


RESEARCH ARTICLE

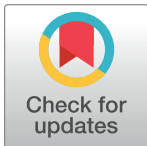
Genome-wide identification and characterization, phylogenetic comparison and expression profiles of *SPL* transcription factor family in *B. juncea* (Cruciferae)

Jian Gao^{1,2}✉, Hua Peng³✉, Fabo Chen^{1,2}, Yi Liu¹, Baowei Chen¹, Wenbo Li^{1,2}* 

1 Department of Life Sciences and Technology, Yangtze Normal University, Fuling, Chongqing, China, **2** Centre for Green Development and Collaborative Innovation in Wuling Mountain Region, Yangtze Normal University, Fuling, Chongqing, China, **3** Sichuan Tourism College, Chengdu, Sichuan, China

✉ These authors contributed equally to this work.

* liverb@163.com



Abstract

SQUAMOSA promoter-binding protein-like (SPL), as plant specific transcription factors, is involved in many plant growth and development processes. However, there is less systematic study for *SPL* transcription factor in *B. juncea* (Cruciferae). Here, a total of 59 *SPL* genes classified into eight phylogenetic groups were identified in *B. juncea*, highly conserved within each ortholog were also found based on gene structure, conserved motif, as well as clustering level. In addition, clustering of *SPL* domain showed that two zinc finger-like structures and NLS segments were identified in almost of *BjuSPLs*. Analyzed of putative cis-elements for *BjuSPLs* demonstrated that *SPL* transcription factors were involved in adverse environmental changes, such as light, plant stresses and phytohormones response. Expression analysis showed that differentially expressed *SPL* genes were identified in flower and stem development of Cruciferae; such as *BjuSPL3a-B*, *BjuSPL2b_B* and *BjuSPL2c_A* were significantly expressed in flower; *BjuSPL3b_B* and *BjuSPL10a_A* were significantly expressed in stem node (VP: vegetative period). Moreover, 28 of the 59 *BjuSPLs* were found involved in their posttranscriptional regulation targeted by *miR156*. We demonstrated that *miR156* negatively regulated *BjuSPL10a_A* and *BjuSPL3b_B* to act for stem development in *B. juncea*.

OPEN ACCESS

Citation: Gao J, Peng H, Chen F, Liu Y, Chen B, Li W (2019) Genome-wide identification and characterization, phylogenetic comparison and expression profiles of *SPL* transcription factor family in *B. juncea* (Cruciferae). PLoS ONE 14(11): e0224704. <https://doi.org/10.1371/journal.pone.0224704>

Editor: Maoteng Li, Huazhong University of Science and Technology, CHINA

Received: July 11, 2019

Accepted: October 18, 2019

Published: November 5, 2019

Copyright: © 2019 Gao et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All relevant data are within the paper and its Supporting Information files.

Funding: The author received no specific funding for this work.

Competing interests: The authors have declared that no competing interests exist.

Introduction

Plant SQUAMOSA promoter-binding protein-like (SPL), acts as fundamental roles in plant growth and development, and are defined by the SBP domains (for SQUAMOSA-PROMOTER BINDING PROTEIN) that have a highly conserved regions of 76 amino acids in length. *SPL* contains a conserved SBP (SQUAMOSA promoter binding protein) domain having two zinc-binding sites consisted of Cys-Cys-Cys-His and Cys-Cys-His-Cys, respectively, and a nuclear localization signal (NLS) motif located at the C-terminal of the SBP domain [1].

Researches have reported that the SBP domain can act as both nuclear import and sequence-specific DNA binding to a consensus-binding site through GTAC core motif and gene-specific flanking regions [2,3]. *SPL* genes are widely distributed in gymnosperms, mosses, single-celled green algae and angiosperms [4], *SPL* genes were firstly detected in *Antirrhinum majus* that were involved in the control of early flower development by binding to the promoter of the floral meristem identity gene *SQUAMOSA* (*SQUA*).

In recent years, numerous studies have characterized *SPL/SBP-box* gene family based on genome-wide identification in many plant species [5–19]. Sixteen *SBP-LIKE* (*SPL*) genes were clustered into eight clades following their conserved SBP domain in *Arabidopsis* [1]. Of those, 10 *SPL* genes belong to five clades are targeted by *miR156* [2, 3], which might highly conserved and acted as important regulators in vegetative phase change in plants [4, 5].

Moreover, *SPL* genes played vital regulatory roles in diversified plant developmental processes in different plant species, such as vegetative phase change [6], male fertility [7], GA biosynthesis [8], plant architecture [9] and response to stress [9] in plant. However, genome-wide analysis of *SPL* genes is scarcely in *B. juncea* owing to lack of genome information, compared with other plant species.

B. juncea (Cruciferae), the tumorous stem mustard, was famous as the raw material for Fuling mustard in China. Currently, the genetics breeding, physiology, biochemistry and classification of *B. juncea* (Cruciferae) have been extensively studied, but little work has been done at the molecular level. Transcription factors *SPL* and their regulation roles in *B. juncea* are poorly understood. Here, genome-wide analysis and molecular dissection of the *BjuSPL* gene family were performed in this study, together with their chromosomal locations. In addition, we constructed phylogenetic tree for *SPL* gene family collected from *A. thaliana*, *B. juncea* and *B. napa*. Moreover, cis-acting elements, conserved motif, gene structure and expression pattern of all identified *BjuSPL* genes in *B. juncea* were also systematically analyzed. This study will give insight into the structure and evolution of the *SPL* genes family in *B. juncea*, thereby further helpful for molecular function analysis of those *SPL* genes in the future.

Methods

Sequence sources

Brassica juncea sequences were kindly provided from Jinhua Yang of Zhejiang University. We retrieved the *SPL* protein sequence of *A. thaliana* and *B. napa* from the Plant Transcription Factor Databases 55 (Plant TFDB v4.0, planttfdb.cbi.pku.edu.cn/) [10], together with their General Feature Format (GFF) file were obtained from Arabidopsis Information Resource (TAIR release 10, <http://www.arabidopsis.org>), and (http://www.genoscope.cns.fr/brassicapapus/data/Brassica_napus.annotation_v5.gff3.gz), respectively. Name and gene IDs of 17 and 59 known *SPL* genes in *A. thaliana* and *B. napa* were shown in [S1 Table](#), respectively.

Identification and distribution of *SPL* transcription factor family in *B. juncea*

SBP domain (PF03110) for *SPL* transcription factor was downloaded from Pfam [11] (Pfam; <http://pfam.xfam.org/>), thereby exploiting for the identification of all possible *SPL* genes from *B. juncea* using HMMER (v3.1b2)[12] (<http://hmmmer.org>) through the e-value <1e-10. All non-redundant hits with expected values were collected and then compared with the *SPL* family in PlnTFDB (<http://plntfdb.bio.uni-potsdam.de/v3.0/>) and PlantTFDB (<http://planttfdb.cbi.pku.edu.cn>). After that, each putative *SPL* gene was confirmed to the presence of SBP domain using SMART [13] (<http://smart.embl-heidelberg.de/>), CDD [14] (<http://www.ncbi>

nslm.nih.gov/Structure/cdd/wrpsb.cgi) and Inter-ProSca (<http://www.ebi.ac.uk/Tools/Inter-ProScan/>). The theoretical MW (molecular weight) and PI (isoelectric point) of *BjuSPLs* were investigated through Expasy57 (<http://web.expasy.org/protparam/>). We fetched the physical location information of the each *SPL* from the corresponding GFF files. Subsequently, we used mg2c (http://mg2c.iask.in/mg2c_v2.0/) to depict their distribution in each *B. juncea* chromosome. In addition, we used BLAST-searching [15] for *BjuSPLs* against each other to identify the duplicated *BjuSPLs* genes, which were defined when both their identity and query coverage was > 80% of their partner sequence [16]. Tandem-duplicated genes were identified as an array of two or more homologous genes within a distance of 100 kb. A chromosome region containing more than two genes within 200 kb was defined as a gene cluster [17]. Moreover, comparison and annotation of orthologous gene clusters among *B. juncea* and *B. napa* were conducted using orthoVenn software [18].

Phylogenetic analysis of *BjuSPLs*

For phylogenetic analysis of multiple sequence, full-length protein sequences of *BjuSPLs* were performed by ClustalW alignment, and then a phylogenetic tree was constructed with MEGA 7.0 software [14] (<https://mega.nz/>) by the Neighbor-Joining (NJ) method, carried out with 1000 replicates boot-strap test.

Gene structure and conserved motif of *BjuSPLs*

Gene structure display server (GSDS 2.0) [19] program (<http://gsds.cbi.pku.edu.cn/>) was used to illustrate the exon/intron structures of *BjuSPLs* through inputting their GFF files. MEME (suite 4.11.4) [20] (<http://meme-suite.org/>) was used for elucidation of conserved motifs of *BjuSPLs*. MEME was run locally through the parameters as follows: optimum width of motif: 6–250 and maximum number of motifs: 9. In addition, MAFFT version 7 (<http://mafft.cbrc.jp/alignment/server/>) was selected for presenting SBP domain. And we used the weblogo platform to generate sequence logos (<http://weblogo.berkeley.edu/>).

