BMC Plant Biology

RESEARCH Open Access

Comprehensive analysis of *SPL* gene family and *miR156a/SPLs* in the regulation of terpenoid indole alkaloid biosynthesis in *Catharanthus roseus* L

Yaojie Zhang[1], Xiaofan Ni[1], Xueqing Fu[2], Ayat Taheri[1], Wei Zhang[1], Pin Liu[1], Hang Liu[1], Ling Li[1], Yuliang Wang[1*] and Kexuan Tang[1*]

Abstract

Background *Catharanthus roseus* (*C. roseus*), belongs to the dogbane family *Apocynaceae*, is the source of various terpenoid indole alkaloids (TIAs) with excellent antitumor properties, including vinblastine, vincristine, ajmalicine, serpentine, vindoline, and catharanthine. However, TIAs are widely used in clinical anticancer research, their content is extremely low in plants, and currently insufficient to meet market demand. Therefore, exploring TIAs regulation is essential to improve TIAs yield. Numerous transcription factors have been shown to be involved in vinblastine biosynthesis, while the function of *SPL* members (*CrSPL*) has not yet been investigated in *C. roseus*. *SPL* genes regulate both plant growth and development as well as metabolite accumulation.

Results In this study, genome-wide analysis identified fourteen *CrSPL* genes. The analysis included their gene structure, phylogenetic relationships, spatio-temporal expression profiles, response to methyl jasmonate and microRNA interactions. The yeast-one-hybrid (Y1H) and Dual-Luciferase (Dual-Luc) assays and transient overexpression show that *CrSPLs* can directly regulate the vinblastine biosynthesis pathway. The Cro-microRNA member, *Cro-miR156a*, can efficiently inhibit the activity of *CrSPLs* members and TIAs pathway genes.

Conclusions *Cro-miR156a* and *CrSPLs* regulate TIAs biosynthesis, and this study provides a promising reference for investigating the roles of *CrSPL* TFs in regulating TIAs biosynthesis in *C. roseus*.

Keywords *Catharanthus roseus*, microRNA/*SPL*, Terpenoid indole alkaloids, Regulation

*Correspondence:
Yuliang Wang
wangyuliang@sjtu.edu.cn
Kexuan Tang
kxtang@sjtu.edu.cn
[1]School of Agriculture and Biology, Shanghai Jiao Tong University, Shanghai 200240, China
[2]School of Design, Shanghai Jiao Tong University, Shanghai 200240, China

Introduction

In plants, the transcription factors (TFs) can specifically bind to the cis-elements of gene promoter regions and activate or repress gene transcription to regulate the expression of downstream genes. Gene expression is essential for the growth and development of plants. Based on their sequences, structure, motif and biological functions, TFs can be classified into several families such as WRKY, AP2/ERF, MYB, MYC, bHLH and so on [1]. The SQUAMOSA Promoter Binding Protein-Like (SPL) family is one of the plant-specific TFs family and their members are widely distributed in plants. The two SPL family members were firstly identified based on their binding activity to the SQUAMOSA gene promoter region of the snapdragon (*Antirrhinum majus*) flower meristem gene [2]which contain the conserved MDAS-BOX motif and can regulate the snapdragon flower development. Members of *SPL* family share a highly conserved SBP domain, generally about 76 amino acid residues and identified by two zinc-figure binding sites known as Cys3His/Cys4 (Zn1) and Cys2HisCys (Zn2) [3] the later overlaps with a conserved nuclear localization signal sequence (NLS) at the C-terminal of SBP domain [4]. The SBP domain has been showed to be necessary for *SPL* genes to bind to the "GTAC" core motif and gene-specific flanking regions [3].

The orthologous *SPL* genes have been identified and analyzed at the genome-wide level in more than 60 species, including 16 *SPL* members in *Arabidopsis thaliana* [5] 28 in *Populus* [6] 31 in maize [7] 19 in rice [8] 56 in wheat [9] 15 in tomato [10] and so on. The specific roles of *SPL* family members have been extensively studied in several model plants, such as *Arabidopsis thaliana*, rice, maize and tomato. The members of the *SPL* family have been found to be responsible for various plant growth and developmental processes [11–14] including vegetative and reproductive phase transitions [13] pollen sac, leaf development, flowering and root development, gibberellic acid (GA) biosynthesis [14] copper homeostasis, trichome formation [12] and stress response [11]. In *A. thaliana*, the *SPL* genes are divided into eight clades *AtSPL1/12/14/16*, *AtSPL2/10/11*, *AtSPL3/4/5*, *AtSPL6*, *AtSPL7*, *AtSPL8*, *AtSPL9/15* and *AtSPL13* [15]. *AtSPL2/10/11* regulate the lateral organ development and the morphogenesis of cauline leaves and flowers [16]. *AtSPL 3/4/5* controled the flowering time via integrating GA signal and photoperiod with *SOC1-SPL* module [17]. *AtSPL8* plays the central role in the pollen sac development, male fertility and GA biosynthesis and signaling [18] and *AtSPL14* participates in the plant architecture development and sensitivity to fumonisin B1 [19]. In *Oryza sativa*, the fine-tuning expression of the *SPL* genes can enhance the rice yield. *OsSPL14* promotes panicle branching and enhances grain productivity [20]; *OsSPL16* controls the grain size, shape and quality [21]. In maize, the SBP-box genes regulate the plant development such as regulating the rate of lateral primordia initiation [22] the bract development, the establishment of meristem boundaries [23] and the naked grains of maize [24]. As more plant genome resources become available, more *SPL* genes and their function will be uncovered.

The microRNAs (miRNAs) have been showed to regulate the gene expression in plants which are 20–24 nucleotides small RNA molecules, can cleavage the target mRNA or bind to their target-genes transcripts and form an RNA-induced silencing complex to repress the translation [25]. Many *SPL* genes identified to date are targeted by the members of miR156/157 family, and the target sites are in the coding-sequence (CDS) and 3'-untranslated regions (3'UTR) [26]. In general, the *SPL* genes can cooperate with the putative miRNAs to form the miRNA-*SPL* module to play a critical role in various plant growth and development processes.

Catharanthus roseus L (Madagascar's periwinkle), from the dogbane family *Apocynaceae*, is the source of a great variety of terpenoid indole alkaloids (TIAs) and regarded as the model plant for TIAs research. The important TIAs vinblastine, vincristine, ajmalicine, serpentine, vindoline, and catharanthine, accumulate in extremely low quantities in *C. roseus* [27] which have pharmacological activities including anti-cancer, anti-diabetic, anti-microbial and so on. Among them, vinblastine and vincristine are widely used in anticancer research and clinical applications. It is usually assumed that the conversion of vinblastine to vincristine occurs in the absence of enzyme involvement. This leads to efforts to enhance production and develop novel sources of these compounds by implementing multiple strategies. As more detailed genetic background and advanced omics tools of *C. roseus*, such as genomic sequencing, single-cell RNA sequencing and multi-omics provide us with more powerful information about the key enzymes, transcription factors (TFs), and regulatory network of vinblastine. In general, key enzymes and transcriptional regulators, expressed individually or in tandem, have been used to engineer the TIAs pathway in *C. roesus* [28].

Many transcription factors (TFs), involved in the regulation of biosynthesis and metabolism of TIAs, have been identified and utilized in different culture methods of *C. roseus*, including hairy root and suspension cell cultures. The TFs such as *CrORCA1/2/3/4/5/6* are from AP2/ERF family [29]; *CrMYC2a/2b/2c* [28] and *CrBIS1/2/3* are from bHLH family [30]; *CrWRKY1/2* from WRKY family [31] *CrZCT1/2/3* are from C2H2 family [32] and *CrBPF1* is from MYB family [33]. The other TF families have still not been systematically described yet. There is some research focused on the identification of the *C. roseus* miRNA and their potential functions in *C. roseus* [34].

However, the research about the *SPL* family members, miRNA-*SPL* modules and their roles in *C. roseus* growth, development, and especially the function in TIAs regulatory are limited and unclear.

Herein, in order to describe and uncover the *SPL* family members and the miRNAs*/SPL* modules function in *C. roseus*. We characterized 14 members of the *C. roseus SPL* gene (*CrSPL*) family, and analyzed the gene structure, conserved domains, phylogenetic relationship, collinear relationship, Cro-miRNAs target sites and expression patterns. Further, we verified the putative Cro-miRNAs activity and investigated the *Cro-miR156a-CrSPL* potential roles in the regulatory of TIAs in *C. roseus* via Yeast-one-hybrid (Y1H) and Dual-Luc assays. Our study aims to explore and elucidate the potential roles of the *SPL* genes in the *C. roseus* TIAs regulatory network, which will expand the TIAs regulatory network.

Method and materials

Plant materials and treatment

Catharanthus roseus and *Nicotiana benthamiana* were used in this study. The *Catharathus roseus* "*SunStorm Red*" seeds ware bought from Syngenta Flowers company (Grilory, USA), and the *Nicotiana benthamiana* seeds were bought from PU JIE BIOLOGY company (Shanghai, China). The *Catharanthus roseus* and *Nicotiana benthamiana* seeds were planted in vermiculite and peaty soil(1: 2)in pots in a growth chamber under 16 light/8 h dark and 65% relative humidity at 25 ± 2 ℃. Three-week-old *Nicotiana benthamiana* plants were prepared for *Agrobacterium tumefaciens* infection. The whole two-month-old *Catharanthus* plants were sprayed by the hormone of 100 umol/L methyl jasmonates solution, the first pair leaves from plants crown were collected after 0, 0.5, 1, 1.5, 3, 6, 9, and 12 h treatments. And then immediately frozen with liquid nitrogen and stored at -80 ℃ for future experiments. Three biological replicates were harvested at each time point.

Genome-wide identification and distribution of *CrSPLs*

To comprehensively identify the *C. roseus SPL* family genes (*CrSPL*), the chromosome level assembly genome information was obtained from the National Center for Biotechnology Information (NCBI) genome database (GenBank: GCA_024505715.1). The SBP family Pfam seed file (PF03110) was downloaded from Pfam database (https://pfam.xfam.org/), then searched and aligned with the *C. roseus* protein database via the hmmsearch function in HMMER.3.2 based on Ubuntu 22.04.3. The NCBI Conserved Domains Database (CDD) was used to predict the SBP domain and exclude the no-typical, structural integrity and redundant sequence. After that, the molecular weight (MW), isoelectric point (pI) and protein subcellular location of the CrSPL proteins were predicted by ExPASy (Expasy - ProtParam tool) and WoLF PSORT (https://wolfpsort.hgc.jp/). The MG2C website (http://mg2c.iask.in/mg2c_v2.1/) was used to locate and align the positions of *CrSPL* genes based on the genome annotation. The proteins sequences were beautified via ESPript 3.0 (https://espript.ibcp.fr/ESPript/cgi-bin/ESPript.cgi).

The phylogenetic, conserved domains, motifs and gene structure analysis

The phylogenetic tree of *Catharanthus roseus*, *Arabidopsis thaliana*, and *Oryza sativa*(japonica)was constructed using the MEGA11.8.0 with the Neighbor Joining (NJ) method and 1000 replications were set for bootstrap test, in which the protein sequences were truncated by TBtools software File trim function. The phylogenetic tree was visualized and beautified using the online software EvolView (https://www.evolgenius.info/). The pro teins sequences of CrSPLs were submitted to the CDD for conserved domains predication, the CrSPLs gene structure analysis was performed via "Visualize Gene Structure" function in TBtools (Supplementary Table 1, Table S1). The conserved motifs of the full-length proteins were performed via MEME Suite (https://meme-suite.org/), with default setting except the number of displayed motifs were set to 10. The other polygenetic tree including CrSPL proteins, motifs distribution, conserved domains, and gene structure were created by TBTools software according to the annotation information of the *Catharanthus* genomic GFF file.

Gene duplication and collinearity analysis

All amino acid sequences and GFF files of *Catharanthus roseus*, rice, and *Arabidopsis* were conducted a two-way comparison using the One step MCScanX of TBtools software, and then the synteny relationship of SPL genes were performed and visualized using the Dual systeny Plot toolkit package of TBtools software.

Analysis of the cis elements of *CrSPL* genes promoters

The upstream 2000 bp promoter regions of the *CrSPL* genes CDS region were extracted and predicted by TBtools software and Plant CARE (https://bioinformatics.psb.ugent.be/), respectively. The cis-elements distribution figure was constructed by TBtools software. Furthermore, the cis elements were emphasized and divided into four categories: stress responsiveness, light signal, cellular development and phytohormones, and the cis-elements' pie figure was constructed by the Graphpad Prism 9.3.0 version.

