Genome-wide characterization of *SPL* gene family in *Codonopsis pilosula* reveals the novel roles of *CpSPL2* and *CpSPL10* in promoting the accumulation of secondary metabolites and growth of *C. pilosula* hairy root

Jing Yang
Shaanxi Normal University
Zhonglong Guo
Peking University
Yao Cao
Shaanxi Normal University
Rui Chen
Shaanxi Normal University
Wentao Wang
Shaanxi Normal University
Xiaozeng Yang
Beijing Academy of Agriculture and Forestry Sciences
xiaoyan cao (caoxiaoyan@snnu.edu.cn)
Shaanxi Normal University https://orcid.org/0000-0002-3188-6663

Research article

Keywords: Biomass, Codonopsis pilosula, Expression patterns, Hairy root, SPLs, Secondary metabolites

Posted Date: November 24th, 2020

DOI: https://doi.org/10.21203/rs.3.rs-81496/v1

Abstract

Background

SQUAMOSA PROMOTER BINDING PROTEIN-LIKE (SPL) transcription factors play critical roles in regulating diverse aspects of plant growth and development, including vegetative phase change, plant architecture, anthocyanin accumulation, lateral root growth, etc. *Codonopsis pilosula* is a famous medicinal plant and its dried root, named Dangshen, is one of the most widely used traditional Chinese medicine. However, little information about *SPL* genes in this species has been reported.

Results

In the present study, 15 *SPL* genes were identified based on the genome data of *Codonopsis pilosula*. Ten of the 15 *CpSPLs* were predicted to be the targets of miR156. Phylogenetic analysis clustered CpSPLs into seven groups (G1-G7) along with 16 SPLs from *Arabidopsis thaliana*. *CpSPLs* in the same group share similar gene structure and conserved motif composition. Cis-acting elements responding to light, stress, and phytohormone widely exist in their promoter regions. Our qRT-PCR results indicated that 15 *CpSPL*s were differentially expressed in different tissues (root, stem, leaf, flower, and calyx), different developmental periods (1, 2 and 3 months after germination), and various conditions (NaCl, MeJA and ABA treatment). Compared with the control, overexpression of *CpSPL2* or *CpSPL10* significantly promoted not only the growth of hairy roots, but also the accumulation of total saponins and lobetyolin.

Conclusions

The SPL genes in the *C. pilosula* genome were identified and their expression patterns were analyzed. The novel roles of *CpSPL2* and *CpSPL10* in promoting the accumulation of secondary metabolites and growth of *C. pilosula* hairy root were revealed. Our results established a foundation for further investigation of *CpSPL*s and provided novel insights into their biological functions.

Background

Transcription factors (TFs) function in various physiological and developmental processes via activating and/or repressing transcription of multiple target genes [1]. They have been usually divided into different families according to the sequence of DNA-binding domains and other conserved motifs [2]. SQUAMOSA-promoter binding protein-like (SPL or SBP) TFs are exclusive to plant and characterized by a highly conserved SBP domain and a nuclear localization signal (NLS) at the C-terminus. The SBP domain is approximately 76 amino acids and includes two zinc-binding sites (one zinc finger is C3H or C4, and the other is C2H4) essential for DNA binding, and the NLS partially overlapps with the second zinc finger [3–6]. AmSBP1 and AmSBP2 from *Antirrhinum majus* were the first discovered SBP-domain proteins in plants, and were found to bind to the floral meristem identity gene SQUAMOSA promoter, so named them [3]. Then *SPL* genes have been identified in many plant species, including single-cell algae, mosses, gymnosperms, and angiosperms [7]. With the rapid implication of high-throughput sequencing technology, more and more plant genome data have been released and genome-wide identification of the *SPL* gene family from model and non-model plants have been identified in *Arabidopsis thaliana* [8], *Oryza sativa* [9], *Glycine max* [10], *Solanum lycopersicum* [11], *Malus domestica* [12], *Salvia miltiorrhiza* [13], *Vitis vinifera* [14], *Phyllostachys edulis* [15], *Capsicum annuum* [16], *Ricinus communis* [17], etc.

The functions of *SPL* genes have been well characterized in the model plant *Arabidopsis* and they play important regulatory roles in diverse developmental progresses, including vegetative to reproductive phase transition, cotyledon- to vegetative-leaf transition, micro- and megasporogenesis, trichome formation, stamen filament elongation, axillary bud formation, and lateral root growth [18–23]. Besides, they are involved in copper homeostasis, abiotic stress response, immune response, and secondary metabolites production [24–27]. The functions of *SPL* genes from other species have also been identified. In rice, *OsSPL14* has been found to promote panicle branching and grain productivity and *OsSPL16* regulates grain yield and quality

[28, 29]. *FvSPL10* from strawberry (*Fragaria vesca*) not only promotes early flowering, but also increases organs size, such as longer root, larger floral organ and seeds [30]. As a class of plant-specific gene family, some *SPL* genes are important candidates for improving plant agronomic traits by genetic engineering.

