between these genes in the three *Brassica* species. We identified three *BoCINV1s*, two *BrCINV1s*, and

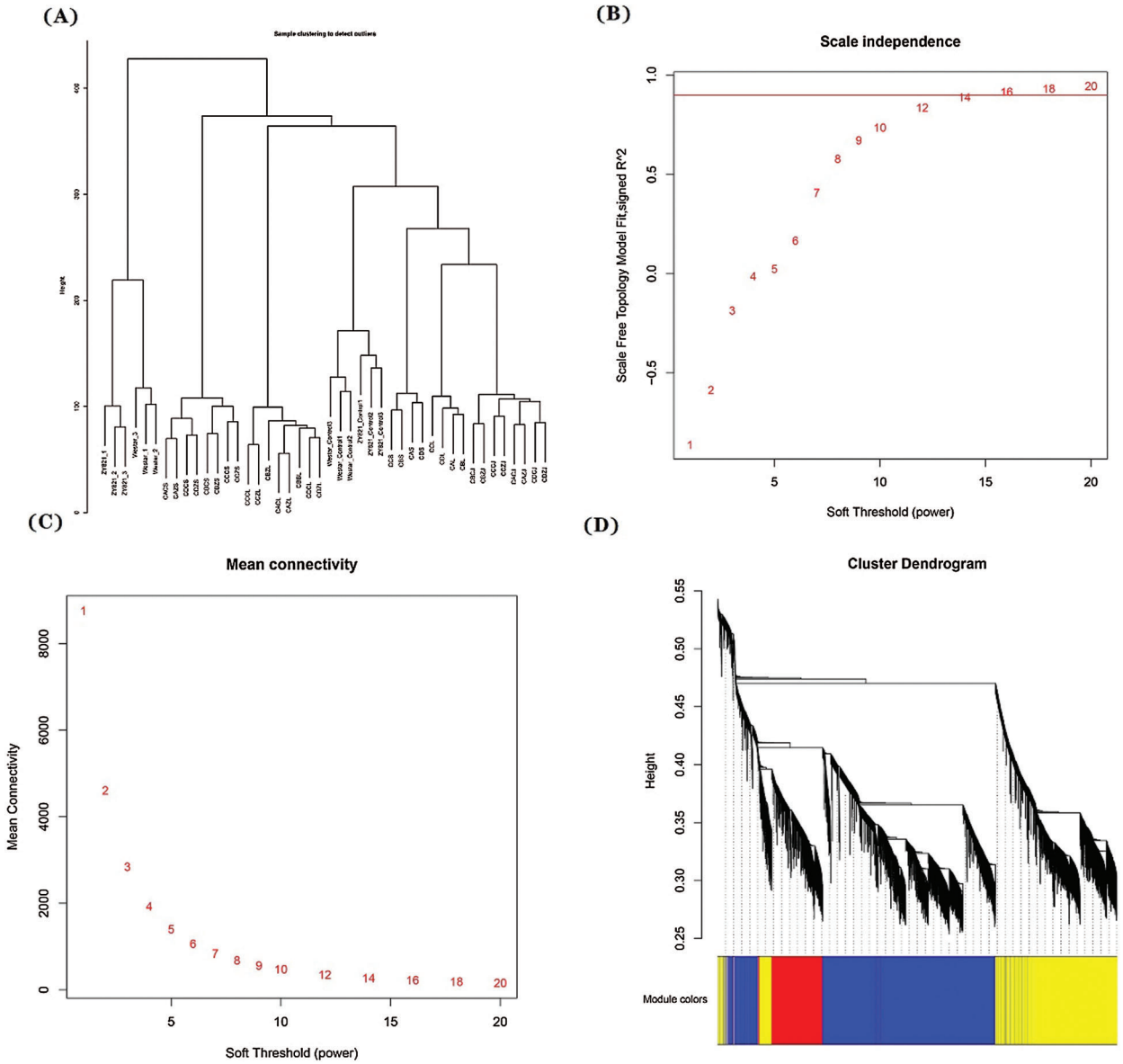

Figure 8: Module identification using WGCNA. (A) Sample clustering to detect outliers. (B and C) Selection of the soft threshold based on scale independence and mean connectivity. (D) Cluster dendrogram of the nine identified modules. The genes included in the gray module were not classified into any other module: their expression patterns were not significantly related

four *BnaCINV1s* in *B. napus* For *CWINV* genes, we detected three *BoCWINV2s*, one *BrCWINV2*, and two *BnaCWINV2s*; two *BoCWINV4s*, two *BrCWINV4s*, and three *BnaCWINV4s*; and two *BoCWINV5s*, one *BrCWINV5*, and two *BnaCWINV5s*. These results suggest that the contraction of the *BnaINV* gene family primarily occurred in the *BnaCWINV2s*, *BnaCWINV4s*, *BnaCWINV5s*, and *BnaCINV1s*.