Expression analysis of *CrSPL* genes in different tissue and response to phytohormone Methyl jasmonate

The mature leaf (MF), immature leaf (IF), stem (S), root (R) and flower (F) in one-month-old *C. roseus* plants

were collected to perform the *CrSPL* genes expression analysis by RT-qPCR. The phytohormone treatment was performed as described in **2.1**. Total RNA was isolated from the samples with the RNApure Plant Kit (DNase I) (CWBIO, China), the quality and concentration of the RNA samples were measured by Thermo Scientific NanoDrop 2000 UV-Vis Spectrophotometer (Thermo Scientific, United States). Takara PrimerScript RT Kit was used to synthesis cDNA with one ug RNA, and the cDNA was diluted to 500 ng/uL. The RT-qPCR (quantitative real-time polymerase chain reaction) was performed on the LightCycle@96, which contained 10 uL Magic SYB mix, 0.4 uL each of forward and reverse primers, 4 uL cDNA and 5.2 uL ddH_2O in a 20 uL reaction system. The reaction conditions include preincubation at 95 ℃ for 30 s, 40 cycles at 95 ℃ for 10 s, and 60 ℃ for 30 s. Three biological replicates were used in each reaction and the relative gene expression levels were normalized with *C. roseus 18 S rRNA* gene and calculated using the $2^{-\Delta\Delta CT}$ method. A detailed list of the primers used in this study can be found in Table S3.

Prediction of putative micro-RNAs and target genes association and activity test

The published *C. rosues* mature miRNAs were downloaded from PmiREN website (https://pmiren.com/) to predict the putative miRNA sites in the *CrSPL* genes, and then the *CrSPL*s CDS sequences and mature miRNAs sequences were submitted to the online psRNA-Target server (https://www.zhaolab.org/) with default parameters. The mature Cro-miRNA sequence, primary Cro-miRNA sequence and the interaction between the *CrSPL*s and putative miRNA can be found in the supplemental documentation (Table S4-S6). To test the putative Cro-miRNAs activity, the CaMV35s promoter was used to overexpress the primary miRNA sequences. The 500 bp were extracted from the upstream and downstream of mature miRNA sequences on *C. roseus* genome were set as the effector. Meanwhile, the modified pGreenII-0800 vector was set as the reporter, in which CaMV35s promoter was used to promote the Firefly-Luciferases infused with the *CrSPL* genes CDS region at N-terminal.

The recombinant plasmids were transformed into *A. tumefaciens* strain GV3101 with helper plasmids pSoup and p19. The overnight-cultured Agrobacterium cells were collected at an optical density of 1 ~ 1.5 and resuspended in MMA buffer (20 mM 2-Morpholinoethanesulphonic acid, 40 mM Magnesium chloride hexahydrate and 100 mM Acetosyringone) and incubated at room temperature for three hours. Then the effectors and reporters were combined and mixed in 1:1 volume ration, infiltrated the 21-day-old *Nicotiana benthamiana* by disposal syringe. Following a 24-hour dark period and a subsequent 24-hour light period, *N. benthamian* leaves were harvested and pulverized. The relative Firefly-Luciferase/Ren-Luciferase activity was then measured to describe the association between the putative miRNA and *CrSPLs* via the Dual-Luciferase Reporter Assay System (Promega, Madison, WI, USA). Six biological replicates were conducted for each sample, and statistical analysis was performed using either one-way ANOVA or Student's t-test.

Yeast one hybrid (Y1H) assays

The whole coding sequences of the *CrSPL2/3/4/5/7/8/9/12* gene were cloned into the pB42AD vector. Additionally, the pLacZ-2u vector was used to clone the three consecutive copies of the core "GTAC" sequence with 4 ~ 6 bp flanking sequence in vinblastine pathway gene promoters. The recombined pB42AD and pLacZ-2u vectors were transformed into the yeast strain EYG48A, the yeast strain, transformed with the pB42AD and "GTAC" motifs alone, was regarded as the negative control. The yeast cells were cultured on SD/-Trp/-Ura medium, then the yeast clones were incubated on SD/-Trp/-Ura/X-β-gal for selection at 30 ℃ for three days.

Dual-luciferase assay for *CrSPL* genes transcription activity on TIAs biosynthesis genes

The 2000 bp DNA sequence of TIAs biosynthesis key enzymes genes upstream were extracted and regarded as the promoters, then cloned into the pGreen-II 0800 vectors. *CrSPL* genes CDS were cloned into the pHB-x-YFP vector. The *A. tumefaciens* strain transformation, culture and *N. benthamiana* infiltrated method were described in **2.7**. Six biological repeats were performed for each sample; statistical analysis was performed via one-way ANOVA or Student's t test.

The transient overexpression of *Cro-miR156a* and *CrSPLs* in *C. roseus* seedlings

The *Cro-miR156a* and *CrSPLs* were transiently overexpressed in *C. roseus* seedlings as described in research [35].

Results

Identification and characters of the *SPL* gene family in the *Catharanthus roseus* genome

SQUAMOSA promoter-binding protein-like transcript factors (*SPL*) are plant-specific transcription factors with the conserved domain SBP that can recognize and bind to the promoter region of MADS box genes. To identify SPL genes in the *C. roseus* genome, the *SBP* HMM file (Pfam: PF03110) was used to blast against the *C. roseus* genome based on the Ubuntu system. Fourteen *SPL* genes were identified and then blasted using NCBI Batch CD-Search

to predict the SBP domain and exclude the atypical, structural integrity and redundant sequences. To avoid naming confusion, the *SPL* genes were named *CrSPL1-CrSPL14* according to their chromosomal location and distributed on linkage group (LG): LG02(4), LG03(1), LG05(2), LG06(2), LG07(2) and LG08(3) (Fig. 1a).

The CDS and protein length of the fourteen *CrSPL* genes ranged from 441 to 3150 bp and 146 to 1084 aa, respectively (Table 1 and Table S1). The highly conserved and intact SBP domain appeared at the N-terminal region of all *CrSPL* genes, it contains two zinc finger structures (Zn-1 and Zn-2) and a nuclear localisation sequence (NLS) (Fig. 1b). In addition, the protein characteristics including basic physical and chemical properties were identified. The molecular weights (M.W.) and theoretical isoelectric points (pI) of the proteins ranged from 16.92 ~ 111.766 kDa and 5.71 to 9.73, respectively. Nine proteins of *CrSPL* family members' pI were greater than 7, which means that these proteins belong to the basic proteins. In terms of subcellular localisation, the WolF PSORT (https://wolfpsort.hgc.jp/) prediction results suggested that CrSPL proteins may be targeted to the nucleus (Table S2).

Phylogenetic relationship, conserved domains, motifs and gene structure of *CrSPLs*

In general, the genes that had a co-adaptation association and relationship can perform similar functions. To explore the functions and revolution relationship of CrSPL proteins, an unrooted phylogenetic tree, contains 14 *C. roseus* SPL(CrSPL), 16 *A. thaliana* SPL(AtSPL) and 19 *O. sativa* SPL(OsSPL) proteins, were constructed by the neighbor-joining method (Fig. 2). These *CrSPL* proteins were divided into eight subgroups: according to the phylogenetic relationship of SPL proteins in *A.thaliana*: CrSPL1/6/14 (Group 1), CrSPL11 (Group 2), CrSPL3 (Group 3), CrSPL2/9 (Group 4), CrSPL5/12(Group 5), CrSPL7/8/10 (Group 6), CrSPL4 (Group 7) and CrSPL13 (Group8).

In addition, to understand the evolutionary significance of domain architecture, a phylogenetic tree of the *CrSPL* gene family combined with gene structure and phylogenetic tree was constructed by MEGA-11.0.8. We performed the conserved motif prediction of 14 CrSPL proteins via the online prediction website MEME, which were set as Motif 1-Motif 10 (Fig. 3). All members contain motif 1, motif 2 and motif 3, which are the most conservative regions in SPL proteins. Interestingly, motif

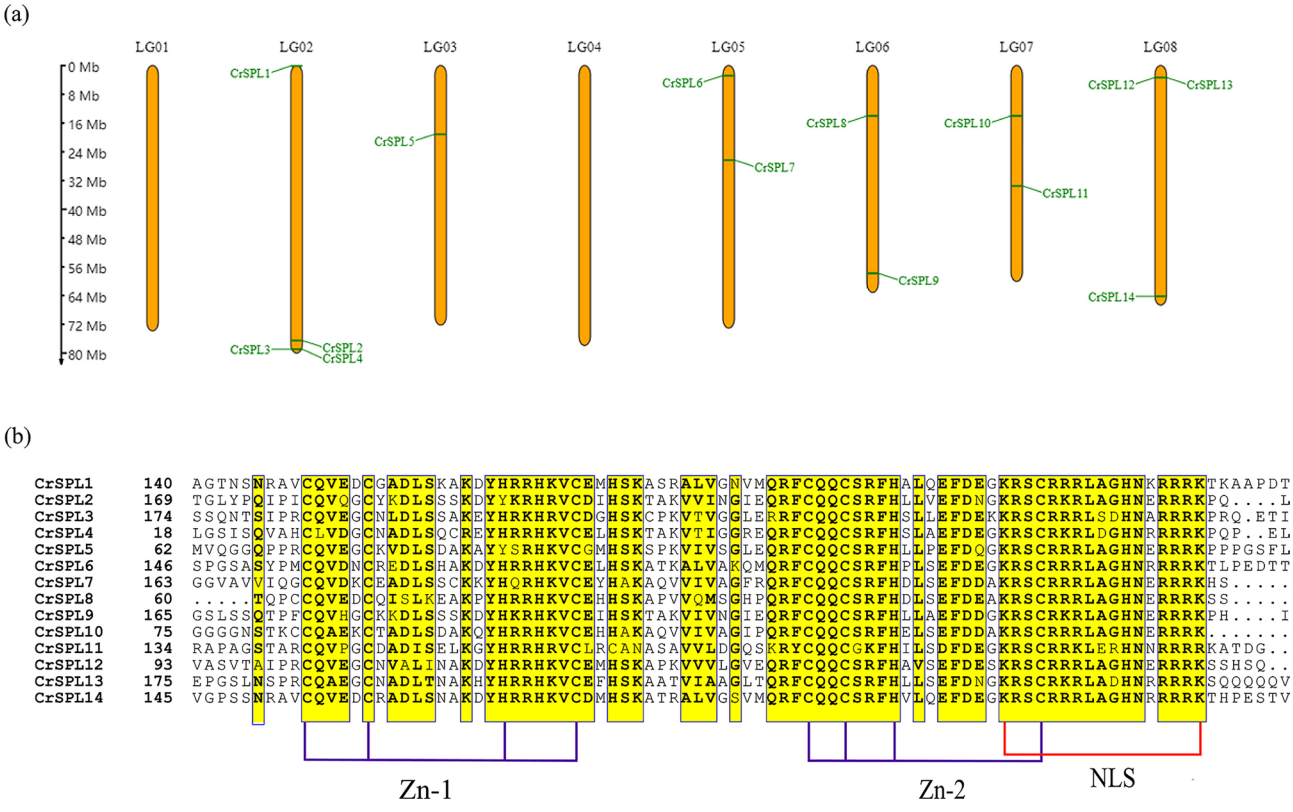

Fig. 1 Identification of *CrSPLs* gene in *Catharanthus roseus*. **(a)** The *CrSPLs* gene location on *Catharanthus roseus* chromosome. The dark green background shape and red text and lines were used to represent the *Catharanthus* chromosome and *CrSPLs* gene location respectively. **(b)** Amino acid sequence alignment of *CrSPL* proteins. The Zn-1 and Zn-2 domains and nuclear localization signal sequence (NLS) are indicated upside, the color background indicated identical amino acids

Table 1 Information about *CrSPL* genes in *Catharanthus roseus*

Gene ID	Locus tag	Protein ID	Gene name [a]	Chromosome position	CDS length(bp)	Protein length(aa)	MW (kDa) [b]	PI	Subcellular localization [c]
EVM0029095.1	M9H77_05808	KAI5674858.1	CrSPL1	LG02:134600–144,161(+)	3025	1008	112.888	6.85	Nucleus
EVM0001126.1	M9H77_09814	KAI5678864.1	CrSPL2	LG02:72983770–77,810,433(+)	1456	485	55.441	7.63	Nucleus
EVM0027799.1	M9H77_10331	KAI5679381.1	CrSPL3	LG02:77676301–77,681,780(-)	1417	472	51.380	8.79	Nucleus
EVM0020734.3	M9H77_10342	KAI5679392.1	CrSPL4	LG02:77804193–77,810,433(-)	1015	338	37.510	8.63	Nucleus
EVM0011066.1	M9H77_11907	KAI5671543.1	CrSPL5	LG03:18912964–18,916,171(-)	1114	371	39.879	8.63	Nucleus
EVM0032430.1	M9H77_20184	KAI5660861.1	CrSPL6	LG05:2865836–2,873,030(-)	3253	1084	119.695	7.15	Nucleus
EVM0027260.1	M9H77_21849	KAI5662526.1	CrSPL7	LG05:26031754–26,033,298(-)	847	282	32.370	7.68	Nucleus
EVM0033139.1	M9H77_25920	KAI5657127.1	CrSPL8	LG06:13720601–13,723,401(+)	441	146	16.923	5.71	Nucleus
EVM0032019.1	M9H77_28837	KAI5660044.1	CrSPL9	LG06:56878716–56,883,734(-)	1669	556	60.870	6.55	Nucleus
EVM0011359.1	M9H77_29721	KAI5652534.1	CrSPL10	LG07:13723417–13,725,672(-)	630	209	22.188	9.73	Nucleus
EVM0023234.1	M9H77_31257	KAI5654070.1	CrSPL11	LG07:32924108–32,936,023(-)	2470	823	91.219	6.59	Nucleus
EVM0035658.1	M9H77_33422	KAI5647417.1	CrSPL12	LG08:3402562–3,406,972(+)	1096	365	40.756	8.39	Nucleus
EVM0021833.1	M9H77_33423	KAI5647418.1	CrSPL13	LG08:3410856–3,414,266(-)	1045	348	38.231	8.78	Nucleus
EVM0005315.1	M9H77_36989	KAI5650984.1	CrSPL14	LG08:63175879–63,186,258(+)	3010	1003	111.766	6.74	Nucleus

[a] The name of SPL genes in this study

[b] M.W., Molecular weight (kDa Kiloadalton)

[c] pI, Isoelectric point. bp, base pair. aa, amino acid

4, motif 5 and motif 6 most only appeared in CrSPL1, CrSPL14, and motif 3 exit in CrSPL11 alone. Meanwhile, motif 3, motif 5 and motif 6 near the C-terminal in the CrSPL1, CrSPL4 and CrSPL11. The motifs' order and range provide the possibility of more functions for the protein.