Cis-acting elements analysis of promoter regions from *BjuSPLs*

To investigate the putative cis-acting elements of 59 candidate *BjuSPLs* genes, their promoter sequences (2,500 bp upstream of the initiation codon “ATG”) were extracted from genome sequences of *B. juncea* and then conducted using PlantCARE [21] (http://bioinformatics.psb.ugent.be/webtools/plantcare/html/search_CARE.html).

Sample collection and RNA isolation

The *B. juncea* ‘YA1’, ‘YA2’ and ‘YA3’ were cultivated in a greenhouse at the experimental farm of the Yihe (Yangtze normal university experiment base) in 2018. Firstly, we sowed seeds of *B. juncea* ‘YA1’, ‘YA2’ and ‘YA3’ in sterilized soil for 2 weeks under normal growth conditions (23°C, 16 h light/8 h dark). After that, 2-week-old plants were transferred and kept for 15 days in the cold room (5 ± 1°C, 12 h light/12 h dark) for vernalization treatment. After the vernalization periods, the plants were grown in a normal growth room under normal growth conditions (23°C, 16 h light/8 h dark). At least three independent biological replicates for leaf of seedling stage (SS), leaf and stem of flowing period (FP), leaf and stem of mature period (MP) from *B. juncea* ‘YA1’, ‘YA2’ and ‘YA3’ were collected for RNA-Seq, as well as qRT-PCR for the candidate differential expression *BjuSPLs* with three replicates. Of those, samples comprised of seedling, leaf (vegetable periods), stem nodule (vegetable periods), stem nodule (Flowing periods), flower and legume from ‘YA1’ were collected for qRT-PCR for miRNAs and their related

targets with three replicates. All harvested tissue were immediately frozen in liquid nitrogen and stored at -80°C for qRT-PCR respectively. Subsequently, using the mirVana™ miRNA Isolation Kit (Ambion) and Trizol Reagent (Invitrogen, Nottingham, UK) kit following the manufacturer's instructions, small RNA and total RNA was isolated from each sample.

Expression pattern analysis of candidate *BjuSPLs* in *B. juncea*

For analyze the expression patterns of *BjuSPLs* genes, our private available RNA expression profile data including leaf of Seedling stage (SS), leaf and stem of flowering period (FP), leaf and stem of mature period (MP) from YA1 (yonganxiaoye1), YA2 (yonganxiaoye2) and YA3 (yonganxiaoye3) were selected respectively (S2 Table). We used Log₂ (FPKM) for calculating the expression levels of *BjuSPL* genes and their expression patterns were clustered by hierarchical clustering model and illustrated using homemade R language.

Prediction of *BjuSPLs* targeted by *miR156*

BjuSPLs targeted by *miR156* collected from literature were predicted using psRNATarget [22] (<http://plantgrn.noble.org/psRNATarget/?function.3>), and the parameters were set as follows: with the maximum expectation of 3 and the target accessibility (UPE) of 50.

Validation of candidate *BjuSPLs* in *B. juncea* using real-time qRT-PCR

To monitor the candidate *BjuSPLs* in *B. juncea*, ten candidate *BjuSPLs* (*BjuSPL10b_A*, *BjuSPL10a_A*, *BjuSPL10b_B*, *BjuSPL11a_B*, *BjuSPL11b_B*, *BjuSPL13a_B*, *BjuSPL2a_A*, *BjuSPL2b_B*, *BjuSPL2c_B*, *BjuSPL3c_A*, *BjuSPL7a_B* and *BjuSPL9b_B* H) were selected for qRT-PCR (ABI 7500 real-time PCR System, United States) based on their expression patterns (S2 Table). The primers are designed by Primer 5.0 software for qRT-PCR experiments and *B. juncea* gene (Actin) is used as a standard control (S3 Table). First-strand cDNA synthesis was performed with 1 μg total RNA from 'Yonganxiaoye 1' *B. juncea* using a M-MLV reverse transcriptase (Promega). The amplification programs were performed according to the standard protocol of the ABI7500 system, and conducted in triplicate as mentioned by Jian et al. [23]. The relative quantitative method ($2^{-\Delta\Delta\text{CT}}$) was used to calculate the fold change in the expression levels of target genes [24].

Validate the Candidate miRNAs by qRT-PCR

To validate of putative miRNAs targeted the candidate *BjuSPLs*, *miR156* was selected based on their corresponding *BjuSPLs* comprised of *BjuSPL10a_A*, *BjuSPL10b_B* and *BjuSPL3b_B*. The primers are designed by Primer 5.0 software for qRT-PCR experiments and radish gene (Actin) is used as a standard control (S2 Table). Firstly, using One Step miRNA 1st cDNA Synthesis Kit (Shenggong, Chengdu, China), microRNA reverse transcription reactions were performed and incubated in an Eppendorf Mastercycler (Eppendorf North America, Westbury, NY) for 60 min at 37°C , followed by 5 min at 95°C , and then 4°C until further use. The qRT-PCR reactions were performed in a 10 μL volume containing 1 μL diluted reverse transcription product, 1 μL PCR buffer, 0.2 mM dNTPs, 2.0 U EasyTaq DNA polymerase (TransGen Biotech, Beijing, China), and 0.5 μM specific primer of *miR156* (5'-TTGACAGAAGA GAGTGAGCAC-3) and universal primer (5'-TTACCTAGCGTATCGTTGAC-3) on Eppendorf Mastercycler. The amplification programs were performed according to the standard protocol of the ABI7500 system, and conducted in triplicate as mentioned by Jian et al. [23]. The relative quantitative method ($2^{-\Delta\Delta\text{CT}}$) was used to calculate the fold change in the expression levels of target genes [24].

Results

Genome-wide identification and distribution analysis of the SPL gene family in *B. juncea*

We used SBP domain (PF03110) to search against protein of *B. juncea* using HMMER software to identify SPL genes in *B. juncea* [12]. All non-redundant hits with expected values were collected and then compared with the SPL family in PlnTFDB (<http://plntfdb.bio.uni-potsdam.de/v3.0/>) and PlantTFDB (<http://plantfdb.cbi.pku.edu.cn>). In addition, we verified the presence of the SBP domain for the candidate SPL genes through SMART, Proscia and CDD [13, 14]. Total of 59 SPL genes were identified in *B. juncea* (S4 Table). Of those, we accurately named *BjuSPL* genes following their closest orthologs in *A. thaliana* and we coded their different paralogs as a, b, c, and so on, together with their order of the homologous chromosomes. The results showed that lengths and predicted molecular weight (Mw) of 59 SPL genes encoding protein ranged from 106 to 1,038 amino acids, as well as 12.161 to 114.796 kDa in *B. juncea* respectively (S4 Table). In addition, chromosome location showed that those *BjuSPL* genes were distributed in 17 of 18 chromosomes present in the *B. juncea* genome (A1-A10, B01-B04, B06-B08), as well as 5 contigs comprised of Contig533, Contig431, Contig4056, Contig1788, Contig137_156605_433750). We found that the numbers of *BjuSPL* genes mapped on each chromosome were uneven and ranged from 0 to 8. Of those, 8 *BjuSPL* genes were distributed in B08 chromosomes in *B. juncea*, but we did not detect SPL gene in B05 of chromosome (Fig 1, S4 Table). In other species, segmental duplication, tandem duplication, and polyploidization were identified determined the genomic locations of the SPL gene family. We identified 17 pairs of segmentally duplicated and three pairs of tandem-duplicated SPL genes in the *B. juncea* genome (S1 Fig). In addition, relationship between *BjuSPLs* with its homologs in *B. napus* were analyzed using OrthoVenn2 software, the results showed that the species of SPLs were identified obtained from 26 clusters, of those, 12 orthologous clusters were found shared between *B. juncea* and *B. napus*, expect for 14 single-copy gene clusters (S5 Table).

Phylogenetic analysis of SPL gene family in *B. juncea*

Total of 135 SPLs (comprised of 59 obtained from this research in *B. juncea*, 17 from *A. thaliana* and 59 from *B. napus*) were used to construct a Neighbour-Joining (NJ) phylogenetic tree using MEGA 7.0 software. We found those SPL genes were clustered into 8 sub-groups named (I to VIII), and we found at least one protein was obtained from two species (*A. thaliana* and *B. napus*) for each group, with differently distinguished in *B. juncea* for their SPL orthologs. Mustard SPL6, SPL7, SPL8 and SPL13 were categorized into group III, VII, VIII and IV, respectively. Mustard SPL5 and SPL9 were categorized into group II, SPL3, SPL4 and SPL5 were categorized into group VI. Mustard SPL2, SPL10 and SPL11 were categorized into group I, and mustard SPL1, SPL14 and SPL16 belong to group V. However, SPL12 were not identified in *B. juncea* in this study (Fig 2).