In addition, we comprehensively analyzed the gene structures, conserved motifs, protein functions, *cis*-acting elements in promoters, and chromosomal locations of the *BnaINV*s. The genetic structures of the *BnaCWINV*s and *BnaVINV*s were similar to those of the other members of their respective categories,

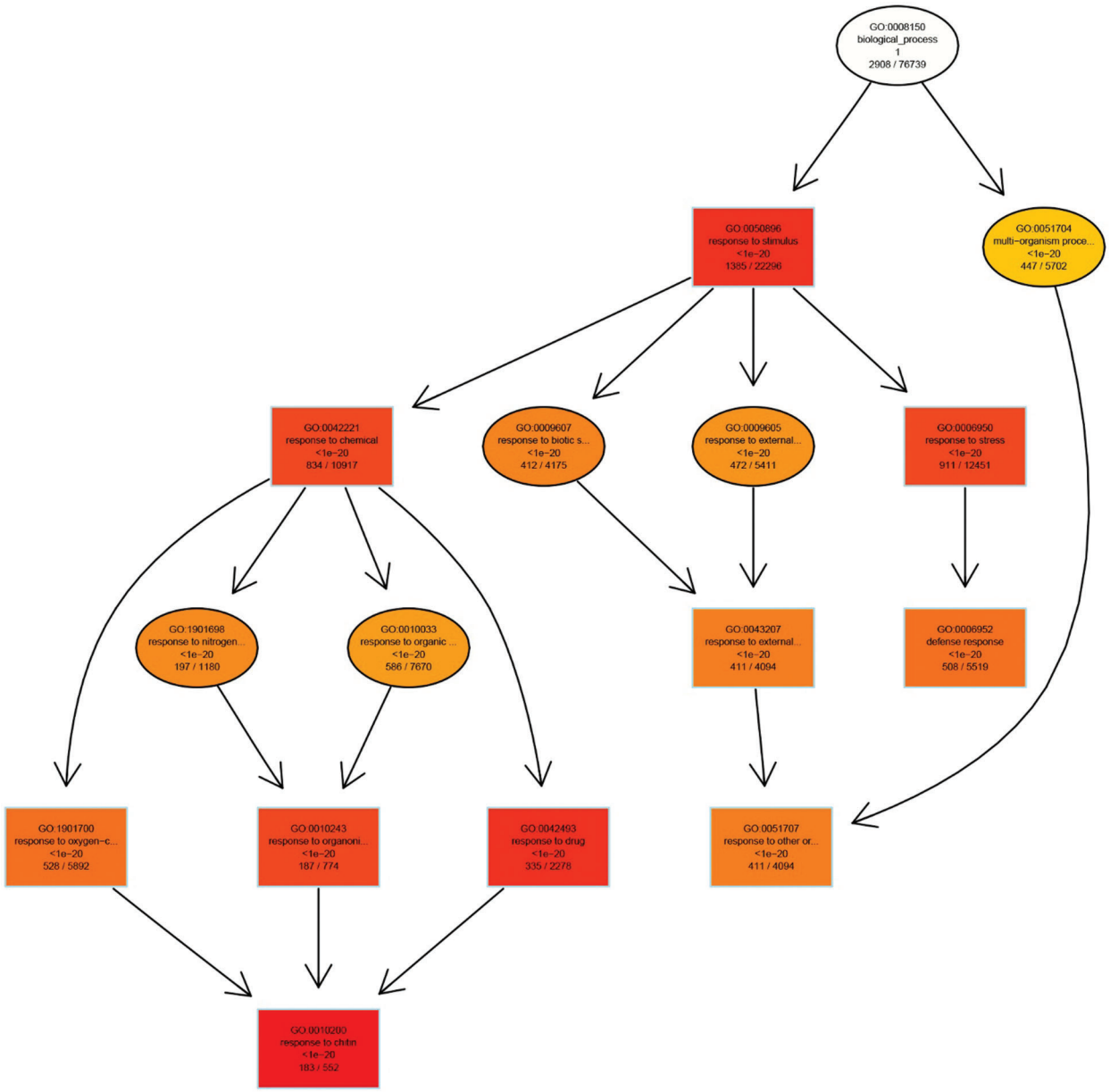

Figure 9: The enriched GO terms in the biological process category for genes in the yellow module

which is reflected in the number, length, and distribution of introns and untranslated regions, such as those of the *BnaCWINV1*s, *BnaVINV2-1*s, *BnaVINV2-3*s, *BnaVINV2-2*s, and *BnaVINV2-4*s. Among proteins, the conserved motifs of the BnaCWINVs and BnaVINVs largely overlapped, while the conserved motifs of the BnaCINVs were quite different. Since CWINVs and VINVs are acid INVs, they share many similarities in protein structure [7], but CINVs are neutral alkaline INVs and are structurally unrelated to acid INVs [60].