To study the gene structure of 14 *CrSPL* genes, CDS/introns patterns were further investigated by aligning the corresponding genomic DNA sequence. Among them, the *CrSPL1*, *CrSPL6*, *CrSPL11* and *CrSPL14* have the most exon numbers about 10–11, and *CrSPL1* and *CrSPL7* have the longest and shortest introns, respectively. Furthermore, the UTR has a variable distribution in these genes, no 5'UTR or 3'UTR regionappeared in *CrSPL2*, *CrSPL5*, *CrSPL8* and *CrSPL11*. In other plants' SPL family, the 5'UTR or 3'UTR also were absent in *SPL* gene structure, such as *TaSPL8/13/15* in wheat, *AtSPL2/5/7/12/13A/13B/15* in *Arabidopsis.*

Gene duplication within *C. roseus* and collinearity analysis of SPL genes between *C. roseus*, *A. thaliana* and *O. sativa*

Segmental and tandem duplication events generally promote gene expansion and evolution, resulting in various paralogous gene pairs that aid plants in coping with diverse growth and developmental processes. Members of *SPL* genes that cluster into the same subgroup share similar gene structure and functions, suggesting a common evolutionary origin. Our study identified only two gene duplication events: *CrSPL7-CrSPL8* and *CrSPL2-CrSPL9*. This suggests that most of the *CrSPL* genes are conserved during the evolutionary process. In other words, the members of the SPL genes might have diverse functions in plant development.

We explored the collinear relationship of *CrSPL* after gene duplication within *C. roseus* by comparing the genomes of *Arabidopsis* and rice (Fig. 4). The results show that the genomes of *C. roseus* and *A. thanliana* have 14 (100%) syntenic gene pairs, which corresponds to the distribution of *CrSPL* on the *C. roseus* chromosome. Three gene pairs were found to be syntenic between *C. roseus* and the Oryza genome.Based on the evolutionary association analysis of gene duplication, a strong collinear correlation was observed between *C. roseus* and *A. thanliana*. These results demonstrate that *C. roseus* and *A. thanliana* have conserved and consistent phylogenetic relationships of SPL genes.

Promoter cis-elements analysis of *CrSPLs*

The nucleotide sequences of 2000 bp were extracted from the upstream of each *CrSPL* CDS region and regarded as promoters. The PlantCARE website was used to predict cis-elements for all sequences. The 553 screened promoter cis-elements can be categorized into four groups: 229 stress-responsive cis-elements, 139 light response and regulatory cis-elements, 148 hormone-responsive cis-elements, and 47 cellular development cis-elements (Fig. 5). The most widely distributed cis-element was MYB, accounting for 26.6%. Additionally, STRE was identified as a significant cis-element in stress-responsive cis-elements. Furthermore, some special cis-elements for MYB binding sites were identified in stress response, such as MBSI, MRE, and MBS, which were involved in regulating flavonoid metabolism, light response, and

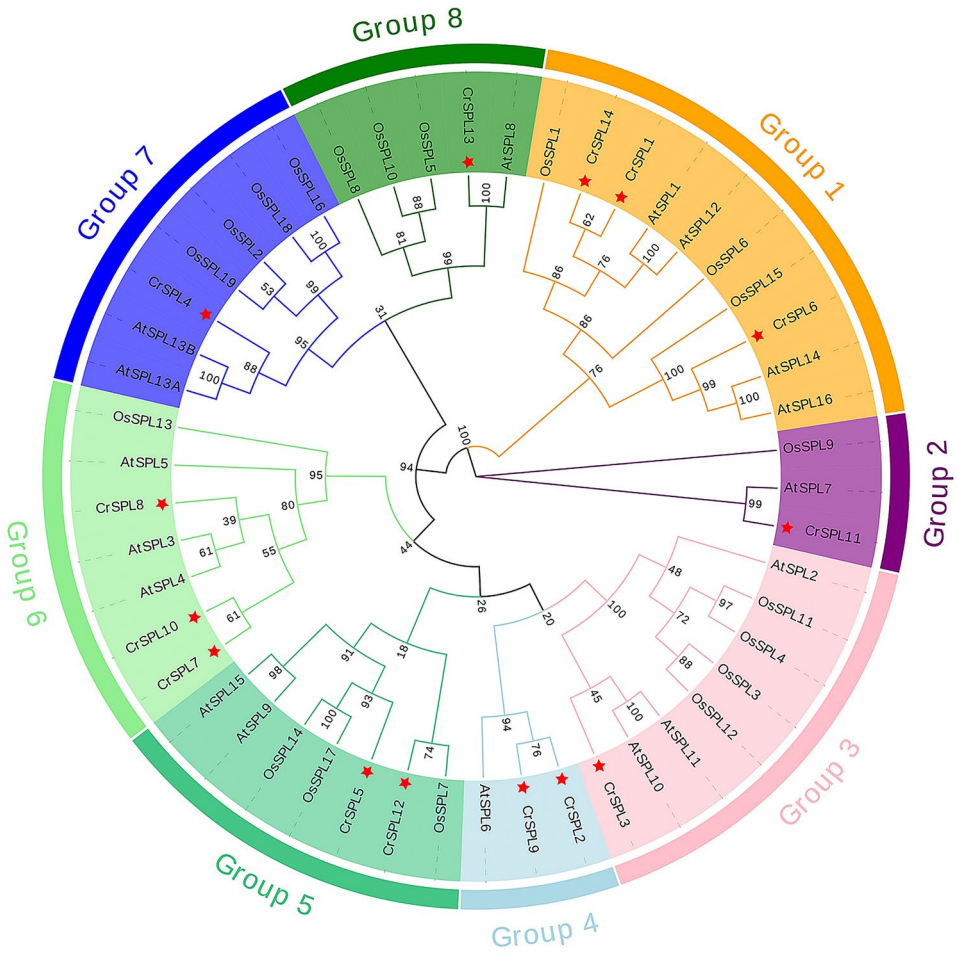

Fig. 2 The neighboring-joining tree between *Catharanthus roseus*, *Arabidopsis thaliana*, and *Oryza sativa*. Different colors were used to distinguish different groups and branches, the *Catharanthus roseus* SPL proteins were decorated by red filled star, and the numbers in the middle of the branches represent the bootstrap scores

diverse stress. The number of stress-responsive cis-elements contributed half of the total cis-elements, suggesting that these genes may be induced in response to biotic and abiotic stress.

Promoters of *CrSPL* were distributed by cis-elements responsible for light signal and hormones such as ABA, JA, Auxin, GA, SA and ETH. The most widely distributed cis-elements were Box 4 and TCT-motif in light signal and MYC and ERE in hormones, accounting for 25.2%, 11.5%, 31.1% and 14.9% respectively. And jasmonic acid elements (MYC, TGACG and CGTCA) were most widely distributed in cis-elements responsible for hormones, accounting for 31.1%, 5.4% and 5.4% respectively, it suggests that *CrSPL* genes might were strongly regulated by JA. The promoter sequence contains several cis-elements related to cellular development, including A-box, CAT-box, CATA-motif, CCGTCC-box/motif, GCN4-motif, and HD-Zip 1. CCGTCC-box/motif and A-box were the most prevalent, accounting for 35.1% and 16.2%, respectively. This suggests that some *CrSPL* genes may play a role in plant growth and development. Overall, we identified several cis-elements in the *CrSPL* promoter sequences. Our findings suggest that *CrSPL* genes are involved in various biological processes in response to stress, light, hormones, and cellular development.

Expression profiles of *CrSPL* genes in different tissues and responses to Methyl jasmonates

To gain an overview of the tissue-specific expression profile of *SPL* genes in *C. roseus*, we analysed the relative expression levels via RT-qPCR in different *C. roseus* tissues, including flowers, stems, immature leaves, and mature leaves (Fig. 6 and Fig. S2a). Most *SPL* genes exhibited low expression levels in the stem. Among them, *CrSPL1*, *CrSPL2*, *CrSPL3*, *CrSPL7*, *CrSPL8*, and *CrSPL10* showed the highest expression levels in mature leaves. *CrSPL5*, *CrSPL6*, *CrSPL11*, *CrSPL12*, *CrSPL13*, and *CrSPL14* show high expression profiles in flowers, while only *CrSPL4* and *CrSPL9* showed high expression levels in immature leaves and roots, respectively. Meanwhile,

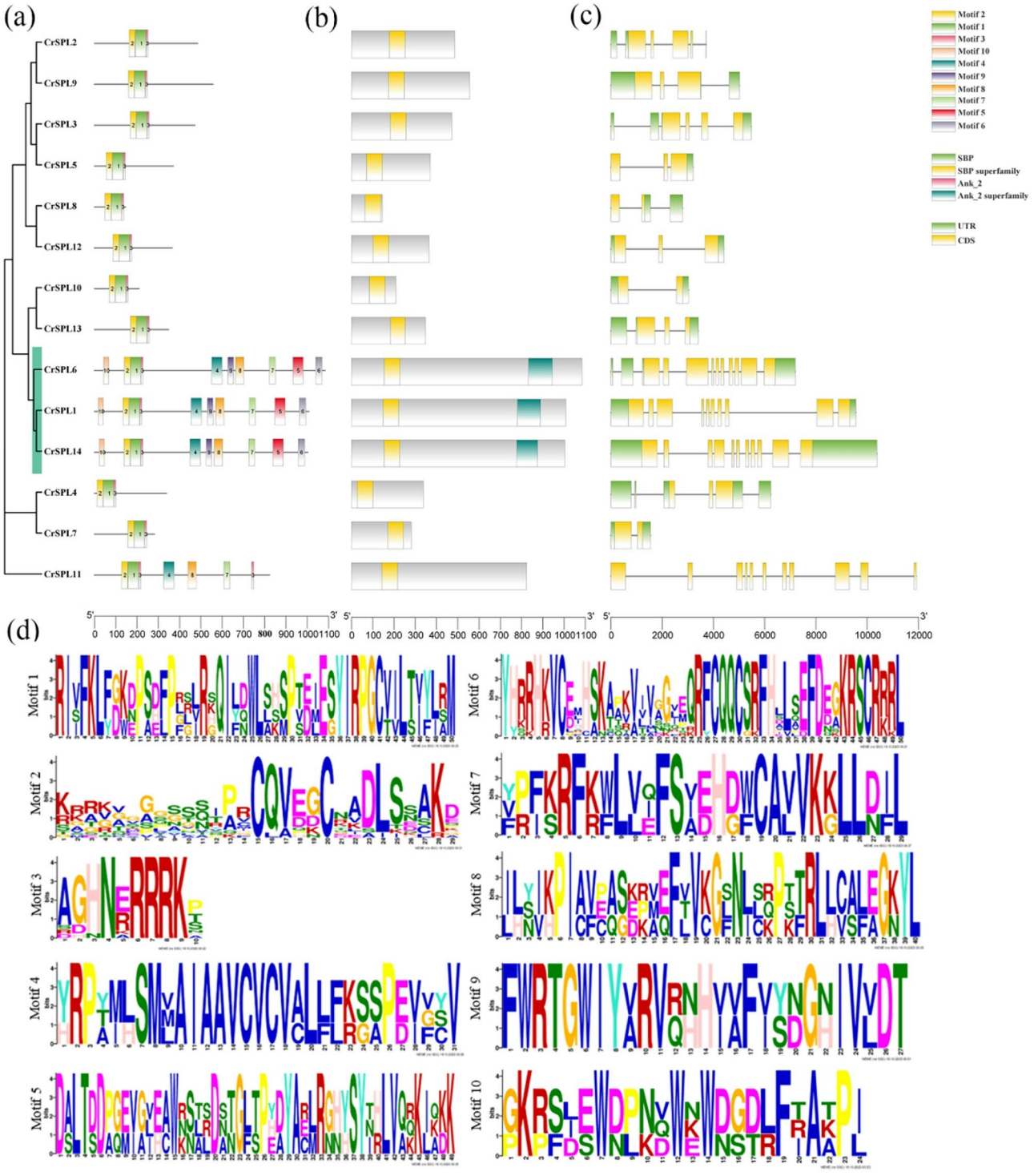

Fig. 3 The unrooted phylogenetic tree, the gene structure and motif analysis of *CrSPL* genes. **(a)** The neighbor-joining tree on the left 14 CrSPL proteins form *C. roseus*. **(b)** Ten conserved motifs were represented by boxes, and different colors were used to represent these motifs. **(c)** The green color boxes indicate the 5′ and 3′ untranslated regions (UTR), and the yellow and gray color represent the exons and introns, respectively. **(d)** Different motifs amino acid sequences of CrSPL proteins