Gene structure and conserved motif analysis of SPLs in *B. juncea*

To demonstrate the structural diversity of the *B. juncea* SPL genes, conserved motif, exon/intron structure and putative *cis*-acting elements from *BjuSPLs* promoters were analyzed. We constructed unrooted NJ tree only using protein sequences of 59 *BjuSPLs* (Fig 3A), as well as their gene structures were analyzed by GSDS 2.0 and displayed in Fig 3B. The results showed that introns number varied from 2 to 12 for those 59 *BjuSPLs*. Of those, 20 *BjuSPLs* genes had 5 introns (5 *BjuSPL2*, 3 *BjuSPL6*, 4 *BjuSPL8*, 2 *BjuSPL9*, 3 *BjuSPL13* and 3 *BjuSPL15*); 6 *BjuSPL10* and 4 *BjuSPL11* had 6 introns, expect for *BjuSPL10b_A*, which contained 10 introns

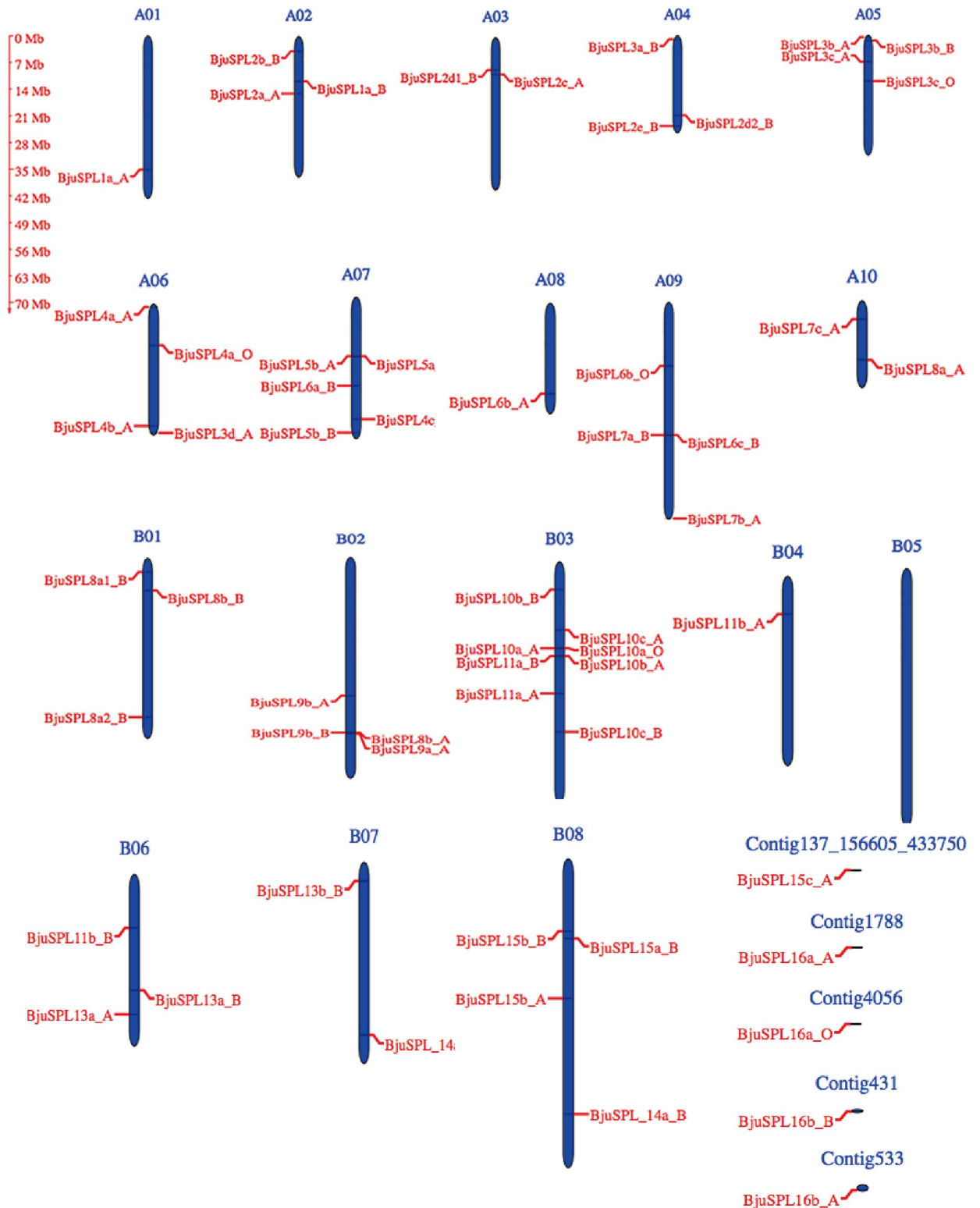


Fig 1. Distribution of SPL genes in the cruciferae genome. Chromosome distribution of SPL genes in *B. juncea* was identified, and the locations of closely linked genes were magnified. The chromosome number is indicated at the top of each chromosome. The scale is in megabases (Mb).

<https://doi.org/10.1371/journal.pone.0224704.g001>

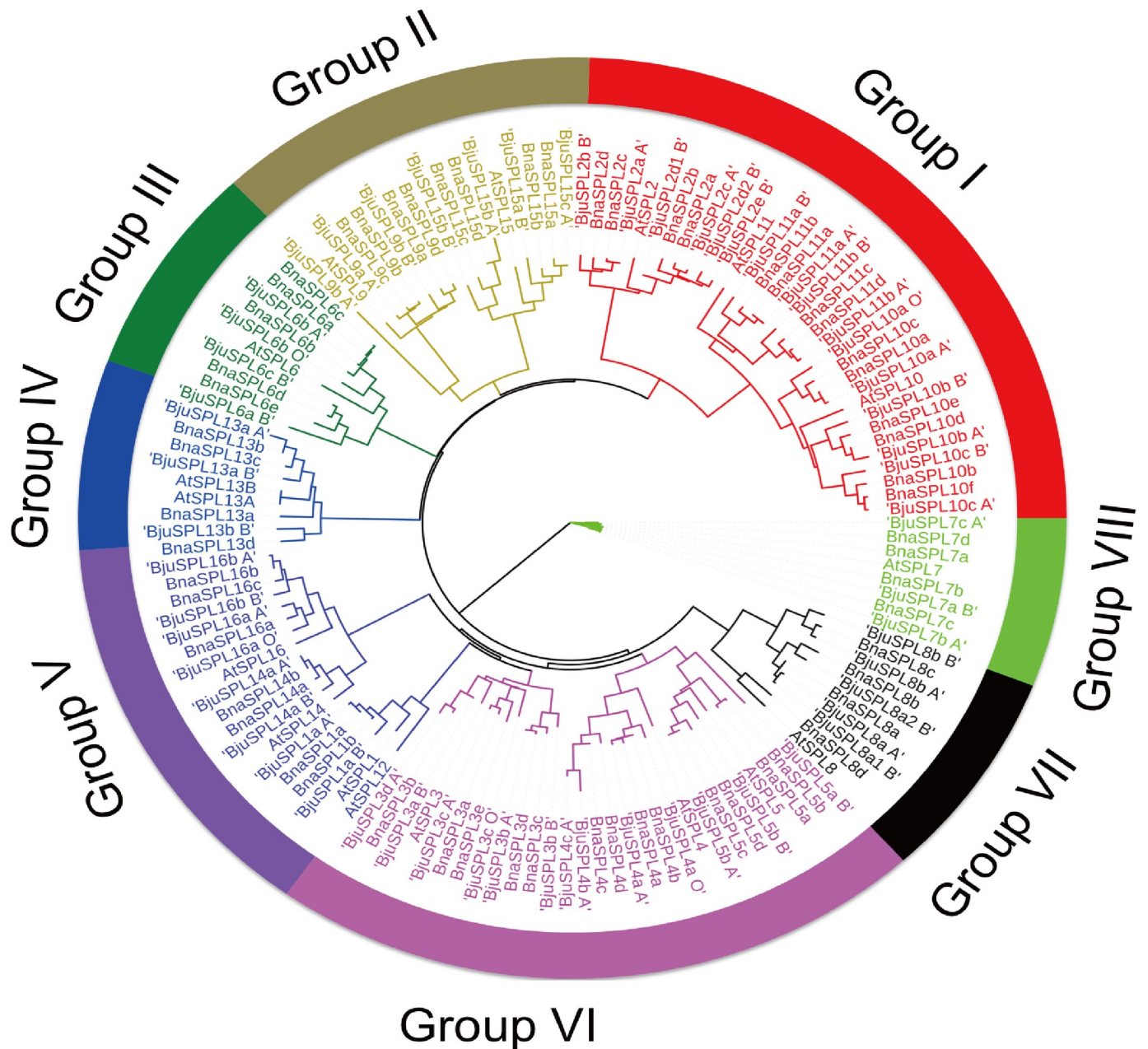


Fig 2. Unrooted phylogenetic tree of the SPL genes from different species (*B. juncea*, *A. thaliana* and *B. napa*). The SBP domain was identified and the phylogenetic tree was constructed using MEGA 7, and the maximum likelihood method with 1000 bootstraps. Different groups of the SBP family are divided with different colors.

<https://doi.org/10.1371/journal.pone.0224704.g002>

owing to segment duplication of some introns; More interestingly, *BjuSPL3*, *BjuSPL4* and *BjuSPL5* contained 2–3 introns, and we found *BjuSPL1*, *BjuSPL14* and *BjuSPL16* contained 12 introns with an ankyrin (ANK) domain along with an SBP domain, suggesting that these proteins represented different functions compared with other groups, which may play a role by interacting with other proteins in plant cells. We revealed that different *BjuSPLs* orthologs were similar to Arabidopsis orthologs and exhibited different exon-intron structures.