INVs play important roles in plant growth and development and in biotic and abiotic stress responses [14,21]. In fact, CWINV activity increases under cold and heat stress [61,62]. Here, we identified 33 LTRs (*cis*-acting elements involved in the low-temperature response) in the promoters of the *BnaINV*s, suggesting these genes play important roles in cold and heat stress responses in rapeseed. Increasing INV

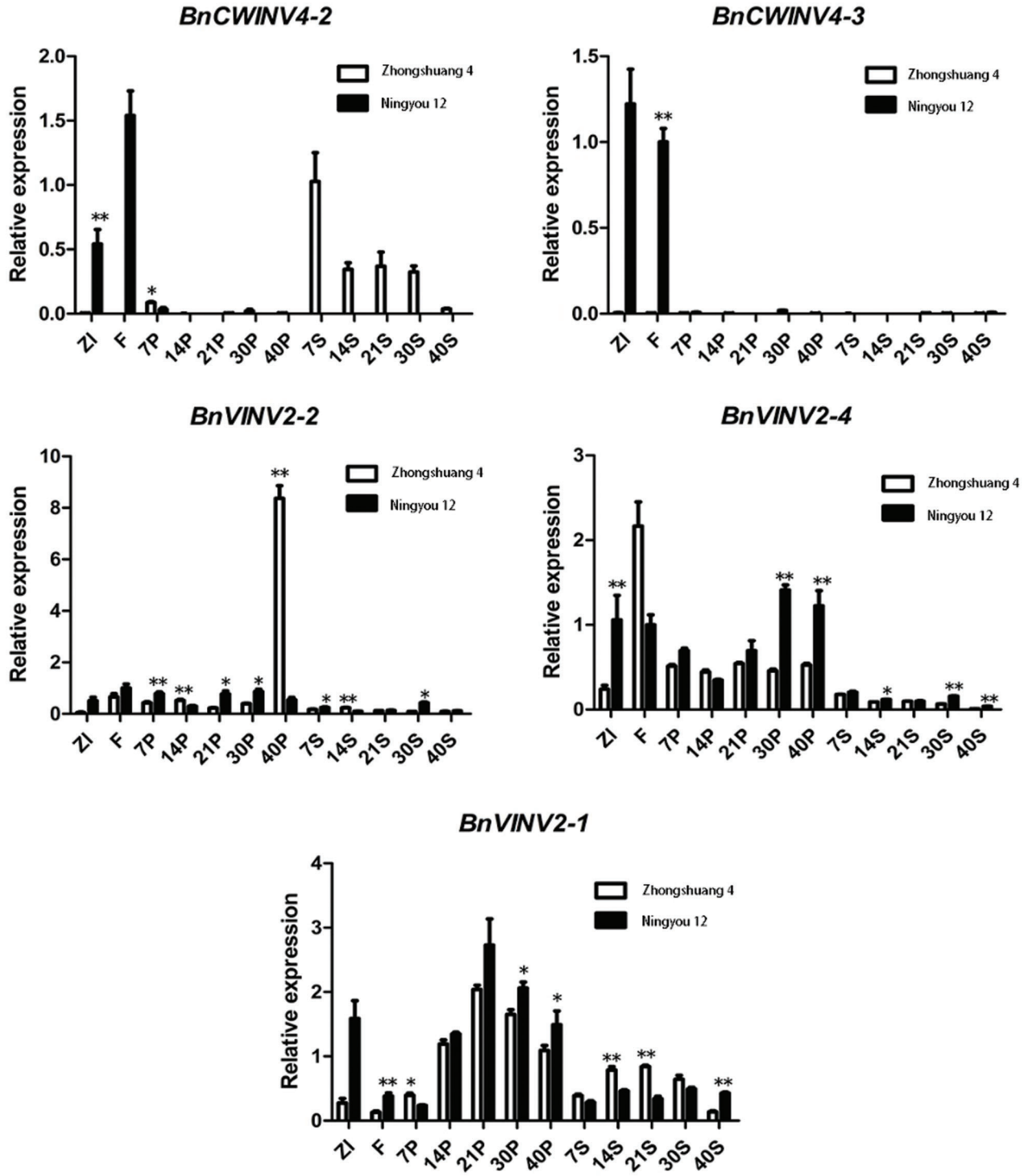

Figure 10: Relative expression levels of selected genes in plant materials with extremely long (Zhongshuang 4) *vs*. short (Ningyou 12) siliques; the average silique length is 9.80 cm in *B. napus* cultivar Zhongshuang 4 and 3.50 cm in Ningyou 12. F, flowers; Zl, buds; 7S, 14S, 21S, 30S, and 40S, seeds collected at 7, 14, 21, 30, and 40 days after flowering, respectively; 7P, 14P, 21P, 30P, and 40P, silique pericarps collected at 7, 14, 21, 30, and 40 days after flowering, respectively. The internal reference gene was *BnUBC21* (EV086936)

activity in tomato cell walls results in inhibited abscisic acid-induced senescence [15]. We also identified 89 ABREs (*cis*-acting elements involved in the abscisic acid response) in the promoters of the *BnaINV*s, pointing to their possible involvement in the abscisic acid signaling pathway. We identified 97 AREs in the promoters of *BnaINV* genes, which may play important roles in maintaining redox balance in plant cells [63]. In *Arabidopsis*, redox balance is important for protecting plants against *S. sclerotiorum* infection [52]. Our RNA-seq analysis indicated that *BnaINV* was upregulated in plants infected with *S. sclerotiorum*, providing new insights into the defense mechanism of *B. napus* against *S. sclerotiorum*. Seed-specific regulatory elements were detected in the promoter of *BnaCINV1-3*, which was expressed during all stages of seed development, suggesting that this gene might function in seed growth and development in *B. napus*.

We used WGCNA to study each *BnaINV* family member, revealing that 14 of the *BnaINV*s were distributed across five different modules. Most of the genes in any given module had similar genetic relationships and expression patterns; for example, *BnaCWINV1-1*, *-2*, *-3*, and *-4*; *BnaCINV1-1* and *-3*; and *BnaVINV2-3* were all grouped into the yellow module, and all of these genes were upregulated following treatment with *S. sclerotiorum*. GO analysis of the *BnaINV*s in the yellow module showed that *BnaCWINV1* is highly