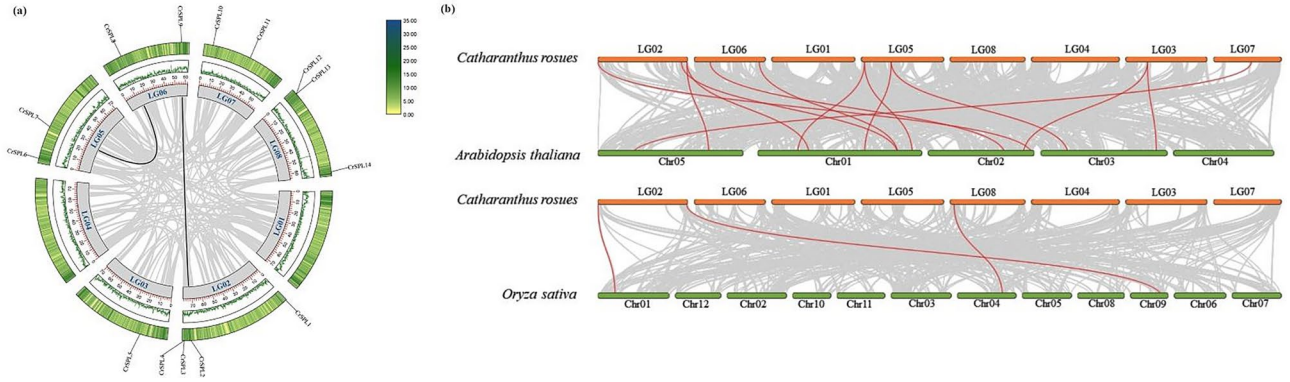

Fig. 4 Orthologous relationships between SPL genes in *Arabidopsis thaliana*, *Oryza japonica*. The orange blocks represent the chromosomes of *C. roseus*, while the green blocks represent chromosomes of other species. Collinear blocks among *SPL* genes were highlighted by red lines, while gray lines represent collinear gene pairs between different species

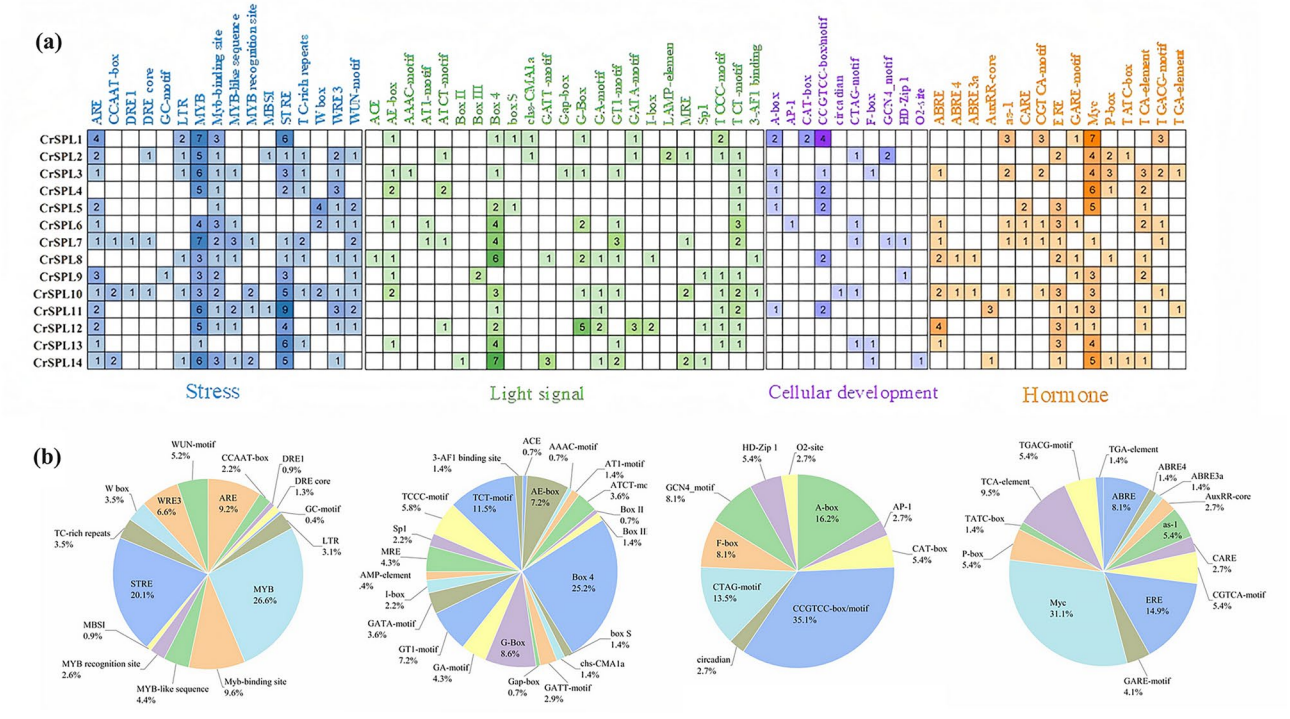

Fig. 5 The analysis of *cis*-elements of 14 *CrSPL* genes promoters. (**a**)The cis-elements classification and numbers of *CrSPL* genes were represented by different colors and grid numbers. (**b**)The four categories of Stress responses, Light signal, Cellular development and Hormone response were shown in different colors. The pie chart size according to the cis-elements ratio (from left to right)

the *CrSPL* gene expression profile from different cell populations is also shown in Fig. S2b. Based on the profile of *CrSPL* genes, it is suggested that *CrSPL* may have various roles in growth processes, and certain members of the *CrSPL* family may play a dominant role at the same time.

TIA biosynthesis can be strongly stimulated by MeJA. Analysis of the promoters of *CrSPL* genes revealed that JA-response elements appear in the promoter region, indicating that *CrSPL* genes could respond to MeJA hormone signals. The RT-qPCR results indicated that the members of the *CrSPL* family members can be induced by MeJA (Fig. 7 and Fig. S3). MeJA strongly inhibited the expression levels of *CrSPL1*, *CrSPL3*, *CrSPL11* and *CrSPL13*, and significantly upregulated the expression level of *CrSPL2*, *CrSPL4*, *CrSPL5*, *CrSPL6*, *CrSPL7*, *CrSPL8*, *CrSPL9*, *CrSPL10* and *CrSPL12*, except for no effect on *CrSPL4* expression level. The *CrSPL1*, *CrSPL3*, *CrSPL4*, *CrSPL5*, *CrSPL6*, *CrSPL7*, *CrSPL9* and *CrSPL10* responded rapidly to MeJA within 1.5 h, while other members responded more slowly.

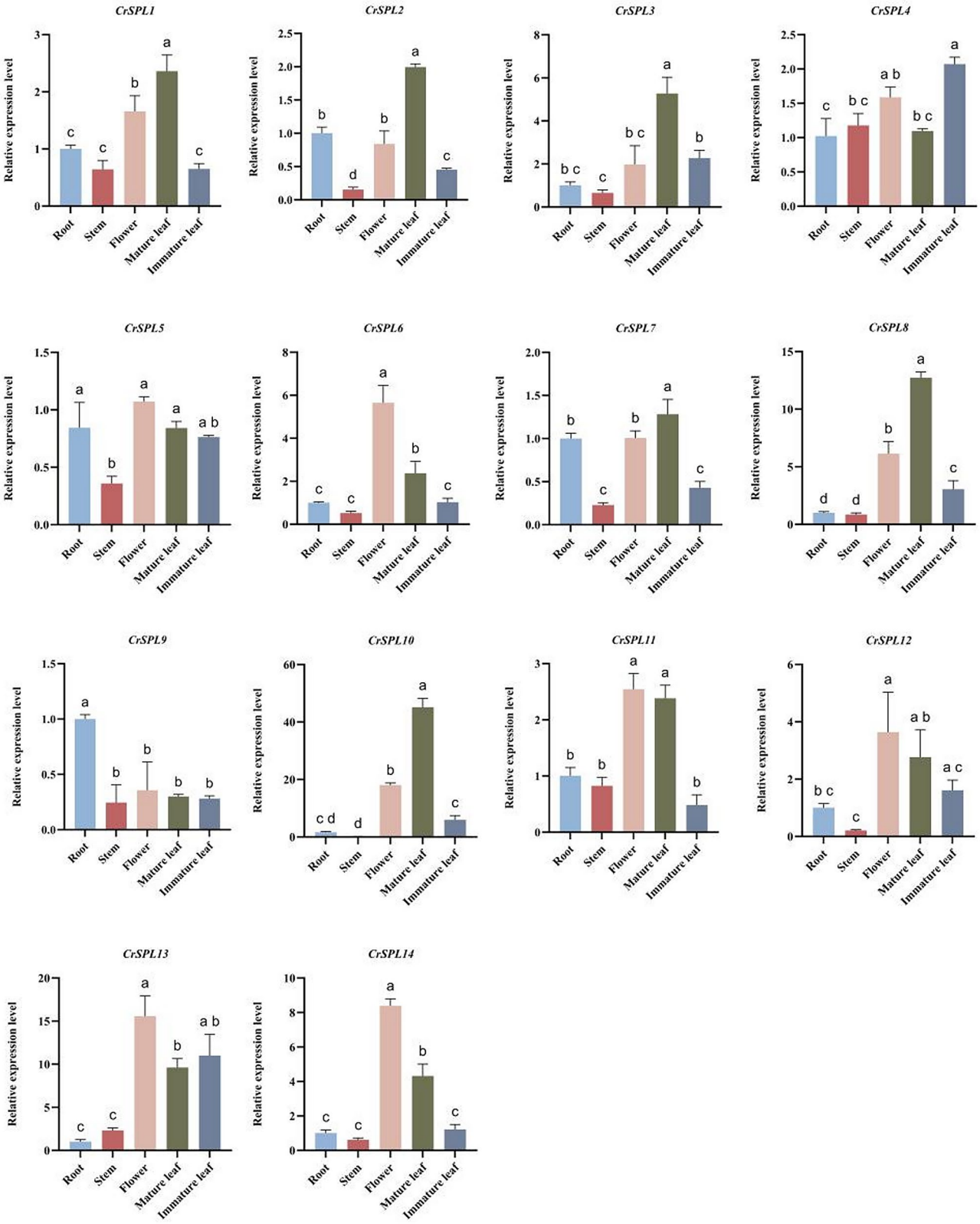

Fig. 6 The expression profiles of *C. roseus SPL* genes in flower (F), stem (S), immature leaf (IF) and mature leaf (MF). Error bars represent mean ± SD ($n = 3$). The results represent the mean ± SD of biological triplicates ($n = 3$), and $2^{-\Delta\Delta ct}$ was used to calculate the relative expression level