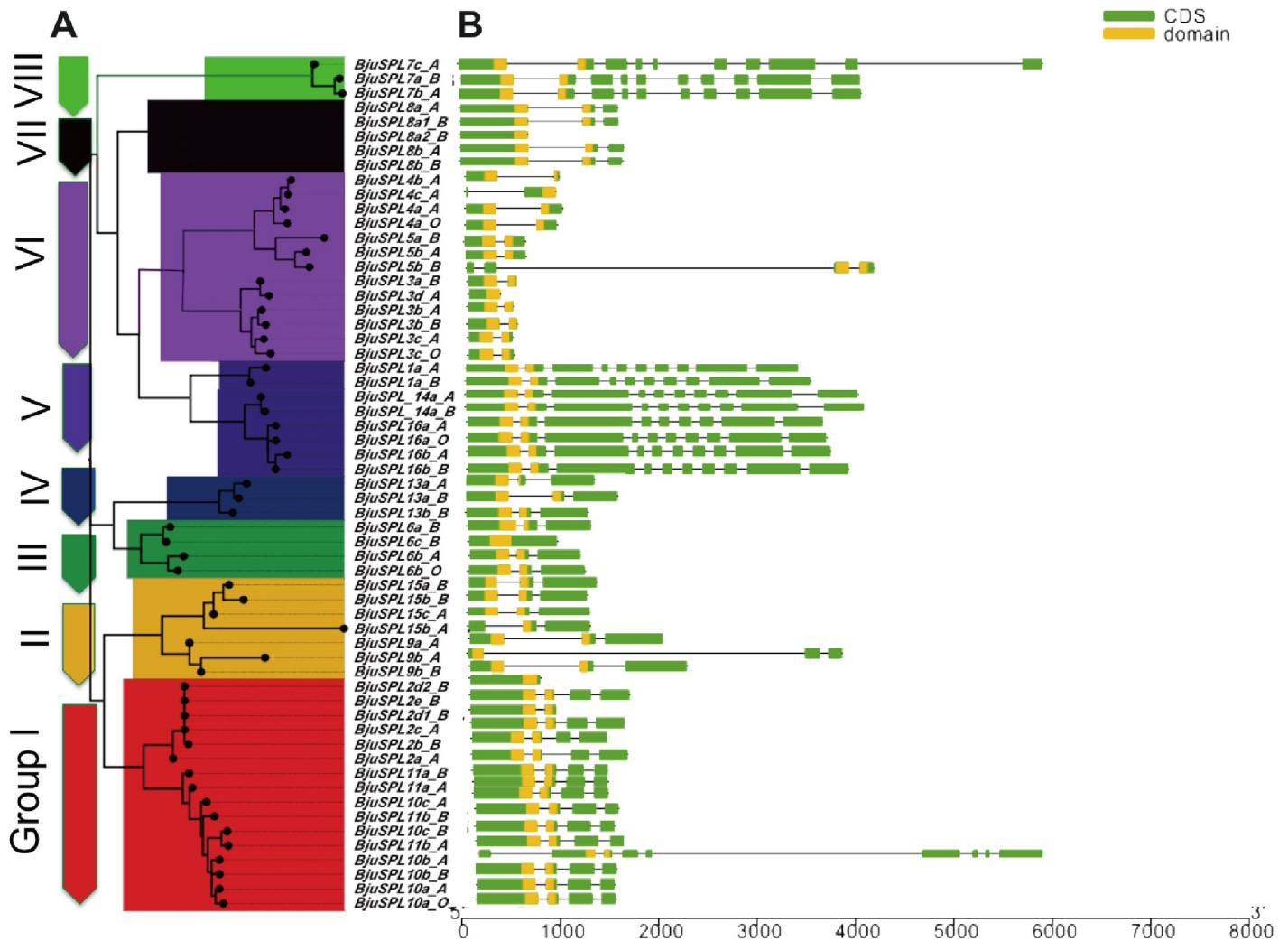


Fig 3. Phylogenetic tree of the SPL genes and genomic organization of SPL genes in *B. juncea*. A: The SBP domain was identified and the phylogenetic tree was constructed using MEGA7, and the maximum likelihood method with 1000 bootstraps. Different groups of the SBP family are divided with different colors. B: Exon/intron structures of *B. juncea* SPL genes. Untranslated 5' and 3' regions and CDS are shown in gray and black. Black lines connecting two CDS represent introns.

<https://doi.org/10.1371/journal.pone.0224704.g003>

Further, protein motifs were identified for *BjuSPLs* through full-length proteins to clarify the characteristics of the SBP domain in *B. Juncea* (S1 Fig). The results showed that 20 motifs with the diversity of sequence structures recognized in *BjuSPLs* proteins among *BjuSPLs*, named motifs 1 to 20 were identified (S2 Fig). The lengths of 29 (motif 7) and 50 amino acids (motif1-6 et.al) were identified in those conserved motifs. The number of the conserved motifs varied from 3 for *BjuSPL2b_B* to 17 for *BjuSPL1a_B* respectively. Of those, we identified that 59 *BjuSPLs* proteins covered two Zn-finger like structure motif, and 56 *BjuSPLs* covered nuclear localization signal motif, except for *BjuSPL4b_A*, *BjuSPL4c_A* and *BjuSPL2_B*. Interestingly, most of *BjuSPL* proteins contained more motifs in SBP domain, expect for *BjuSPL15b_A* (had motif 1, 2 and 15), such as *BjuSPL2* (6–8 motifs), *BjuSPL7* (6–7 motifs), *BjuSPL14* (9–10 motifs) and *BjuSPL1* (11–13 motifs).

Moreover, we presented the SBP domain structures by constructing the multiple alignments of all 59 *BjuSPLs* proteins using MAFFT version 7. The results revealed that two zinc finger-like structures (Cys3His-type, Cys2HisCys), together with NLS segments were

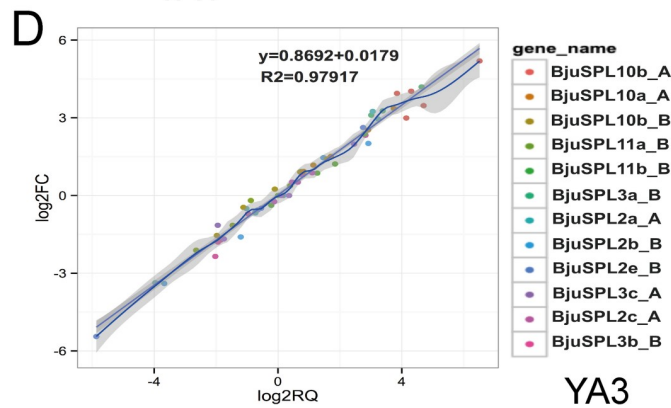
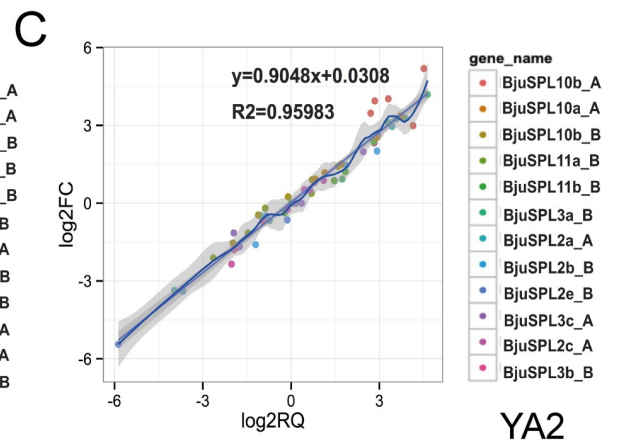
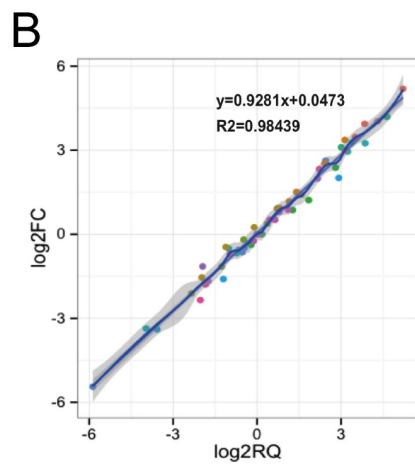
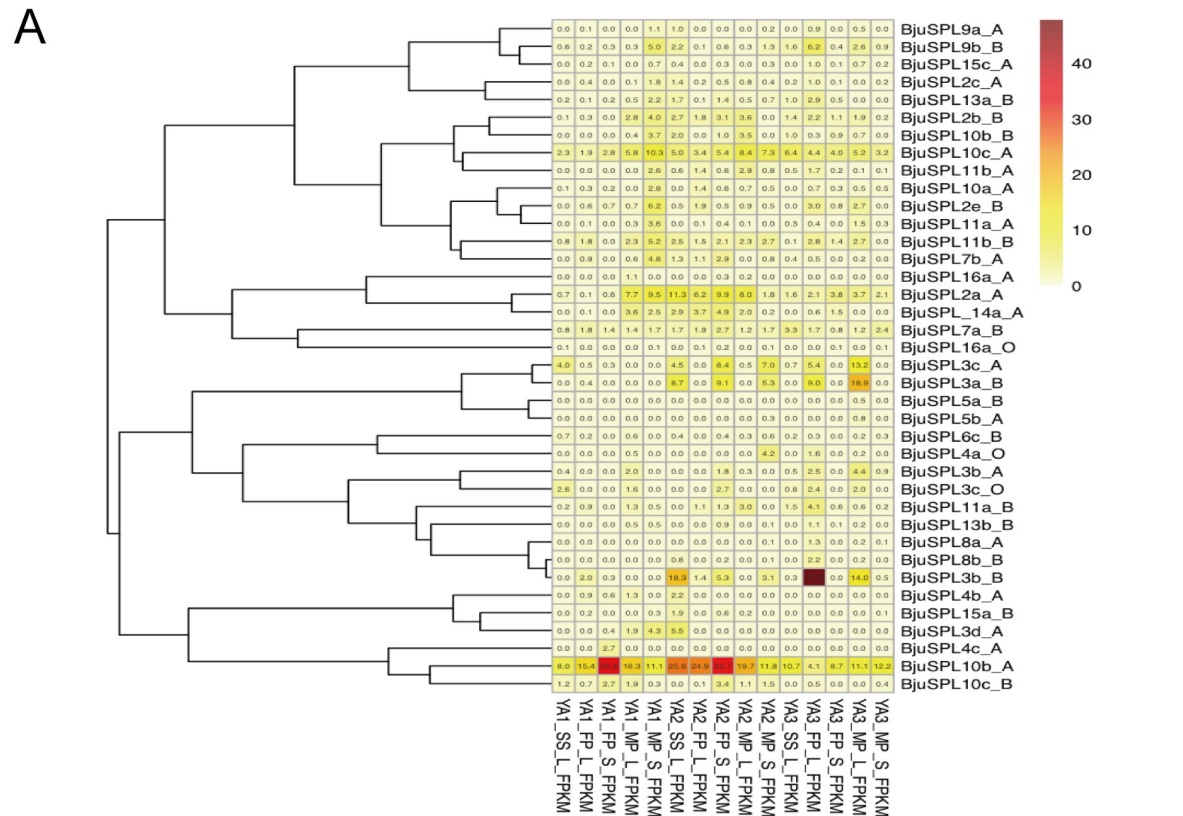