correlated with the response to external biological stimuli or biotic stress. Indeed, CWINV plays an important role in carbohydrate acquisition and promotes the defense response during plant-pathogen interactions in tobacco [64]. Our findings provide new insights into the response mechanisms of *B. napus* to pathogen infection. *BnaCWINV4-2* and *BnaCWINV4-3* in the brown module were specifically expressed in main inflorescence buds and flowers at relatively high levels. GO analysis showed that these genes were highly correlated with the formation of the pollen outer wall. In rice, the CWINV gene *OsINV4* is expressed in an anther-specific pattern, and its downregulation results in pollen sterility [62]. *BnaCWINV4-2* and *BnaCWINV4-3* were specifically expressed in flowers and buds and were expressed at significantly higher levels in Ningyou 12 (short siliques) than in Zhongshuang 4 (long siliques). These results suggest that BnaCWINV4-2 and BnaCWINV4-3 influence silique development by affecting bud and flower development, as flowers give rise to siliques.

The expression of *BnaVINV2-2* and *BnaVINV2-4* in the green module was not significantly affected by *S. sclerotiorum* infection; rather, these genes were highly expressed in silique pericarps and stems. The expression levels of *BnaVINV2-2* and *BnaVINV2-4* in silique pericarps significantly differed in Ningyou 12 (short siliques) *vs*. Zhongshuang 4 (long siliques). VINVs play major roles in cell expansion and sugar accumulation [17]. Therefore, *BnaVINV2-2* and *BnaVINV2-4* might play important roles in the elongation of silique pericarps. GO analysis suggested that these genes are related to seed maturation, seed oil body biogenesis, and lipid localization. However, no previous studies have shown that these proteins are related to seed maturation or seed oil body biosynthesis. Our results provide important knowledge that could guide the breeding and production of *B. napus* in the future. The turquoise module contained only one gene, *BnaVINV2-1*, which may functionally differ from the other genes. *BnaVINV2-1* was more highly expressed in silique pericarps than in other tissues. Silique pericarps are important photosynthetic organs in *B. napus*, supplying most of the nutrients needed for silique development [65]. Our GO analysis indicated that *BnaVINV2-1* is functionally involved in photosynthesis, implying that *BnaVINV2-1* plays a role in photosynthesis in the siliques of this species.