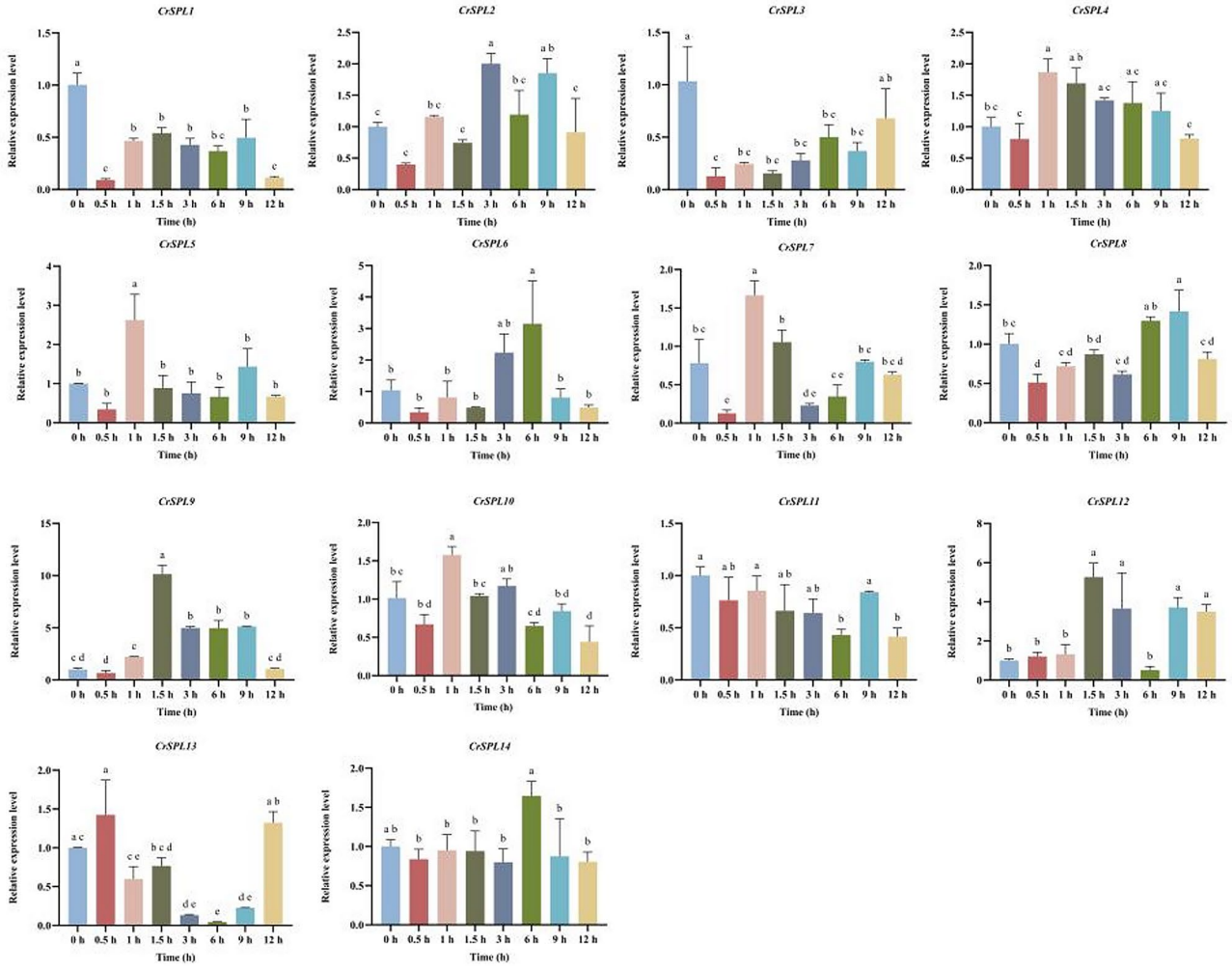

Fig. 7 *CrSPLgenes* response to methyl jasmonates (MeJA). The RT-qPCR was used to detect the relative expression levels of *CrSPL* genes under MeJA treatment. The results represent the mean ± SD of biological triplicates ($n = 3$), and $2^{-\Delta\Delta ct}$ was used to calculate the relative expression level.

Prediction of putative miRNA-target interactions and activity test

MicroRNAs (miRNAs) are involved in the regulation of gene expression in plants, acting as post-transcriptional repressors by targeting downstream genes for cleavage or degradation [36]. The psRNATarget tool (https://www.zhaolab.org/) was employed to ascertain the presence of the target sites within the *CrSPL* family members. The results suggested that ten members of Cro-miRNA family members can be paired with certain *CrSPL* genes in the CDS region (Fig. 8a). These members include *Cro-miR156a-d*, *Cro-miR159c*, *Cro-miR2111a*, *Cro-miR2275a*, *Cro-miR828a*, *Cro-miRN5319* and *Cro-miR391a*. The *Cro-miR156a-d* can target *CrSPL2*, *3*, *4*, *5*, *9* and *12*, while *Cro-miR391a* targets *CrSPL1* and *Cr-miR2111a* targets *CrSPL10*. The association relationship between *CrSPL* genes and Cro-miRNAs is not specific. For instance, *CrSPL2* can be targeted by *Cro-miR156a-d*, *CrSPL3* by *Cro-miR159e* and *Cro-miR156a-d*, and *CrSPL4* by *Cro-miR2775a* and *Cro-miR828a*.

To test the activity of Cro-miRNAs, *Cro-miR156a* was overexpressed using the CaMV35s promoter due to its mature sequence homology as the effector. Additionally, we cloned members of the *CrSPL* family members (*CrSPL2*, *3*, *4*, and *5*) into the modified pGreenII-0800 vector as the reporter, in which the CrSPL CDS were infused at the N-terminator of Firefly-Luciferase protein. (Fig. 8b). The relative Dual-Luciferase level was measured, and the results showed that Cro-miR156a significantly reduced the relative Luciferase activity of the reporter. The results indicated that there was a decrease in activity of *CrSPL2* by approximately 38%, *CrSPL3* by 55%, *CrSPL4* by 44%, and *CrSPL5* by 69% (Fig. 8c). This suggests that the primary *Cro-miR156a* sequence can be transcribed and processed into mature micro small RNAs, which can then affect the activity of the *SPL* genes. The difference in the level of decrease among the target

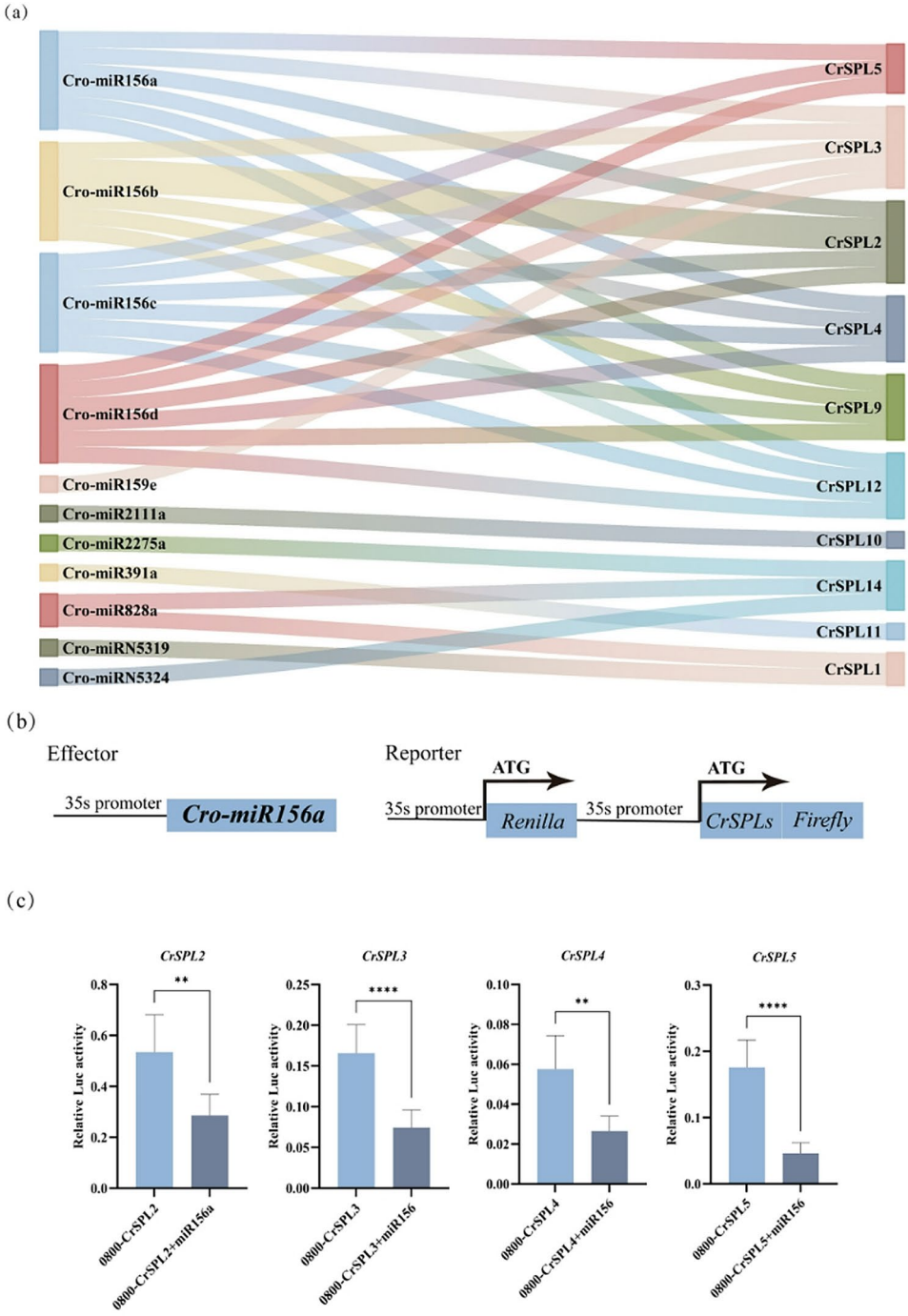

Fig. 8 The interaction and association analysis between putative miRNAs and their target *CrSPL* genes. **(a)**The putative miRNAs and targeted genes are in the left and right, respectively. The different colors were used to connect the miRNAs and targeted genes to describe the interaction and association relationship. **(b)** The vectors constructed for the *Cro-miR156* function test; **(c)** *Cro-miR156a* activity tests in *N. benthamiana*. The data were showed as the mean ± standard error ($n = 6$). Graphmad prism 9.3.0 and Excel were used for data processing and statistical analysis, and *Student's t test* was used for significance test. *$P \le 0.05$, **$P \le 0.01$,***$P \le 0.001$, **** $P \le 0.0001$

SPL genes may be due to the binding and cleavage capabilities of Cro-miR156 or the limited effect of *SPL* mRNA change on *SPL*/FireflyLuc fusion protein activity.

The effects of CrSPL2/3/4/5/7/8/9/12 on the promoter activity of part enzymes in vinblastine biosynthesis

The biosynthesis of vinblastine is a complex process, comprising three distinct phases: the shikimate-chorismate-indole pathway, the MEP-secoiridoid pathway, and the monoindole alkaloid pathway (Fig. S1). However, the involvement of SPL genes in the regulation of vinblastine biosynthesis remains to be demonstrated. A Dual-Luciferase and Yeast One-Hybrid assay were conducted to investigate the potential involvement of *CrSPL* genes in the biosynthesis of TIAs.

Initially, a Dual-LUC assay was conducted to ascertain whether CrSPL proteins exert an influence on the activity of specific vinblastine synthesis enzymes (*Cr7DLGT*, *CrLAMT*, *CrSGD*, *CrGS*, *CrT3R*). The CrSPL2/3/4/5/7/8/9/12 were selected to perform the Dual-LUC, based on their interaction with *Cro-miR156a* and the expression profiles form different tissues and cell populations RNA-seq data (Fig. S2). As illustrated in Fig. 9, the relative Luc activities demonstrate that CrSGD activity can be activated by CrSPL2/3/4/5/9/12. Similarly, CrSPL2/3/4/5/9/12 activated the CrGS LUC activity, with the exception of CrSPL7/8, which had no effect. Additionally, only CrSPL9 was observed to activate Cr7DLGT and CrLAMT activity. A similar trend was observed for CrT3R, whereby CrSPL2/9 activated the Luc activity, while CrSPL3/4/5/7/8/12 had no effect.

The function of *CrSPL* in the biosynthesis of TIAs was further substantiated through the utilization of a Yeast One-Hybrid assay (Fig. 10). The SPL family member has been demonstrated to possess the capacity to specifically recognize and bind the "GTAC" box. The PlantCARE analysis has revealed the presence of 5, 8, 2, 3 and 1 "GTAC" boxes on CrSGD, CrGS, Cr7DLGT, CrLAMT and CrT3R promoter region, respectively. The yeast one-hybrid assay demonstrated that the second "GTAC" box on CrSGD was directly bound by CrSPL8 and CrSPL12, while the third "GTAC" box on CrLAMT was directly bound by CrSPL2/3/4/5/8/9. Additionally, the sole "GTAC" box on CrT3R was shown to be directly bound by CrSPL3 and CrSPL8. In conclusion, the *CrSPL* genes may exert a direct influence on the activity of key enzymes involved in the biosynthesis of TIAs.