Fig 4. Expression patterns of *BjuSPL* genes and validate of candidate *BjuSPLs* in different tissues and cultivars of *B. juncea*. A. Expression patterns of *BjuSPL* genes in different tissues and cultivars of *B. juncea*. The color represents *BjuSPLs* expression levels: Log₂ (FPKM). The phylogenetic relationship was showed on the left. RNA expression profile data were selected as follows: leaf of seedling stage (SS), leaf and stem of flowing period (FP), leaf and stem of mature period (MP) from YA1 (yonganxiaoye1), YA2 (yonganxiaoye2) and YA3 (yonganxiaoye3) respectively. B. Validate of candidate *BjuSPLs* in 'YA1' using qRT-PCR and then correlation between RNA-seq and qRT-PCR data were conducted. Each RNA-seq expression data was plotted against that from quantitative real-time PCR and fit into a linear regression. Both x- and y-axes were shown in log₂ scale and each colour represented a different gene. C. Validate of candidate *BjuSPLs* in 'YA2' using qRT-PCR and then correlation between RNA-seq and qRT-PCR data were conducted. D. Validate of candidate *BjuSPLs* in 'YA3' using qRT-PCR and then correlation between RNA-seq and qRT-PCR data were conducted.

<https://doi.org/10.1371/journal.pone.0224704.g004>

identified in all *BjuSPLs*, except for three SPLs (*BjuSPL4b_A*, *BjuSPL4c_A* and *BjuSPL2B2_B*) without NLS (S3 Fig). We demonstrated that two Zn-finger structures and NLS section in the sequences SBP domain of *BjuSPLs* were conserved in *B. juncea*, resulted from a similar number and type of SBP motifs identified in each of *BjuSPL* orthologs. We concluded that similar gene structure, motif architecture of *BjuSPL* orthologs were significantly clustered into the same phylogenetic tree.

The *cis*-acting elements analysis of *BjuSPLs* gene promoter regions

To investigate putative functional of 59 *BjuSPL* genes in *B. juncea*, we collected their upstream sequences (2,500 bp upstream of the initiation codon) and then used for *cis*-acting element prediction by PlantCARE. More than 45 types of putative *cis*-elements were identified present in the promoters of *BjuSPLs*, such as light responsive elements, stress responsive elements, phytohormone response element, gibberellin, salicylic acid (TCA-element) and MeJA (S4 Fig, S6 Table). This suggests the most of *BjuSPLs* genes were involved in different biological processes, including response to abiotic stresses (defense and stress) and phytohormones in *B. juncea*.

Expression profiles of *SPL* genes in *B. juncea*

The expression profiles of all the identified 59 *BjuSPLs* were analyzed available from our private available RNA expression profile data. The expression of *BjuSPLs* was constructed by heat map using homemade R and showed in Fig 4. More interestingly, only *BjuSPL10b_A* and *BjuSPL10c_A* were found highly expressed in all tissues comprised of leaf of seedling stage (SS), leaf and stem of flowing period (FP), leaf and stem of mature period (MP) from YA1 (yonganxiaoye1), YA2 (yonganxiaoye2) and YA3 (yonganxiaoye3) respectively. Especially for *BjuSPL10c_A*, which were found highly accumulated in stem of flowing period (FP). In addition, *BjuSPL3c_A*, *BjuSPL3a_B* and *BjuSPL3b_B* were found specifically expressed in leaf of seedling stage (SS), flowing period (FP), mature period (MP) from YA2 (yonganxiaoye2) and YA3 (yonganxiaoye3), expect for YA1 (yonganxiaoye1), especially for *BjuSPL3b_B*. However, other *BjuSPL* genes were no significantly detected in all tested tissues. Expression patterns of *BjuSPL* genes with great alteration showed that *BjuSPL3a-B*, *BjuSPL2b_B* and *BjuSPL2c_A* were significantly expressed in flower; *BjuSPL 3b_B* and *BjuSPL10a_A* were significantly expressed in stem node (VP: vegetative period), but only *BjuSPL10b_B* were found cumulative in seedling (Fig 5).

MiR156-mediated posttranscriptional regulation of *BjuSPLs*

To understand the miR156-mediated posttranscriptional regulation of the *BjuSPL* genes, we searched the coding regions and 3' UTRs of all *BjuSPLs* for the targets of mustard *miR156a-miR156h*. The results showed that 28 *BjuSPL* genes, belonged to all groups expect for *BjuSPL9* that are complementary association with the *Bju-miR156* mature sequences (Fig 6), suggesting that *Bju-miR156* may specifically regulated these *BjuSPL* genes in *B. juncea*. In addition, *Bju-*