In summary, our WGCNA, GO, and RNA-seq analyses suggested that *BnaCWINV4-2*, *BnaCWINV4-3*, *BnaVINV2-1*, *BnaVINV2-2*, and *BnaVINV2-4* may be involved in the growth and development of various tissues in *B. napus*, while *BnaCWINV1-1*, *-2*, *-3*, and *-4*; *BnaCINV1-1* and *-3*; and *BnaVINV2-3* may play important roles in plant defense against pathogens. These findings provide valuable insights to guide the production of *B. napus* in the future.

Funding Statement: This work was supported by grants from the National Natural Science Foundation of China (31830067), National Key Research and Development Program of China (2018YFD0100504-05), and the “111” Project (B12006). The author of these projects is J.-N.L.

Conflicts of Interest: The authors declare that they have no conflicts of interest to report regarding the present study.

References

1. Andriunas, F. A., Zhang, H. M., Xia, X., Patrick, J. W., Offler, C. E. (2013). Intersection of transfer cells with phloem biology—broad evolutionary trends, function, and induction. *Frontiers in Plant Science, 4(221),* 221. DOI 10.3389/fpls.2013.00221.
2. Winter, H., Huber, S. C. (2000). Regulation of sucrose metabolism in higher plants: localization and regulation of activity of key enzymes. *Critical Reviews in Biochemistry and Molecular Biology, 35(4),* 253–289. DOI 10.1080/10409230008984165.
3. Zhang, Y. L., Zhang, A. H., Jiang, J. (2013). Gene expression patterns of invertase gene families and modulation of the inhibitor gene in tomato sucrose metabolism. *Genetics & Molecular Research, 12(3),* 3412–3420.
4. Chandra, A., Jain, R., Solomon, S. (2012). Complexities of invertases controlling sucrose accumulation and retention in sugarcane. *Current Science, 102(6),* 857–866.
5. Bocock, P. N., Morse, A. M., Dervinis, C., Davis, J. M. (2008). Evolution and diversity of invertase genes in *Populus trichocarpa*. *Planta, 227(3),* 565–576. DOI 10.1007/s00425-007-0639-3.
6. Bracho, G. E., Whitaker, J. R. (1990). Purification and partial characterization of potato (*Solanum tuberosum*) invertase and its endogenous proteinaceous inhibitor. *Plant Physiology, 92(2),* 386–394. DOI 10.1104/pp.92.2.386.
7. Sturm, A. (1999). Invertases. Primary structures, functions, and roles in plant development and sucrose partitioning. *Plant Physiology, 121(1),* 1–7. DOI 10.1104/pp.121.1.1.
8. Roitsch, T., Gonzalez, M. C. (2004). Function and regulation of plant invertases: sweet sensations. *Trends in Plant Science, 9(12),* 606–613. DOI 10.1016/j.tplants.2004.10.009.
9. Jung-Il, C., Sang-Kyu, L., Seho, K., He-Kyung, K., Sung-Hoon, J. et al. (2005). Molecular cloning and expression analysis of the cell-wall invertase gene family in rice (*Oryza sativa* L.). *Plant Cell Reports, 24(4),* 225–236. DOI 10.1007/s00299-004-0910-z.
10. Kim, J. Y., Mahé, A., Guy, S., Brangeon, J., Roche, O. et al. (2000). Characterization of two members of the maize gene family, Incw3 and Incw4, encoding cell-wall invertases. *Gene, 245(1),* 89–102. DOI 10.1016/S0378-1119(00)00034-2.
11. Chen, Y. C., Yang, C. S., Yang, C. C., Wang, A. Y., Sung, H. Y. et al. (2009). Insights into the catalytic properties of bamboo vacuolar invertase through mutational analysis of active site residues. *Phytochemistry, 70(1),* 25–31. DOI 10.1016/j.phytochem.2008.10.004.
12. Lammens, W., Le Roy, K., Schroeven, L., Van Laere, A., Rabijns, A. et al. (2009). Structural insights into glycoside hydrolase family 32 and 68 enzymes: functional implications. *Journal of Experimental Botany, 60(3),* 727–740. DOI 10.1093/jxb/ern333.
13. Xie, J., Cai, K., Hu, H. X., Jiang, Y. L., Yang, F. et al. (2016). Structural analysis of the catalytic mechanism and substrate specificity of anabaena alkaline invertase InvA reveals a novel glucosidase. *Journal of Biological Chemistry, 291(49),* 25667–25677. DOI 10.1074/jbc.M116.759290.
14. Ruan, Y. L., Jin, Y., Yang, Y. J., Li, G. J., Boyer, J. S. (2010). Sugar input, metabolism, and signaling mediated by invertase: roles in development, yield potential, and response to drought and heat. *Molecular Plant, 3(6),* 942–955. DOI 10.1093/mp/ssq044.
15. Jin, Y., Ni, D. A., Ruan, Y. L. (2009). Posttranslational elevation of cell wall invertase activity by silencing its inhibitor in tomato delays leaf senescence and increases seed weight and fruit hexose level. *Plant Cell, 21(7),* 2072–2089. DOI 10.1105/tpc.108.063719.
16. Hayes, M. A., Angela, F., Dry, I. B. (2010). Involvement of abscisic acid in the coordinated regulation of a stress-inducible hexose transporter (VvHT5) and a cell wall invertase in grapevine in response to biotrophic fungal infection. *Plant Physiology, 153(1),* 211–221. DOI 10.1104/pp.110.154765.
17. Ruan, Y. L. (2014). Sucrose metabolism: gateway to diverse carbon use and sugar signaling. *Annual Review of Plant Biology, 65(1),* 33–67. DOI 10.1146/annurev-arplant-050213-040251.