Transient ovexpression of *Cro-miRNA* and *CrSPLs* in *C. roseus* cotyledon

To further validate the function of *Cro-miR156a* and the *CrSPLs* in vivo, the transient overexpression was performed in the *C. roseus* cotyledon system. The transient overexpression of *Cro-miR156a* was observed

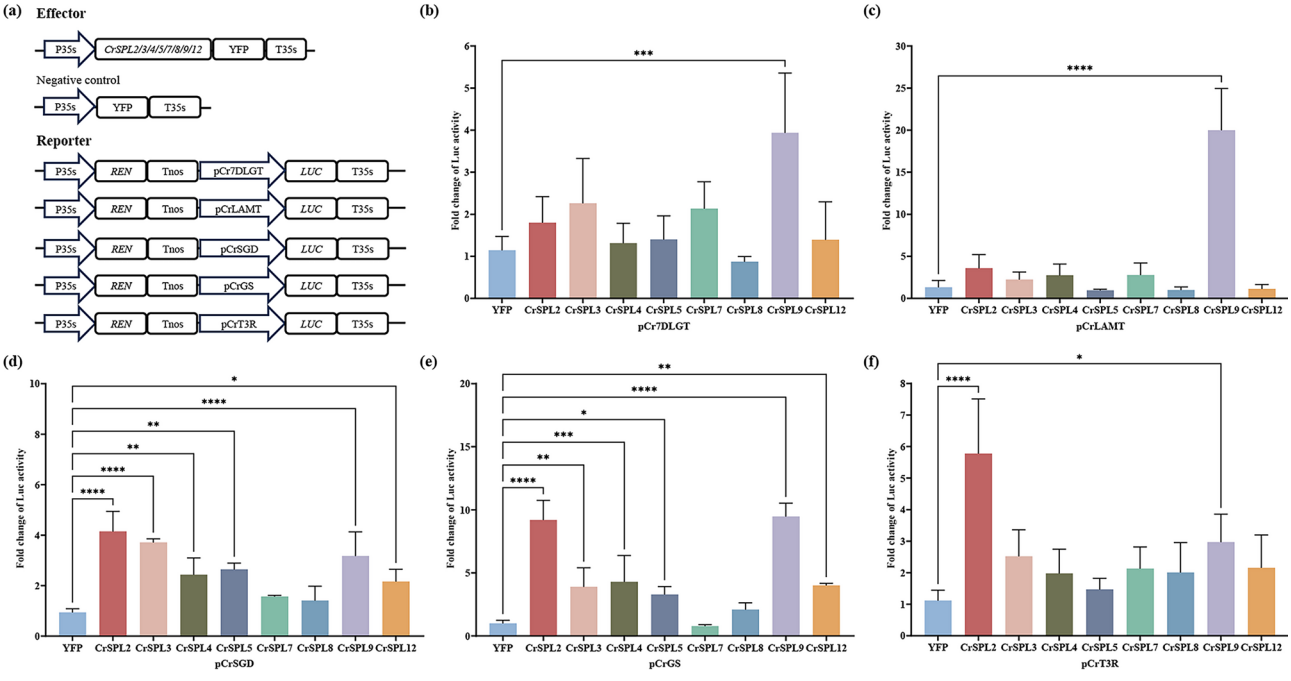

Fig. 9 *CrSPLs* gene can activate the TIAs pathway gene promoter activity. **(a)** The vector construction for Dual-Luciferase assay. **(b)-(f)** The effect of *CrSPLs* gene on TIAs pathway gene promoter. *Cr7DLGT*, 7-deoxyloganetic acid glucosyl transferase; *CrLAMT*, loganic acid O-methyltransferase; *CrSGD*, strictosidine β-D-glucosidase; *CrGS*, geissoschizine synthase; *CrT3R*, tabersonine 3-reductase. The data showed as the mean ± standard error ($n = 6$). Graphmad prism 9.3.0 and Excel were used for data processing and statistical analysis, and *Student's t test* was used for significance test. *$P \le 0.05$, **$P \le 0.01$, ***$P \le 0.001$, **** $P \le 0.0001$

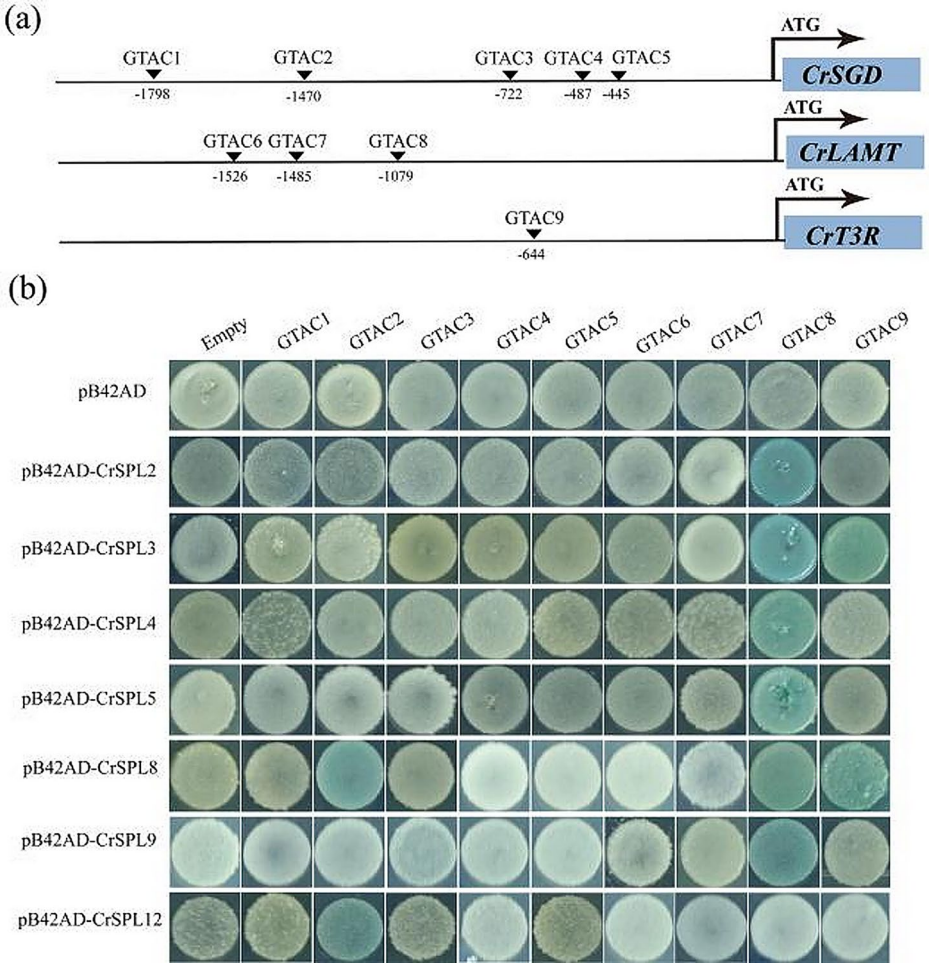

Fig. 10 *CrSPLs* can directly bind TIAs pathway gene promoter region. **(a)** Schematic diagrams of the CrSGD, CrLAMT and CrT3R promoters, and the "GTAC" cis-elements position are marked by rectangle. (**b**) Yeast one-hybrid (Y1H) assay for protein-DNA interaction, the Y1H assays results were one of six times repeated assay

to effectively reduce the relative expression levels of *CrSPL2*, *CrSPL3*, *CrSPL5* and *CrSPL9* by 61.4%, 21%, 59% and 69.7%, respectively (Figrue. 11). These results are consistent with those obtained from the Dual-Luc assay in *N. benthamiana* in above. Nevertheless, no statistically significant difference was identified in the relative expression levels of *CrSPL4* and *CrSPL12*. As anticipated, the transient overexpression of *Cro-miR16a* was observed to alter the expression levels of genes involved in the vinblastine pathway. In comparison to the YFP transient overexpression lines, there was a notable reduction in the expression levels of *CrLAMT*, *CrSLS*, *CrTDC*, *CrSGD*, *CrRedox1*, *CrDAT* and *CrPRX1* excepted *CrSTR*, *CrGO*, *CrSAT* and *CrT3R*. This suggests that *Cro-miR16a* may be capable of reducing vinblastine pathway gene expression by decreasing the activity of *CrSPLs*, as well as by directly binding to the vinblastine pathway gene promoter and regulating its expression level.

In addition, *CrSPL2*, *CrSPL3* and *CrSPL5* were expressed in cotylen, respectivley. Interesting, the overexpression of *CrSPL2*, *CrSPL3* and *CrSPL5* have limited effect on TIAs biosynthesis pathway genes expression level. In *CrSPL2*-overexpressed line, *CrSPL2* can lightly increase *CrSGD*, *CrGO*, *CrT3R* and *CrDAT* expression level. as with the effect of *CrSPL2*, *CrSPL5* only had slight influence on *CrGS* and *CrGO*. However, it seems that *CrSPL3* had no effficent effect on these genes (Fig. 12). These resluts indicated that overexpressing of *CrSPL* member alone might have limited effect on TIAs biosynthesis in *C. roseus* cotyleon.

Discussion

Catharanthus roseus is a widely cultivated plant known for its cancer-fighting properties and attractive appearance. It is an important source of terpenoid indole alkaloids (TIA) and is often used as a research model plant. Vinblastine and vincristine are two of the most widely used natural plant antitumor drugs, widely used in the treatment of Hodgkin's disease, acute lymphoblastic leukemia, malignant lymphoma and other disease s,

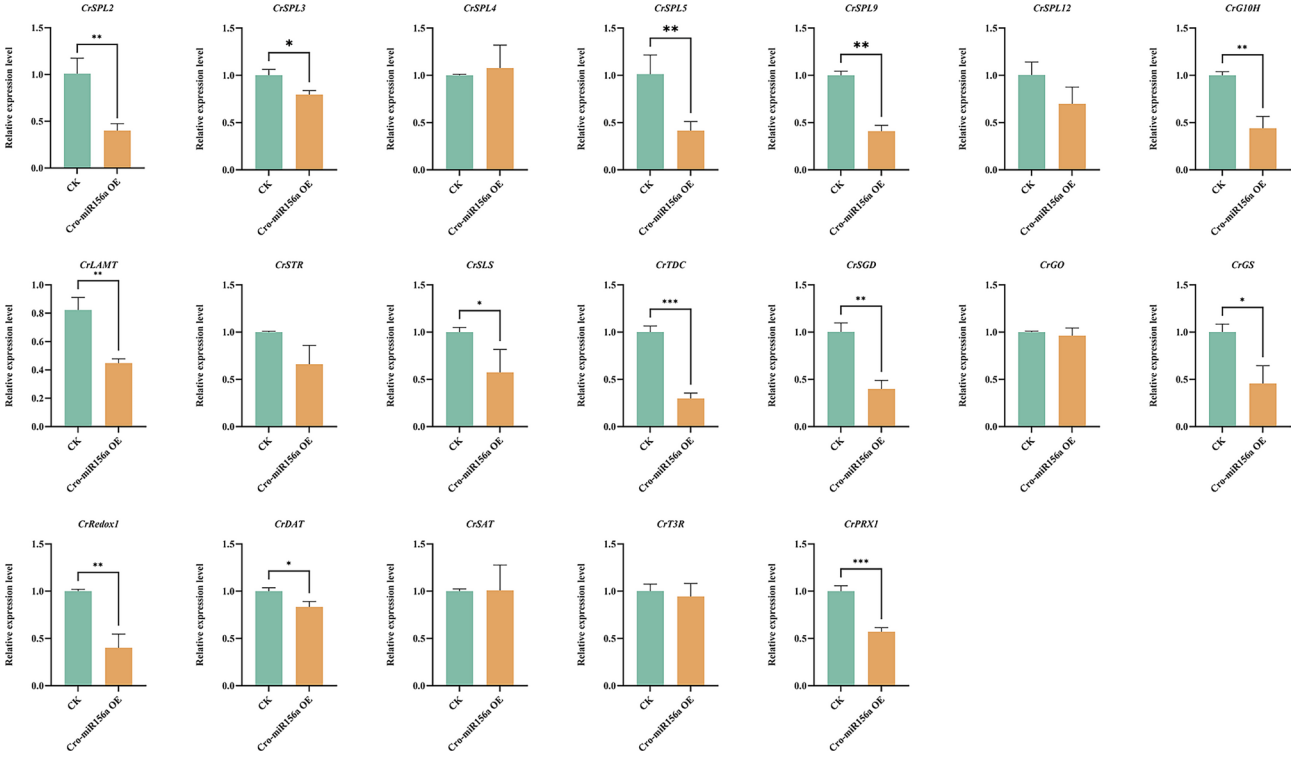

Fig. 11 The overexpression of *Cro-miR156a* in *C.roseus* seedlings. The results represent the mean ± SD of biological triplicates ($n = 3$), and $2^{-\Delta\Delta ct}$ was used to calculate the relative expression level and the *Student's t test* was used for significance test. *$P \le 0.05$, **$P \le 0.01$, ***$P \le 0.001$, **** $P \le 0.0001$. *CrG10H*, geraniol 10-hydroxylase; *CrLAMT*, loganic acid O-methyltransferase; *CrSTR*, strictosidine synthase; *CrSLS*, secologanin synthase; *CrTDC*, tryptophan decarboxylase; *CrSGD*, strictosidine-O-β-d-glucosidase; *CrGO*, geissoschizine oxidase; *CrGS*, geissoschizine synthase; *CrRedox1*, protein redox 1; *CrDAT*, deacetylvindoline-O-acetyltransferase; *CrSAT*, stemmadenine-O-acetyltransferase; *CrT3R*, 16-methoxy-2,3-dihydro-3-hydroxytabersonine synthase; *CrPRX1*, class III peroxidase

vindoline and catharanthine are used in the treatment of diabetes, and serpentine and ajmalicine are used in the clinic as a highly effective antihypertensive drug [37, 38]. The availability of high-quality genome and transcriptome data has made it easier to explore gene function and the TIAs regulatory network in *C. roseus*. Recently, the ZCT family [39]MAPK family [40] and GRF family [41] members' roles in TIAs regulatory have been illustrated. Previous research has shown that *SPL* genes are important for plant growth and development, including flower development, phase transition, signaling transduction, and regulation of secondary metabolites. However, information on the functions of *SPL* genes is still limited and unclear. In this study, we identified 14 *SPL* genes in *C. roseus* through genome-wide identification, without considering alleles (Table 1; Fig. 1). The number of *SPL* genes (Figs. 2 and 3) in *C. roseus* was similar to that in *A. thaliana* (16) and grape (15), but smaller than in wheat (56) [9]. This suggests that *SPL* genes have undergone different gene duplication events in different species.