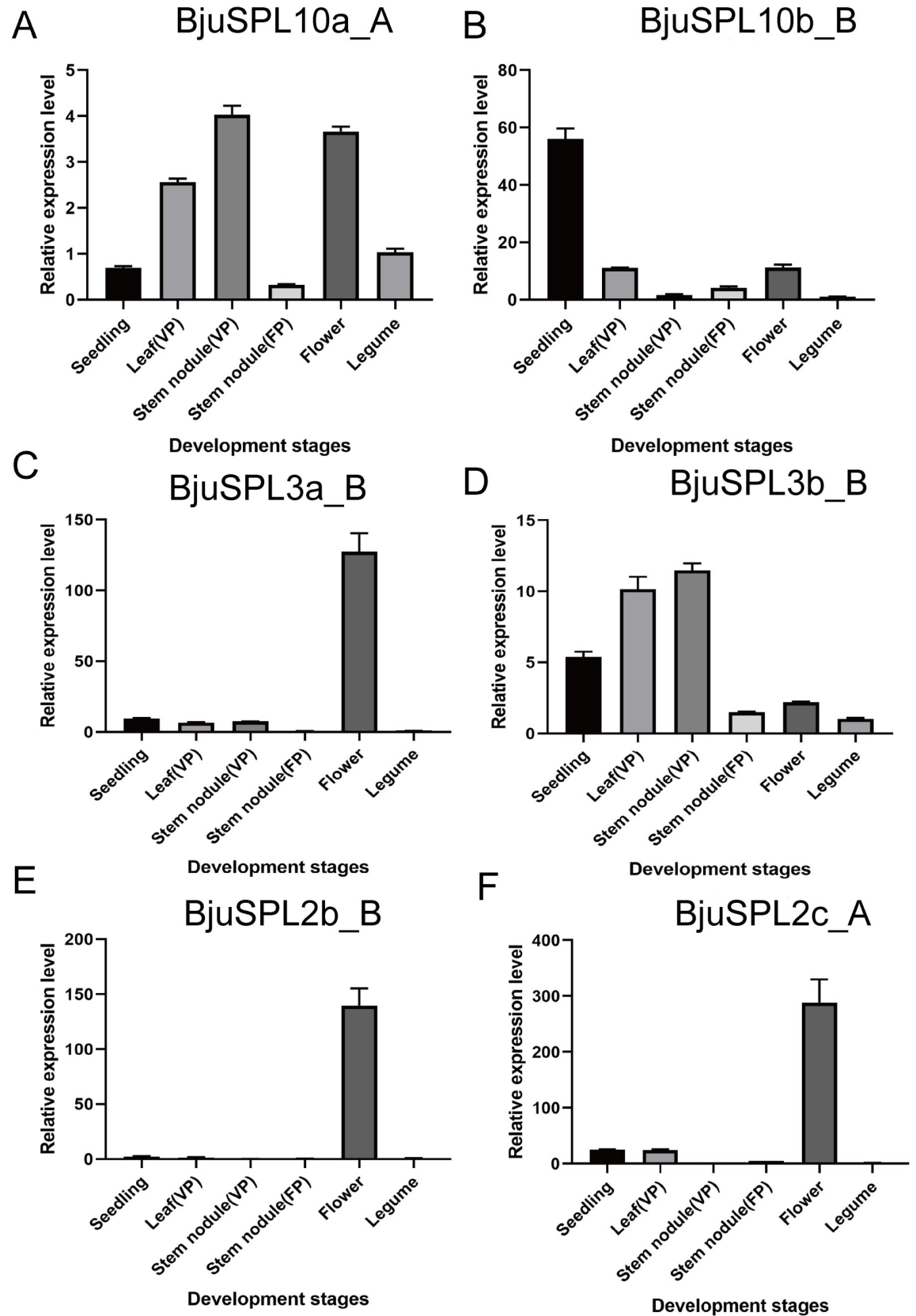


Fig 5. Validated of candidate *BjuSPLs* in the organ development using qRT-PCR based on their great alteration expression levels. From left to right represent seedling, leaf (Vegetative period), stem (Vegetative period), stem (Flower period), flower and legume. Error bars show the standard error calculated from three biological replicates. Values are presented as the mean \pm SEM.

<https://doi.org/10.1371/journal.pone.0224704.g005>

miR156 was found acted on the location in the coding region of *BjuSPL10*, *BjuSPL11* and *BjuSPL2* belong to group I, *BjuSPL1*, *BjuSPL14* and *BjuSPL16* belong to group V and *BjuSPL3*, *BjuSPL4* and *BjuSPL5* belong to group VI, respectively. These results showed that *miR156* conserved in plants by mediated posttranscriptional regulation of the SPLs.

Discussions

In recent years, researchers have focused on *SPL* transcription factors owing to their conservation and specially functions in plants. In addition, the number of *SPLs* have been identified varies from species to species through genome-wide analysis in plants, including 15 *SPLs* in potato [25] and tomato [26], 17–20 in rice and Arabidopsis [2, 27], as well as 27–30 in Chinese cabbage [28]. However, there is no report about *SPL* gene family in *B. juncea* based on genome-wide identification, owing to lack of sequenced genome information. In this research, we firstly reported systematically identification of *SPL* genes and investigation of their functional structure in *B. juncea*. Based on our results, 59 *SPL* genes were identified in *B. juncea* (S1 Table). Here, phylogenetic tree of *SPL* proteins in *B. juncea*, Arabidopsis and *B. napus* demonstrated that all *SPL* genes were categorized into eight groups (I–VIII) (Fig 2). Among 59 mustard *SPLs*, each kind of orthologs from 10 Arabidopsis *SPLs* orthologs was clustered together. In recent years, researches have demonstrated that segmental duplications were existed and play vital roles in the gene expansion of *SPL* transcription factor gene family in many plants [29, 30]. Here, 17 pairs of segmentally duplicated and three pairs of tandem-duplicated *SPL* genes were identified in the *B. juncea* genome. In addition, there were 12 ortholog pairs in syntenic regions between *B. juncea* and *B. napa* (S5 Table).

By conducting GSDS structure of *SPLs* gene in mustard, different gene structures were found in different *SPL* orthologs contained with the diverse number of introns, including 5 introns in *BjuSPL2*, *BjuSPL8*, *BjuSPL9*, *BjuSPL13* and *BjuSPL15*; 6 introns in *BjuSPL10* and *BjuSPL11*; 3 introns in *BjuSPL3*, *BjuSPL4* and *BjuSPL5*, as well as 12 introns in *BjuSPL1*, *BjuSPL14* and *BjuSPL16*. In addition, similar motifs were found shared in different cotton *SPL* orthologs through MEME analyze. 59 *BjuSPL* proteins were identified contained two Zn-finger like structure motif, and 56 *BjuSPLs* contained nuclear localization signal motif, except for *BjuSPL4b_A*, *BjuSPL4c_A* and *BjuSPL2B2_B*. Interestingly, most of *BjuSPL* proteins contained more motifs in SBP domain, expect for *BjuSPL15b_A* (had motif 1, 2 and 15), such as *BjuSPL2* (6–8 motifs), *BjuSPL7* (6–7 motifs), *BjuSPL14* (9–10 motifs) and *BjuSPL1* (11–13 motifs). Moreover, we found *BjuSPL1*, *BjuSPL14* and *BjuSPL16* contained 12 introns with an ankyrin (ANK) domain, suggesting that these proteins represented different functions compared with other groups. We speculated that similar function were probably shared among different *SPL* orthologs in mustard, based on their similar conserved motifs and exon/intron structure, expect for some specially *SPL* gene, such as *BjuSPL1*, *BjuSPL14* and *BjuSPL16* with an ankyrin (ANK) domain, *BjuSPL4b_A*, *BjuSPL4c_A* and *BjuSPL2B2_B* without nuclear localization signal, NLS.

Functional analyzed of cis-acting elements for *BjuSPLs* showed that *SPL* transcription factors may be involved in light, phytohormones and stresses responsiveness. RNA-seq profiles demonstrated that several of putative identified *SPL* genes show a tissue-dependent and development-dependent expression patterns (Fig 4). Of those, we found that *BjuSPL10b_A* and *BjuSPL10c_A* were significantly accumulated in the developmental stage and all tested tissues.

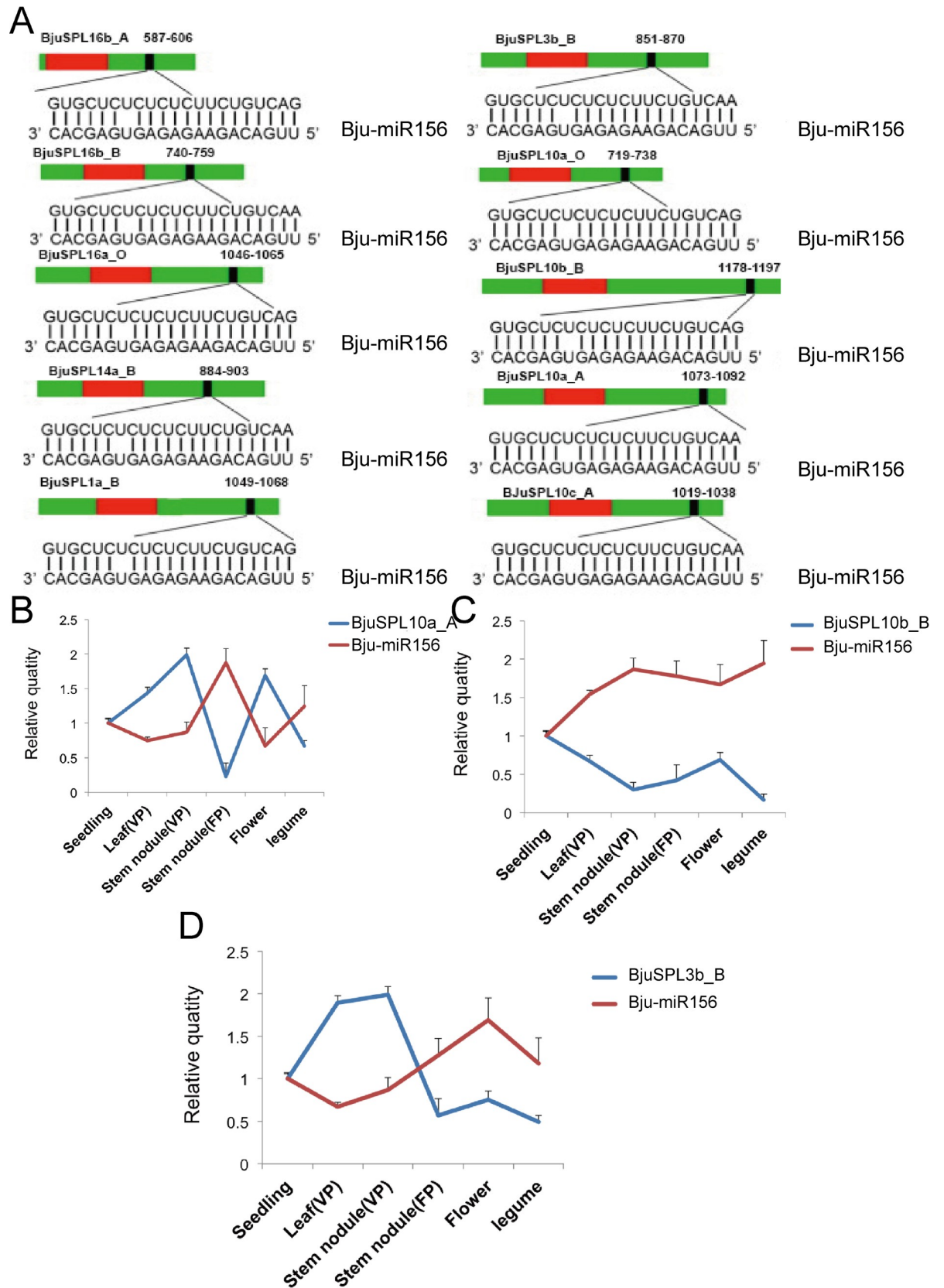


Fig 6. Prediction of *BjuSPLs* regulated by *miR156* and qRT-PCR analysis of *miRNA156* acted in *BjuSPL10a_A*, *BjuSPL10b_B* and *BjuSPL3b_B* in *B. juncea*. A. Predict of *BjuSPLs* regulated by *miR156*. The miRNA target sites (red) with the nucleotide positions of *BjuSPL* transcripts are shown. The RNA sequence of each complementary site from 5' to 3' and the predicted miRNA sequence from 3' to 5' are indicated in the expanded regions. B. qRT-PCR analysis of *miRNA156* acted in *BjuSPL10a_A* in *B. juncea*. Standard error bars are provided for three biological repeats. C. qRT-PCR analysis of *miRNA156* acted in *BjuSPL10b_B* in *B. juncea*. D. qRT-PCR analysis of *miRNA156* acted in *BjuSPL3b_B* in *B. juncea*.