18. Wang, L., Li, X. R., Lian, H., Ni, D. A., He, Y. K. et al. (2010). Evidence that high activity of vacuolar invertase is required for cotton fiber and *Arabidopsis* root elongation through osmotic dependent and independent pathways, respectively. *Plant Physiology, 154(2),* 744–756. DOI 10.1104/pp.110.162487.
19. Klann, E. M., Hall, B., Bennett, A. B. (1996). Antisense acid invertase (TIV1) gene alters soluble sugar composition and size in transgenic tomato fruit. *Plant Physiology, 112(3),* 1321–1330. DOI 10.1104/pp.112.3.1321.
20. Morey, S. R., Hirose, T., Hashida, Y., Miyao, A., Hirochika, H. et al. (2018). Genetic evidence for the role of a rice vacuolar invertase as a molecular sink strength determinant. *Rice, 11(1),* 6. DOI 10.1186/s12284-018-0201-x.
21. Tauzin, A. S., Giardina, T. (2014). Sucrose and invertases, a part of the plant defense response to the biotic stresses. *Frontiers in Plant Science, 5(5),* 293. DOI 10.3389/fpls.2014.00293.
22. Li, Z., Palmer, W. M., Martin, A. P., Wang, R., Rainsford, F. et al. (2012). High invertase activity in tomato reproductive organs correlates with enhanced sucrose import into, and heat tolerance of, young fruit. *Journal of Experimental Botany, 63(3),* 1155–1166. DOI 10.1093/jxb/err329.
23. Kauffmann, S., Legrand, M., Geoffroy, P., Fritig, B. (1987). Biological function of 'pathogenesis-related' proteins: four PR proteins of tobacco have 1,3-β-glucanase activity. *EMBO Journal, 6(11),* 3209–3212. DOI 10.1002/j.1460-2075.1987.tb02637.x.
24. Loon, L. C. V., Strien, E. A. V. (1999). The families of pathogenesis-related proteins, their activities, and comparative analysis of PR-1 type proteins. *Physiological and Molecular Plant Pathology, 55(2),* 85–97. DOI 10.1006/pmpp.1999.0213.
25. Welham, T., Pike, J., Horst, I., Flemetakis, E., Katinakis, P. et al. (2009). A cytosolic invertase is required for normal growth and cell development in the model legume, *Lotus japonicus*. *Journal of Experimental Botany, 60(12),* 3353–3365. DOI 10.1093/jxb/erp169.
26. Barnes, W. J., Anderson, C. T. (2018). Cytosolic invertases contribute to cellulose biosynthesis and influence carbon partitioning in seedlings of *Arabidopsis thaliana*. *Plant Journal, 94(6),* 956–974. DOI 10.1111/tpj.13909.
27. Lou, Y., Gou, J. Y., Xue, H. W. (2007). PIP5K9, an *Arabidopsis* phosphatidylinositol monophosphate kinase, interacts with a cytosolic invertase to negatively regulate sugar-mediated root growth. *Plant Cell, 19(1),* 163–181. DOI 10.1105/tpc.106.045658.
28. Battaglia, M. E., Martin, M. V., Lechner, L., Gma, M. N. N. L., Salerno, G. L. (2017). The riddle of mitochondrial alkaline/neutral invertases: a novel *Arabidopsis* isoform mainly present in reproductive tissues and involved in root ROS production. *PLoS One, 12(9),* e0185286. DOI 10.1371/journal.pone.0185286.
29. Ji, X., Van den Ende, W., Van Laere, A., Cheng, S., Bennett, J. (2005). Structure, evolution, and expression of the two invertase gene families of rice. *Journal of Molecular Evolution, 60(5),* 615–634. DOI 10.1007/s00239-004-0242-1.
30. Sheila, J. C., Cristal, L. G., Cecilia, M. E. N., Armando, M. S. J., Samuel, T. et al. (2018). Genome-wide analysis of the invertase gene family from maize. *Plant Molecular Biology, 97(4-5),* 385–406.
31. Shen, L. B., Yuan, Y., Huang, H., Qin, Y. L., Liu, Z. J. et al. (2018). Genome-wide identification, expression, and functional analysis of the alkaline/neutral invertase gene family in pepper. *International Journal of Molecular Sciences, 19(1),* 224. DOI 10.3390/ijms19010224.
32. Chen, Z., Gao, K., Su, X., Rao, P., An, X. et al. (2015). Genome-wide identification of the invertase gene family in populus. *PLoS One, 10(9),* e0138540. DOI 10.1371/journal.pone.0138540.
33. Wang, L., Zheng, Y., Ding, S., Zhang, Q., Chen, Y. et al. (2017). Molecular cloning, structure, phylogeny and expression analysis of the invertase gene family in sugarcane. *BMC Plant Biology, 17(1),* 18903. DOI 10.1186/s12870-017-1052-0.
34. Mun, J. H., Kwon, S. J., Yang, T. J., Seol, Y. J, Jin, M. et al. (2009). Genome-wide comparative analysis of the *Brassica rapagene* space reveals genome shrinkage and differential loss of duplicated genes after whole genome triplication. *Genome Biology, 10(10),* R111.
35. Mahoonak, A. S., Swamylingappa, B. (2007). A new method for preparation of non-toxic, functional protein hydrolysate from commercial mustard cake. *12th International Rapeseed Congress, 1,* 1728–1731.