In *C. roseus*, two gene duplication events occurred: *CrSPL7-CrSPL8* and *CrSPL2-CrSPL9* (Fig. 4). Additionally, we identified fourteen gene pairs with *A. thaliana* and three gene pairs with *O. sativa* during the gene duplication and collinear relationship analysis. These results suggest that *C. roseus* and *A. thanliana* have a conserved evolutionary relationship of *SPL* genes and a close evolutionary distance.

The analysis of cis-elements in the *CrSPL* gene promoter regions revealed the presence of stress, light-signal, cellular development, and plant hormone signal response cis-elements (Fig. 5). The CrSPL genes' expression profiles from different tissue and cell population show their roles in plants growth and development (Fig. 6). Among them, Jasmonate (JA) and its methyl esters, MeJA, are regulators of plant development and key elicitors in metabolite production. MeJA has been shown to participate in the synthesis of TIAs in *C. roseus*, and several *C. roseus* TF members can respond to MeJA (Fig. 7). *CrSPL* genes were also treated by MeJA and responded positively, indicating that JA has the potential to expand its applications by inducing the expression of *CrSPL* genes.

MiRNAs can reduce the expression level of their targets through transcript cleavage or translation repression. The SPL family is typically targeted by miR156/157 and form miRNA-SPL module, which is conserved in plants [36]. There are numerous reports on the multifaceted roles of the miR156-SPL module in different plants. This study

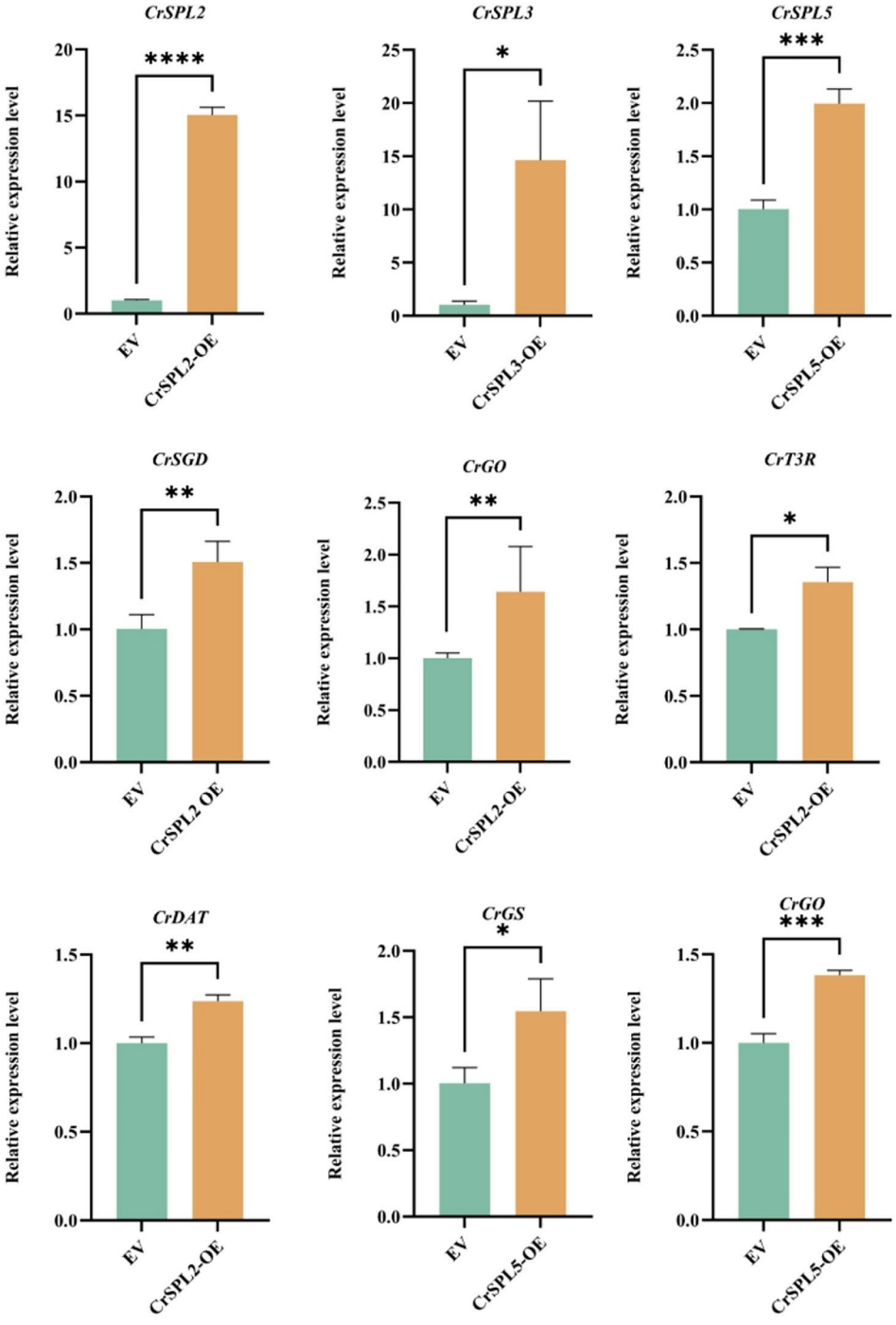

Fig. 12 The overexpression of *CrSPL2/3/5* in *C. roseus* seedling. The results represent the mean ± SD of biological triplicates ($n = 3$), and $2^{-\Delta\Delta ct}$ was used to calculate the relative expression leve and the *Student's t* test was used for significance test. *$P \le 0.05$, **$P \le 0.01$, ***$P \le 0.001$, **** $P \le 0.0001$

focuses on the potential role of the miR156-SPL module in metabolite regulation, rather than growth and development. The SPL genes are involved in some of the secondary metabolic processes, among which the synthesis of anthocyanins is more regulated. The miR156-SPL module can promote the biosynthesis of anthocyanins, flavonoids and flavanols in Populus tremula by targeting and regulating the mRNA expression level of the SPL gene [42]. Fifteen SPL TFs were identified in *Salvia miltiorrhiza*, SmSPL6 can directly bind to the *Sm4CL9* and *SmCYP98A14* and promote the synthesis of Salvianolic acid B (Sal B) and Rosmarinic acid (Ros A) [43]. However, SmSPL7 inhibits the expression of *SmTAT1* and *Sm4CL9* by directly binding to their promoters, blocking Sal B

biosynthesis, and promotes anthocyanin accumulation [44]; AaSPL2 can directly bind the key enzyme *AaDBR2* and promote artemisinin biosynthesis in *Artemisia annua* [45] and SPL9 activates the expression of the sesquiterpene synthase gene *TPS21* by directly binding to its promoter and promote sesquiterpene production in *Pogostemon cablin* [46].

There are limited reports describing the role of miR156/SPL module in TIAs in plants, including *C. roseus*. In this study, we predicted interactions between putative Cro-miRNA and CrSPL. The miR156 family members, *Cro-miR156a*-e, target *CrSPL2*, *CrSPL3*, *CrSPL4*, *CrSPL5*, *CrSPL9*, and *CrSPL12*. We verified that *Cro-miR156a* can efficiently suppress the expression level of *CrSPL2*, *CrSPL3*, *CrSPL4* and *CrSPL5* genes in *N. benthamiana* (Fig. 8). Combining the Y1H (Fig. 10) and Dual-Luc assays (Fig. 9), it was found that CrSPLs members can bind to the specific TIAs gene promoter region and directly activate the Cr7DLGT, and CrLAMT, CrsSGD, CrGS and CrT3R promoter activity, which means these CrSPLs might directly regulate the TIAs biosynthesis.

Obtaining a stable transgenic *C. roseus* is a challenging process, which limits the exploration of the functions of the Cro-miR156/CrSPL modules in *C. roseus*. In this study, we employed the recently described transient overexpression method in the *C. roseus* cotyledon system to ascertain the potential function of *Cro-miR156a* (Fig. 11) and CrSPLs in the regulatory of TIAs. *Cro-miR156a* effectively reduced the expression levels of its targets *CrSPL2/3/5* with the exception of *CrSPL4* and *CrSPL12*. Concurrently, *Cro-miR156a* is also capable of effectively reducing the expression levels of key genes involved in TIAs biosynthesis, including: *CrG10H*, *CrLAMT*, *CrSLS*, *CrTDC*, *CrSGD*, *CrGS*, *CrRedox1*, *CrDAT* and *CrPRX1*. Given the limited regulatory effect of CrSPLs on the TIAs pathway (Fig. 12), it seems plausible to suggest that the observed decrease in the expression level of TIAs pathway genes may be attributed to the reduction in the expression levels of the CrSPLs member genes and their potential coordinators. Nevertheless, further investigation is required to ascertain the potential coordinator of CrSPLs through high-throughput sequencing techniques, including ATAC-seq, single-cell-seq, chip-seq etc. The direct measurement of TIAs via an HPLC-MS system in stable transformed systems, such as hairy roots and suspension cells, will provide more efficient evidence. Collectively, our findings offer novel insights into the potential role of CrSPLs in regulating TIAs biosynthesis.

Conclusion

A comprehensive analysis of the *SPL* gene family was conducted on a whole-genome scale in *Catharanthus roseus*. Fourteen *CrSPL* members were identified and subjected to comprehensive analysis, encompassing gene structure, conservative motifs, evolutionary relationships, gene duplications, microRNA interactions, spatio-temporal expression patterns and MeJA response. Members of the *CrSPL* family have been observed to respond to methyl jasmonate (MeJA). Cro-miR156a was verified to target and decrease the expression levels of the *CrSPL2*, *CrSPL3*, *CrSPL4* and *CrSPL5* genes. It has been demonstrated that CrSPL2/3/4/5/7/8/9/12 can directly activate the activity of certain genes involved in the partial vinblastine pathway and thus can be regarded as a positive regulator of vinblastine. The overexpression of Cro-*miR156a* and *CrSPL2*/3/5 indicates that *Cro-miR156a* functions as a negative regulator of vinblastine, reducing the expression level of vinblastine pathway genes. Additionally, the regulatory influence of *CrSPL2*/5 on TIAs biosynthesis is limited, suggesting the need for additional factors. Overall, the findings of this study contribute to the understanding of TIAs regulation by Cro-*miR156a* and *CrSPLs*.

Abbreviations

C. roseus	*Catharanthus roseus*
A. thaliana	*Arabidopsis thaliana*
O. sativa	*Oryza sativa*
N. benthamiana	*Nicotiana benthamiana*
TIAs	Terpenoid indole alkaloids
MeJA	Methyl jasmonate
Y1H	Yeast One Hybrid

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12870-025-06802-5.

Supplementary Material 1

Supplementary Material 2

Acknowledgements

Not applicable.

Author contributions

Yaojie Zhang: Conceptualization, Data curation, Investigation, Methodology, Software, Writing - original draft. Xiaofan Ni: Data curation, Visualization, Writing - original draft. Xueqing Fu: Investigation, Conceptualization, Software. Ayat Taheri: Data curation, Software, Writing - review & editing. Wei Zhang: Data curation, Software, Writing - review & editing. Pin Liu: Writing-review & editing. Hang Liu: Investigation, Methodology. Ling Li: Investigation, Conceptualization. Yuliang Wang: Project administration. Kexuan Tang: Funding acquisition, Conceptualization, Project administration, Supervision, Writing– review & editing.All authors read and approved the final manuscript.

Funding

This work was supported by National Key R&D Program of China (2023YFC3503901, 2023); National Key R&D Program of China (2018YFA0900600, 2018); the Bill & Melinda Gates Foundation (INV-027291). Under the grant conditions of the Foundation, a Creative Commons Attribution 4.0 Generic License has already been assigned to the Author Accepted Manuscript version that might arise from this submission.

Data availability

All data generated or analysed during this study are included in this published article.

Declarations

Ethics approval and consent to participate

We all affirm that the manuscript reporting studies do not include any human participants, human data, or human tissue.