<https://doi.org/10.1371/journal.pone.0224704.g006>

BjuSPL3c_A, *BjuSPL3a_B* and *BjuSPL3b_B* were found specifically expressed in leaf of Seedling stage (SS), flowing period (FP), mature period (MP) from YA2 (yonganxiaoye2) and YA3 (yonganxiaoye3), except for YA1 (yonganxiaoye1), especially for *BjuSPL3b_B*. This result suggests that only several of mustard *SPL* gene family may play vital roles in leaf and stem development, and sub-functionalization, gained new functions and lost functions possibly existed in paralog genes of *SPL* gene family. To date, less reports about expression and functional analysis of *SPL* genes were found in mustard, such as *SPL* gene family were found to be down-regulated in male-sterility lines [31] and *SPL13* was found to be cold-inducible in early stages of development. In addition, the target genes of *miR156* such as *CAT*, *ABCB*, *CBS*, and *SPL* showed their involvement in stress response and temporal expression changes during leaf development in *B. juncea* [32–34].

We proposed that *miR156* might also involve in *SPL*-regulated gene networks in *B. juncea*. The previous showed that 10 *SPLs* in Arabidopsis were reported to be potential targets of *miR156/157* [5], as well as 17 *SPLs* in soybean [35], 18 *SPLs* in Populus [30] and 11 *SPLs* in rice [36]. Here, we also demonstrated that 28 *SPLs* were putative targeted by *miR156*. More importantly, we found that a potential target site (motif 7) was found located in those *SPLs* targeted by the *miR156*. These findings were also identified in rice and Arabidopsis, all *miR156*-targeted *SPLs* were also found *miR156* recognition element located in motif [36]. In Arabidopsis, researches have demonstrated that the most of *AtSPLs* have been acted as diverse function, such as *SPL2*, *SPL9*, *SPL10*, *SPL11*, *SPL13* and *SPL15* were specifically association with shoot development [37], *SPL3*, *SPL4* and *SPL5* primarily control flowering time [38]. *SPL3*, *SPL9* and *SPL10* specifically expressed in lateral root [39]. *SPL1* and *SPL12* acted as important regulator at the reproductive stage [40]. *SPL7* were found involved in the Cu deficiency response [41] and *SPL8* conferred to male fertility, as well as regulated gynoecium differential patterning [7, 42]. In our study, we found *BjuSPL2c_A*, *BjuSPL3a_B* and *BjuSPL2b_B* were found specifically expressed in flower in mustard and *BjuSPL10b_B* targeted by *miR156* were found highly accumulated in seedling, together with *BjuSPL3b_b* and *BjuSPL10a_A* in stem node (VP). More interestingly, we demonstrated that *BjuSPL3b_b* and *BjuSPL10a_A* were negatively regulated by *miR156* that specially involved in stem node (VP) development. However, the detailed expression patterns and function of those tissue-dependent and development-dependent *SPL* transcription factor, as well as their regulators *miR156* remains to be further investigated in mustard. This is firstly report about identification of *SPL* gene family in mustard; it would be helpful for well functional analysis of tissue-dependent and development-dependent *SPL* transcription factor in the future.

Supporting information

S1 Table. Gene name and gene ID of *SPLs* in Arabidopsis, rice and *B. napa*.
(XLSX)

S2 Table. Expression patterns of the *BjuSPL* genes in various tissues, including leaf of seedling stage (SS), leaf and stem of flowing period (FP), leaf and stem of mature period (MP) from YA1 (yonganxiaoye1), YA2 (yonganxiaoye2) and YA3 (yonganxiaoye3) respectively.
(XLSX)

S3 Table. List of primers for qRT-PCR analysis of candidate *BjuSPLs*.
(XLSX)

S4 Table. Characterization of SPL family genes identified in *B. juncea*.
(XLSX)

S5 Table. 12 ortholog pairs in syntenic regions between *B. juncea* and *B. napa* identified using orthoVenn software.
(XLSX)

S6 Table. *Cis*-acting elements identified in the promoter regions of *BjuSPL* genes.
(XLSX)

S1 Fig. Phylogenetic tree of the SPL gene family in *B. juncea* annotated with collinear and tandem relationships. Curves connecting pairs of gene names suggest either the collinear relationship (red) or tandem relationship (blue). This annotated tree is output from 'family tree plotter' of MCscanX software.
(TIF)

S2 Fig. Conserved motifs of *BjuSPLs* proteins identified in this study. The conserved motifs were identified using MEME (suite 4.11.4) based on the protein sequences of *BjuSPLs*, and each motif is indicated with a colored box numbered (1 to 20) at the bottom. Motif 1 and motif 2 represent two Zn-finger like structure and nuclear localization signal (NLS).
(TIF)

S3 Fig. Alignment of the SBP domain in *BjuSPL* proteins. Multiple sequences alignment was performed using MAFFT version 7. Two Zn finger like structure (Cys3His-type, Cys2HisCys) and NLS are indicated. In addition, motif logo and protein sequence of the SBP domain and NLS segment were showed.
(TIF)

S4 Fig. Analyze of *cis*-acting elements related to abiotic stresses and phytohormones response in *BjuSPLs* promoters.
(TIF)

Author Contributions

Conceptualization: Jian Gao, Wenbo Li.

Data curation: Hua Peng, Yi Liu.

Formal analysis: Yi Liu.

Methodology: Baowei Chen.

Project administration: Wenbo Li.

Resources: Fabo Chen.

Software: Jian Gao, Baowei Chen.

Validation: Hua Peng, Fabo Chen.

Writing – original draft: Jian Gao.

Writing – review & editing: Wenbo Li.