36. Yon, R. S., William, B., Berardini, T. Z., Chen, G., David, D. et al. (2003). The *Arabidopsis* Information Resource (TAIR): a model organism database providing a centralized, curated gateway to Arabidopsis biology, research materials and community. *Nucleic Acids Research, 1,* 1.
37. Cheng, F., Liu, S., Wu, J., Fang, L., Sun, S. et al. (2011). BRAD, the genetics and genomics database for Brassica plants, *11(1),* 136–130.
38. Altschul, S. F., Madden, T. L., Schäffer, A. A., Zhang, J., Zhang, Z. et al. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Research, 25(17),* 3389–3402. DOI 10.1093/nar/25.17.3389.
39. Salvador, C. G., Silla-Martínez, J. M., Toni, G. (2009). trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics, 25(15),* 1972–1973. DOI 10.1093/bioinformatics/btp348.
40. Xia, X., Xie, Z. (2001). DAMBE: software package for data analysis in molecular biology and evolution. *Journal of Heredity, 30(7),* 1720–1728.
41. Guindon, S., Dufayard, J. F., Lefort, V., Anisimova, M., Hordijk, W. et al. (2010). New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Systematic Biology, 59(3),* 307–321.
42. Kumar, S., Stecher, G., Tamura, K. (2016). MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Molecular Biology & Evolution, 33(7),* 1870–1874. DOI 10.1093/molbev/msw054.
43. Gasteiger, E., Gattiker, A., Hoogland, C., Ivanyi, I., Appel, R. D. et al. (2003). ExPASy: the proteomics server for in-depth protein knowledge and analysis. *Nucleic Acids Research, 31(13),* 3784–3788.
44. Voorrips, R. E. (2002). MapChart: software for the graphical presentation of linkage maps and QTLs. *Journal of Heredity, 93(1),* 77–78. DOI 10.1093/jhered/93.1.77.
45. Guo, A. Y., Zhu, Q. H., Chen, X., Luo, J. C. (2007). GSDS: a gene structure display server. *Hereditas, 29(8),* 1023–1026. DOI 10.1360/yc-007-1023.
46. Bailey, T. L., Nadya, W., Chris, M., Li, W. W. (2006). MEME: discovering and analyzing DNA and protein sequence motifs. *Nucleic Acids Research, 34(Web Server),* W369–W373. DOI 10.1093/nar/gkl198.
47. Quevillon, E., Silventoinen, V., Pillai, S., Harte, N., Mulder, N. et al. (2005). InterProScan: protein domains identifier. *Nucleic Acids Research, 33(Web Server),* W116–W120. DOI 10.1093/nar/gki442.
48. Chalhoub, B., Denoeud, F., Liu, S., Parkin, I. A. P., Tang, H. et al. (2014). Early allopolyploid evolution in the post-Neolithic *Brassica napus* oilseed genome. *Science, 345(6199),* 950–953. DOI 10.1126/science.1253435.
49. Magali, L., Patrice, D., Gert, T., Kathleen, M., Yves, M. et al. (2002). PlantCARE, a database of plant cis-acting regulatory elements and a portal to tools for in silico analysis of promoter sequences. *Nucleic Acids Research, 30 (1),* 325–327. DOI 10.1093/nar/30.1.325.
50. O'Connor, B. P. (2000). SPSS and SAS programs for determining the number of components using parallel analysis and Velicer's MAP test. *Behavior Research Methods, Instruments & Computers, 32(3),* 396–402. DOI 10.3758/BF03200807.
51. Lu, K., Xiao, Z., Jian, H., Peng, L., Qu, C. et al. (2016). A combination of genome-wide association and transcriptome analysis reveals candidate genes controlling harvest index-related traits in *Brassica napus*. *Scientific Reports, 6(1),* 36452. DOI 10.1038/srep36452.
52. Girard, I. J., Tong, C., Becker, M. G., Mao, X., Huang, J. et al. (2017). RNA sequencing of *Brassica napus* reveals cellular redox control of Sclerotinia infection. *Journal of Experimental Botany, 68(18),* 5079–5091. DOI 10.1093/jxb/erx338.
53. Lalitha, S. (2000). Primer Premier 5. *Biotech Software & Internet Report, 1(6),* 270–272. DOI 10.1089/152791600459894.
54. Pan, T. C., Sasaki, T., Zhang, R. Z., Fässler, R., Chu, M. L. (1994). Structure and expression of fibulin-2, a novel extracellular matrix protein with multiple EGF-like repeats and consensus motifs for calcium binding. *Journal of Cell Biology, 123(5),* 1269–1277. DOI 10.1083/jcb.123.5.1269.
55. Li, D., Dye, T. D. (2013). Power and stability properties of resampling-based multiple testing procedures with applications to gene oncology studies. *Computational and Mathematical Methods in Medicine, 23,* 610297.