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

Received: 11 April 2025 / Accepted: 28 May 2025
Published online: 02 July 2025

References

1. Riechmann JL, Ratcliffe OJ. A genomic perspective on plant transcription factors. Curr Opin Plant Biol. 2000;3(5):423–34. https://doi.org/10.1016/s1369-5266(00)00107-2.
2. Klein J, Saedler H, Huijser P. A new family of DNA binding proteins includes putative transcriptional regulators of the Antirrhinum majus floral meristem identity gene SQUAMOSA. Mol Gen Genet. 1996;250(1):7–16. https://doi.org/10.1007/BF02191820
3. Birkenbihl RP, Jach G, Saedler H, Huijser P. Functional dissection of the plant-specific SBP-domain: overlap of the DNA-binding and nuclear localization domains. J Mol Biol. 2005;352(3):585–96. https://doi.org/10.1016/j.jmb.2005.07.013.
4. Li Y, Wang S, Adhikari PB, et al. Evolutionary assessment of SQUAMOSA PROMOTER BINDING PROTEIN-LIKE genes in citrus relatives with a specific focus on flowering. Mol Hortic. 2023;3(1):13. https://doi.org/10.1186/s43897-023-00061-4.
5. Cardon G, Höhmann S, Klein J, Nettesheim K, Saedler H, Huijser P. Molecular characterisation of the Arabidopsis SBP-box genes. Gene. 1999;237(1):91–104. https://doi.org/10.1016/s0378-1119(99)00308-x.
6. Tuskan GA, Difazio S, Jansson S, et al. The genome of black cottonwood, Populus trichocarpa (Torr. & Gray). Science. 2006;313(5793):1596–604. https://doi.org/10.1126/science.1128691
7. Wei H, Zhao Y, Xie Y, Wang H. Exploiting SPL genes to improve maize plant architecture tailored for high-density planting. J Exp Bot. 2018;69(20):4675–88. https://doi.org/10.1093/jxb/ery258.
8. Xie K, Wu C, Xiong L. Genomic organization, differential expression, and interaction of SQUAMOSA promoter-binding-like transcription factors and microRNA156 in rice. Plant Physiol. 2006;142(1):280–93. https://doi.org/10.1104/pp.106.084475.
9. Zhu T, Liu Y, Ma L, et al. Genome-wide identification, phylogeny and expression analysis of the SPL gene family in wheat. BMC Plant Biol. 2020;20(1):420. https://doi.org/10.1186/s12870-020-02576-0.
10. Salinas M, Xing S, Höhmann S, Berndtgen R, Huijser P. Genomic organization, phylogenetic comparison and differential expression of the SBP-box family of transcription factors in tomato. Planta. 2012;235(6):1171–84. https://doi.org/10.1007/s00425-011-1565-y.
11. Cui LG, Shan JX, Shi M, Gao JP, Lin HX. The miR156-SPL9-DFR pathway coordinates the relationship between development and abiotic stress tolerance in plants. Plant J. 2014;80(6):1108–17. https://doi.org/10.1111/tpj.12712.
12. Ioannidi E, Rigas S, Tsitsekian D, et al. Trichome patterning control involves TTG1 interaction with SPL transcription factors. Plant Mol Biol. 2016;92(6):675–87. https://doi.org/10.1007/s11103-016-0538-8.
13. Jung JH, Seo PJ, Kang SK, Park CM. miR172 signals are incorporated into the miR156 signaling pathway at the SPL3/4/5 genes in Arabidopsis developmental transitions. Plant Mol Biol. 2011;76(1–2):35–45. https://doi.org/10.1007/s11103-011-9759-z
14. Zhang Y, Schwarz S, Saedler H, Huijser P. SPL8, a local regulator in a subset of gibberellin-mediated developmental processes in Arabidopsis. Plant Mol Biol. 2007;63(3):429–39. https://doi.org/10.1007/s11103-006-9099-6
15. Yang Z, Wang X, Gu S, Hu Z, Xu H, Xu C. Comparative study of SBP-box gene family in Arabidopsis and rice. Gene. 2008;407(1–2):1–11. https://doi.org/10.1016/j.gene.2007.02.034
16. Shikata M, Koyama T, Mitsuda N, Ohme-Takagi M. *Arabidopsis* SBP-box genes SPL10, SPL11 and SPL2 control morphological change in association with shoot maturation in the reproductive phase. Plant Cell Physiol. 2009;50(12):2133–45. https://doi.org/10.1093/pcp/pcp148
17. Jung JH, Ju Y, Seo PJ, Lee JH, Park CM. The SOC1-SPL module integrates photoperiod and gibberellic acid signals to control flowering time in *Arabidopsis*. Plant J. 2012;69(4):577–88. https://doi.org/10.1111/j.1365-313X.2011.04813.x
18. Unte US, Sorensen AM, Pesaresi P, et al. SPL8, an SBP-box gene that affects pollen sac development in *Arabidopsis*. Plant Cell. 2003;15(4):1009–19. https://doi.org/10.1105/tpc.010678
19. Stone JM, Liang X, Nekl ER, Stiers JJ. *Arabidopsis* AtSPL14, a plant-specific SBP-domain transcription factor, participates in plant development and sensitivity to Fumonisin B1. Plant J. 2005;41(5):744–54. https://doi.org/10.1111/j.1365-313X.2005.02334.x
20. Miura K, Ikeda M, Matsubara A, et al. OsSPL14 promotes panicle branching and higher grain productivity in rice. Nat Genet. 2010;42(6):545–9. https://doi.org/10.1038/ng.592.
21. Wang S, Wu K, Yuan Q, et al. Control of grain size, shape and quality by OsSPL16 in rice. Nat Genet. 2012;44(8):950–4. https://doi.org/10.1038/ng.2327.
22. Chuck GS, Brown PJ, Meeley R, Hake S. Maize SBP-box transcription factors unbranched2 and unbranched3 affect yield traits by regulating the rate of lateral primordia initiation. Proc Natl Acad Sci U S A. 2014;111(52):18775–80. https://doi.org/10.1073/pnas.1407401112.
23. Chuck G, Whipple C, Jackson D, Hake S. The maize SBP-box transcription factor encoded by tasselsheath4 regulates bract development and the establishment of meristem boundaries. Development. 2010;137(8):1243–50. https://doi.org/10.1242/dev.048348.
24. Wang H, Nussbaum-Wagler T, Li B, et al. The origin of the naked grains of maize. Nature. 2005;436(7051):714–9. https://doi.org/10.1038/nature03863.
25. Bartel DP. MicroRNAs: genomics, biogenesis, mechanism, and function. Cell. 2004;116(2):281–97. https://doi.org/10.1016/s0092-8674(04)00045-5.
26. Ott CE, Grünhagen J, Jäger M, et al. MicroRNAs differentially expressed in postnatal aortic development downregulate Elastin via 3' UTR and coding-sequence binding sites. PLoS ONE. 2011;6(1):e16250. https://doi.org/10.1371/journal.pone.0016250.
27. Liu Y, Patra B, Singh SK, et al. Terpenoid Indole alkaloid biosynthesis in *Catharanthus roseus*: effects and prospects of environmental factors in metabolic engineering. Biotechnol Lett. 2021;43(11):2085–103. https://doi.org/10.1007/s10529-021-03179-x
28. Schweizer F, Colinas M, Pollier J, et al. An engineered combinatorial module of transcription factors boosts production of monoterpenoid Indole alkaloids in *Catharanthus roseus*. Metab Eng. 2018;48:150–62. https://doi.org/10.1016/j.ymben.2018.05.016
29. Singh SK, Patra B, Paul P, Liu Y, Pattanaik S, Yuan L. Revisiting the ORCA gene cluster that regulates terpenoid Indole alkaloid biosynthesis in *Catharanthus roseus*. Plant Sci. 2020;293:110408. https://doi.org/10.1016/j.plantsci.2020.110408
30. Singh SK, Patra B, Paul P, Liu Y, Pattanaik S, Yuan L. BHLH IRIDOID SYNTHESIS 3 is a member of a bHLH gene cluster regulating terpenoid Indole alkaloid biosynthesis in *Catharanthus roseus*. Plant Direct. 2021;5(1):e00305. https://doi.org/10.1002/pld3.305
31. Schluttenhofer C, Pattanaik S, Patra B, Yuan L. Analyses of *Catharanthus roseus* and *Arabidopsis thaliana* WRKY transcription factors reveal involvement in jasmonate signaling. BMC Genomics. 2014;15(1):502. https://doi.org/10.1186/1471-2164-15-502
32. Pauw B, Hilliou FA, Martin VS, et al. Zinc finger proteins act as transcriptional repressors of alkaloid biosynthesis genes in *Catharanthus roseus*. J Biol Chem. 2004;279(51):52940–8. https://doi.org/10.1074/jbc.M404391200
33. van der Fits L, Zhang H, Menke FL, Deneka M, Memelink J. A *Catharanthus roseus* BPF-1 homologue interacts with an elicitor-responsive region of the secondary metabolite biosynthetic gene Str and is induced by elicitor via a JA-independent signal transduction pathway. Plant Mol Biol. 2000;44(5):675–85. https://doi.org/10.1023/a:1026526522555.
34. Shen EM, Singh SK, Ghosh JS, et al. The MiRNAome of *Catharanthus roseus*: identification, expression analysis, and potential roles of MicroRNAs in regulation of terpenoid Indole alkaloid biosynthesis. Sci Rep. 2017;7:43027. https://doi.org/10.1038/srep43027
35. Liu Y, Lyu R, Singleton JJ, Patra B, Pattanaik S, Yuan L. A Cotyledon-based Virus-Induced gene Silencing (Cotyledon-VIGS) approach to study specialized metabolism in medicinal plants. Plant Methods. 2024;20(1):26. https://doi.org/10.1186/s13007-024-01154-x.

36. Yang X, Dong W, Ren W, Zhao Q, Wu F, He Y. Cytoplasmic HYL1 modulates miRNA-mediated translational repression. Plant Cell. 2021;33(6):1980–96. https://doi.org/10.1093/plcell/koab090.
37. Noble CO, Guo Z, Hayes ME, et al. Characterization of highly stable liposomal and immunoliposomal formulations of vincristine and vinblastine. Cancer Chemother Pharmacol. 2009;64(4):741–51. https://doi.org/10.1007/s00280-008-0923-3.
38. Wang Q, Zhao X, Jiang Y, Jin B, Wang L. Functions of representative terpenoids and their biosynthesis mechanisms in medicinal plants. Biomolecules. 2023;13(12):1725. https://doi.org/10.3390/biom13121725.
39. Traverse KKF, Breselge S, Trautman JG, et al. Characterization of the zcts, a subgroup of Cys2-His2 zinc finger transcription factors regulating alkaloid biosynthesis in *Catharanthus roseus*. Plant Cell Rep. 2024;43(9):209. https://doi.org/10.1007/s00299-024-03295-8
40. Gao X, Zhu X, Wang Z, et al. Modulation of terpenoid Indole alkaloid biosynthesis in *Catharanthus roseus* by *Sphingomonas Sp* Y503 via the CrMAPKKK1-CrMAPKK1/CrMAPKK2-CrMPK3 signaling cascade. Plant Cell Environ. 2025;48(2):1692–704. https://doi.org/10.1111/pce.15253
41. Chang C, Guo X, Wang B et al. CrGRF1/4 mediating light signal to regulate monoterpenoid Indole alkaloid biosynthesis in *Catharanthus roseus*. Plant Cell Environ. https://doi.org/10.1111/pce.15484
42. Wang Y, Liu W, Wang X, et al. MiR156 regulates anthocyanin biosynthesis through SPL targets and other MicroRNAs in Poplar. Hortic Res. 2020;7:118. https://doi.org/10.1038/s41438-020-00341-w.
43. Cao Y, Chen R, Wang WT, Wang DH, Cao XY. SmSPL6 induces phenolic acid biosynthesis and affects root development in Salvia miltiorrhiza. Int J Mol Sci. 2021;22(15):7895. https://doi.org/10.3390/ijms22157895
44. Chen R, Cao Y, Wang W, et al. Transcription factor SmSPL7 promotes anthocyanin accumulation and negatively regulates phenolic acid biosynthesis in Salvia miltiorrhiza. Plant Sci. 2021;310:110993. https://doi.org/10.1016/j.plantsci.2021.110993
45. Lv Z, Wang Y, Liu Y, et al. The SPB-Box transcription factor AaSPL2 positively regulates Artemisinin biosynthesis in Artemisia annua L. Front Plant Sci. 2019;10:409. https://doi.org/10.3389/fpls.2019.00409
46. Yu ZX, Wang LJ, Zhao B, et al. Progressive regulation of sesquiterpene biosynthesis in *Arabidopsis* and Patchouli (*Pogostemon cablin*) by the miR156-targeted SPL transcription factors. Mol Plant. 2015;8(1):98–110. https://doi.org/10.1016/j.molp.2014.11.002

Publisher's note