References

1. Liu N, Tu L, Wang L, Hu H, Xu J, Zhang X. MicroRNA 157-targeted SPL genes regulate floral organ size and ovule production in cotton. *Bmc Plant Biology*. 2017; 17(1): 7. <https://doi.org/10.1186/s12870-016-0969-z> PMID: 28068913
2. Yang Z, Wang X, Gu S, Hu Z, Hua X, Xu C. Comparative study of SBP-box gene family in Arabidopsis and rice. *Gene*. 2008; 407(1): 1–11.
3. Wang H, Wang H. The miR156/SPL Module, a Regulatory Hub and Versatile Toolbox, Gears up Crops for Enhanced Agronomic Traits. *Molecular Plant*. 2015; 8(5): 677–688. <https://doi.org/10.1016/j.molp.2015.01.008> PMID: 25617719
4. George C, A Mark C, Koy S, Sarah H. The heterochronic maize mutant Corngrass1 results from overexpression of a tandem microRNA. *Nature Genetics*. 2007; 39(4): 544–549. <https://doi.org/10.1038/ng2001> PMID: 17369828
5. Wu G, Poethig RS. Temporal regulation of shoot development in Arabidopsis thaliana by miR156 and its target SPL3. *Development*. 2006; 133(18): 3539–3547. <https://doi.org/10.1242/dev.02521> PMID: 16914499
6. Tsubosa Y, Sato H, Tachimori Y, Hokamura N, Hosokawa M, Kinoshita Y, et al. Functional analysis of the Arabidopsis thaliana SBP-box gene SPL3: a novel gene involved in the floral transition. *Plant Journal*. 2010; 12(2): 367–377.
7. Shuping X, María S, Susanne HH, Rita B, Peter H. miR156-targeted and nontargeted SBP-box transcription factors act in concert to secure male fertility in Arabidopsis. *Plant Cell*. 2010; 22(12): 3935–3950. <https://doi.org/10.1105/tpc.110.079343> PMID: 21177480
8. Zhang Y, Schwarz S, Saedler H, Huijser P. SPL8, a local regulator in a subset of gibberellin-mediated developmental processes in Arabidopsis. *Plant Molecular Biology*. 2007; 63(3): 429. <https://doi.org/10.1007/s11103-006-9099-6> PMID: 17093870
9. Stone JM, Xinwen L, Nekl ER, Stiers JJ. Arabidopsis AtSPL14, a plant-specific SBP-domain transcription factor, participates in plant development and sensitivity to fumonisin B1. *Plant Journal*. 2010; 41(5): 744–754.
10. Jin J, Tian F, Yang DC, Meng YQ, Kong L, Luo J, et al. PlantTFDB 4.0: toward a central hub for transcription factors and regulatory interactions in plants. *Nucleic Acids Research*. 2017; 45(D1): D1040–D1045. <https://doi.org/10.1093/nar/gkw982> PMID: 27924042
11. Finn RD, Alex B, Jody C, Penelope C, Eberhardt RY, Eddy SR, et al. Pfam: the protein families database. *Nucleic Acids Research*. 2014; 42(D1): 222–230.
12. Finn RD, Jody C, Eddy SR. HMMER web server: interactive sequence similarity searching. *Nucleic Acids Research*. 2011; 39(Web Server): 29–37.
13. Ivica L, Tobias D, Peer B. SMART: recent updates, new developments and status in 2015. *Nucleic Acids Research*. 2015; 43(D1): 257–260.
14. Aron MB, Derbyshire MK, Gonzales NR, Shennan L, Farideh C, Geer LY, et al. CDD: NCBI's conserved domain database. *Nucleic Acids Research*. 2015; 43(D1): D222.
15. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Altschul S F. Basic local alignment search tool (BLAST)[J]. *Journal of Molecular Biology*, 1990, 215(3):403–410. [https://doi.org/10.1016/S0022-2836\(05\)80360-2](https://doi.org/10.1016/S0022-2836(05)80360-2) PMID: 2231712
16. Kong X, Lv W, Jiang S, Zhang D, Cai G, Pan J, et al. Genome-wide identification and expression analysis of calcium-dependent protein kinase in maize. *Bmc Genomics*. 2013; 14(1): 433–433.
17. Holub EB. The arms race is ancient history in Arabidopsis, the wildflower. *Nature Reviews Genetics*. 2001; 2(7): 516–527. <https://doi.org/10.1038/35080508> PMID: 11433358
18. Yi W, Colemanderr D, Chen G, Gu YQ. OrthoVenn: a web server for genome wide comparison and annotation of orthologous clusters across multiple species. *Nucleic Acids Research*. 2015; 43(1): 78–84.
19. Hu B, Jin J, Guo AY, Zhang H, Luo J, Gao G. GSDS 2.0: an upgraded gene feature visualization server. *Bioinformatics*. 2014; 31(8): 1296. <https://doi.org/10.1093/bioinformatics/btu817> PMID: 25504850
20. Bailey TL, Nadya W, Chris M, Li WW. MEME: discovering and analyzing DNA and protein sequence motifs. *Nucleic Acids Research*. 2006; 34(Web Server): 369–373.
21. Magali L, Patrice D, Gert T, Kathleen M, Yves M, Yves VDP, et al. PlantCARE, a database of plant cis-acting regulatory elements and a portal to tools for in silico analysis of promoter sequences. *Nucleic Acids Research*. 2002; 30(1): 325–327. <https://doi.org/10.1093/nar/30.1.325> PMID: 11752327
22. Dai X, Zhao PX. psRNATarget: a plant small RNA target analysis server. *Nucleic Acids Research*. 2011; 39(suppl): W155–W159.

23. Jian G, Mao L, Zhang C, Hua P, Lin H, Shen Y, et al. A putative pathogen-resistant regulatory pathway between MicroRNAs and candidate target genes in maize. *Journal of Plant Biology*. 2015; 58(4): 211–219.
24. Scheffe JH, Lehmann KE, Buschmann IR, Unger T, Funke-Kaiser H. Quantitative real-time RT-PCR data analysis: current concepts and the novel “gene expression’sCTdifference” formula. *Journal of Molecular Medicine*. 2006; 84(11): 901–910. <https://doi.org/10.1007/s00109-006-0097-6> PMID: 16972087
25. Kavas M, Kızıldoğan AK, Abanoz B. Comparative genome-wide phylogenetic and expression analysis of SBP genes from potato (*Solanum tuberosum*). *Computational Biology & Chemistry*. 2017; 67: 131–140.
26. Geraldo Felipe FES, Eder Marques S, Azevedo MDS, Guivin MAC, Daniel Alves R, Cassia Regina F, et al. microRNA156-targeted SPL/SBP box transcription factors regulate tomato ovary and fruit development. *Plant Journal*. 2014; 78(4): 604–618. <https://doi.org/10.1111/tj.12493> PMID: 24580734
27. Xu M, Hu T, Zhao J, Park MY, Earley KW, Wu G, et al. Developmental Functions of miR156-Regulated SQUAMOSA PROMOTER BINDING PROTEIN-LIKE (SPL) Genes in *Arabidopsis thaliana*. *Plos Genetics*. 2016; 12(8): e1006263. <https://doi.org/10.1371/journal.pgen.1006263> PMID: 27541584
28. Hua-Wei T, Xiao-Ming S, Wei-Ke D, Yan W, Xi-Lin H. Genome-wide analysis of the SBP-box gene family in Chinese cabbage (*Brassica rapa* subsp. *pekinensis*). *Genome / National Research Council Canada = Genome / Conseil national de recherches Canada*. 2015; 58(11): 463–477.
29. Hongmin H, Jun L, Min G, Singer SD, Hao W, Linyong M, et al. Genomic organization, phylogenetic comparison and differential expression of the SBP-box family genes in grape. *Plos One*. 2013; 8(3): e59358. <https://doi.org/10.1371/journal.pone.0059358> PMID: 23527172
30. Li C, Lu S. Molecular characterization of the SPL gene family in *Populus trichocarpa*. *Bmc Plant Biology*. 2014; 14(1): 131.
31. Yang J, Liu X, Yang X, Zhang M. Mitochondrially-targeted expression of a cytoplasmic male sterility-associated orf220 gene causes male sterility in *Brassica juncea*. *Bmc Plant Biology*. 2010; 10(1): 231.
32. Jia-Wei W, Benjamin C, Detlef W. miR156-regulated SPL transcription factors define an endogenous flowering pathway in *Arabidopsis thaliana*. *Cell*. 2009; 138(4): 738–749. <https://doi.org/10.1016/j.cell.2009.06.014> PMID: 19703399
33. Kabin X, Jianqiang S, Xin H, Jialing Y, Xianghua L, Jinghua X, et al. Gradual increase of miR156 regulates temporal expression changes of numerous genes during leaf development in rice. *Plant Physiology*. 2012; 158(3): 1382–1394. <https://doi.org/10.1104/pp.111.190488> PMID: 22271747
34. Stief A, Altmann S, Hoffmann K, Pant BD, Scheible WR, Bäurle I. *Arabidopsis* miR156 Regulates Tolerance to Recurring Environmental Stress through SPL Transcription Factors. *Plant Cell*. 2014; 26(4): 1792. <https://doi.org/10.1105/tpc.114.123851> PMID: 24769482
35. Tripathi RK, Goel R, Kumari S, Dahuja A. Genomic organization, phylogenetic comparison, and expression profiles of the SPL family genes and their regulation in soybean. *Development Genes & Evolution*. 2017; 227(2): 1–19.
36. Kabin X, Congqing W, Lizhong X. Genomic organization, differential expression, and interaction of SQUAMOSA promoter-binding-like transcription factors and microRNA156 in rice. *Plant Physiology*. 2006; 142(1): 280–293. <https://doi.org/10.1104/pp.106.084475> PMID: 16861571
37. Xu M, Hu T, Zhao J, Park MY, Earley KW, Wu G, et al. Developmental Functions of miR156-Regulated SQUAMOSA PROMOTER BINDING PROTEIN-LIKE (SPL) Genes in *Arabidopsis thaliana*. *Plos Genetics*. 2016; 12(8): e1006263–. <https://doi.org/10.1371/journal.pgen.1006263> PMID: 27541584
38. Jae-Hoon J, Yun J, Pil Joon S, Jae-Hyung L, Chung-Mo P. The SOC1-SPL module integrates photoperiod and gibberellic acid signals to control flowering time in *Arabidopsis*. *Plant Journal*. 2012; 69(4): 577–588. <https://doi.org/10.1111/j.1365-313X.2011.04813.x> PMID: 21988498
39. Yu N, Niu QW, Ng KH, Chua NH. The role of miR156/SPLs modules in *Arabidopsis* lateral root development. *Plant Journal*. 2015; 83(4): 673–685. <https://doi.org/10.1111/tj.12919> PMID: 26096676
40. Chao LM, Liu YQ, Chen DY, Xue XY, Mao YB. *Arabidopsis* Transcription Factors SPL1 and SPL12 Confer Plant Thermotolerance at Reproductive Stage. *Molecular Plant*. 2017; 10(5): 735–748. <https://doi.org/10.1016/j.molp.2017.03.010> PMID: 28400323
41. Garcia-Molina A, Xing S, Huijser P. Functional characterisation of *Arabidopsis* SPL7 conserved protein domains suggests novel regulatory mechanisms in the Cu deficiency response. *Bmc Plant Biology*. 2014; 14(1): 231.
42. Shuping X, María S, Antoni GM, Susanne HH, Rita B, Peter H. SPL8 and miR156-targeted SPL genes redundantly regulate *Arabidopsis* gynoecium differential patterning. *Plant Journal*. 2013; 75(4): 566–577. <https://doi.org/10.1111/tj.12221> PMID: 23621152