56. Poole, R. L. (2007). The TAIR database. *Methods in Molecular Biology, 406,* 179–212.
57. Zhao, J., Buchwaldt, L., Rimmer, S. R., Sharpe, A., Mcgregor, L. et al. (2010). Patterns of differential gene expression in *Brassica napus* cultivars infected with *Sclerotinia sclerotiorum*. *Molecular Plant Pathology, 10 (5),* 635–649. DOI 10.1111/j.1364-3703.2009.00558.x.
58. Zhang, J., Li, G., Li, H., Pu, X., Jiang, J. et al. (2015). Transcriptome analysis of interspecific hybrid between *Brassica napus* and *Brassica rapa* reveals heterosis for oil rape improvement. *International Journal of Genomics, 2015,* 1–11.
59. Cheng, F., Wu, J., Wang, X. (2014). Genome triplication drove the diversification of *Brassica* plants. *Horticulture Research, 1,* 115–120.
60. Wan, H., Wu, L., Yang, Y., Zhou, G., Ruan, Y. L. (2018). Evolution of sucrose metabolism: the dichotomy of invertases and beyond. *Trends in Plant Science, 23(2),* 163–177. DOI 10.1016/j.tplants.2017.11.001.
61. Xu, X. X., Qin, H., Yang, W. N., Ye, J. (2017). The roles of cell wall invertase inhibitor in regulating chilling tolerance in tomato. *BMC Plant Biology, 17(1),* 195. DOI 10.1186/s12870-017-1145-9.
62. Oliver, S. N., Dongen, J. T. V., Alfred, S. C., Mamun, E. A., Zhao, X. et al. (2005). Cold-induced repression of the rice anther-specific cell wall invertase gene OSINV4 is correlated with sucrose accumulation and pollen sterility. *Plant, Cell and Environment, 28(12),* 1534–1551. DOI 10.1111/j.1365-3040.2005.01390.x.
63. Liu, Y., Offler, C. E., Ruan, Y. L. (2016). Cell wall invertase promotes fruit set under heat stress by suppressing ROS-independent plant cell death. *Plant Physiology, 172(1),* 163–180. DOI 10.1104/pp.16.00959.
64. Jutta, E., Ina, S. T., Hardy, S. N., Sophia, S., Engelbert, W. et al. (2008). RNA interference-mediated repression of cell wall invertase impairs defense in source leaves of tobacco. *Plant Physiology, 147(3),* 1288–1299. DOI 10.1104/pp.108.121418.
65. Gammelvind, L. H., Schjoerring, J. K., Mogensen, V. O., Jensen, C. R., Bock, J. G. H. (1996). Photosynthesis in leaves and siliques of winter oilseed rape (*Brassica napus* L.). *Plant and Soil, 186(2),* 227–236. DOI 10.1007/BF02415518.