Codonopsis pilosula is a member of the Campanulaceae family. Its dried root, named "Dangshen" in Chinese, is one of the most widely used traditional Chinese medicine for replenishing qi (vital energy), strengthening body immunity, improving appetite, promoting gastrointestinal function, reducing blood pressure, and curing gastric ulcers [31]. In addition, Dangshen is also a well-known health-care food in China and is listed in the “Food and Drug Homology Catalogue” approved by the National Health Commission of People’s Republic of China. Consequently, the demand for Dangshen is growing, and the yield and accumulation of bioactive metabolites is attracting more and more attention in planting field [32, 33]. Lobetyolin, alkaloids, polysaccharides, and saponins are the major active ingredients in Dangshen, which are responsible for most of the pharmacological functions found in the medicine [34]. Lobetyolin, a general marker compound in Dangshen, has been well reported to exert multiple bioactivities, such as anti-cancer, antiviral, anti-inflammatory, anti-oxidative, mucosal protective, and xanthine oxidase inhibiting properties [35, 36].

Although *C. pilosula* has received great attention on the chemical constituents and their pharmacological activities, relevant study of this species at the genetic level is lagging behind and only a few literatures involved genes in *C. pilosula* [37–40]. Until now, *SPL* gene family has never been reported in *C. pilosula*. Most recently, we have developed an efficient *Agrobacterium rhizogenes*-mediated transformation approach for transgenic hairy roots with this species [39], which lay a good foundation for genetic engineering of that species. Here, we identified 15 *SPL* genes based on the genome sequence of *C. pilosula* (data unpublished). Gene structure, conserved motif, target prediction of miR156, and cis-acting elements of 15 *CpSPLs* were systematically analyzed. And their spatiotemporal expression profiles in different tissues and expression patterns under various conditions (NaCl, MeJA and ABA treatment) were analyzed by qRT-PCR. Furthermore, we obtained *CpSPL2* or *CpSPL10* overexpressing transgenic hairy roots, and a significant increase was observed in the biomass and concentrations of total saponins and lobetyolin. As far as we know, this is the first experimental research on gene function in this species. These findings demonstrate that *CpSPL2* and *CpSPL10* positively regulate the growth of hairy roots and accumulation of active ingredients, which have great potential in improving the yield and quality of Dangshen.

Methods

Identification of SPL genes in C. pilosula and bioinformatic analysis

The sequences of SBP domain (ID: PF03110), which were downloaded from Pfam database (http://pfam.xfam.org/), was used to search possible SPL genes in *C. pilosula* genome sequences (data unpublished) by HMMER (http://hmmer.org) with the e-value < 1e-10. A total of 15 *CpSPL* genes containing a complete SBP domain were identified.

The online analysis software psRNATarget (http://plantgrn.noble.org/psRNATarget/analysis?function=3) was used to predict *CpSPL* genes directly targeted by miR156, with the maximum expectation value of 3.0 and UPE value of 16. MEGA X software (https://mega.nz/) was used to construct the phylogenetic tree of 31 full-length SPL amino acid sequences, 15 from *C. pilosula* and 16 from *A. thaliana*, with 1000 bootstraps in the Maximum-likelihood (ML) method. Gene Structure Display Server (http://gsds.cbi.pku.edu.cn/) was used for gene structure analysis. The MEME program (http://meme-suite.org/) was used to for identification of the conserved motifs. The cis-acting elements of 15 *CpSPL* genes promoter regions (2000 bp upstream of the translation initiation codon “ATG”) were analyzed online (http://bioinformatics.psb.ugent.be/webtools/plantcare/html/).

Plant materials and treatments

Seeds of *C. pilosula* were collected from Gansu Province, China. The botanical origin of the materials was identified by Professor ZheZhi Wang in Shaanxi Normal University. The specimens of the seeds were deposited in the herbarium of National Engineering Laboratory for Resource Development of Endangered Crude Drugs in Northwest of China, Shaanxi Normal University, Xi’an, China. The Seed owner allowed the collection. The seeds of *C. pilosula* were germinated and incubated according to the method that we described previously [39].

For gene spatiotemporal expression analysis, the leaves, stems, and roots were collected separately from one-, two-, and three-month-old seedlings, and the flower and calyx were collected from the plants at the flowering stage. To test *CpSPL*s responses to hormonal and stress treatments, two-week-old seedings were treated with 200 mmol NaCl, 200 μmol MeJA, and 100 μmol ABA, respectively, as we have described previously [40]. The control group was treated with the same amount of ddH_2O and all the samples were collected 6 h after treatment.

Gene expression analysis

For qRT-PCR analysis, total RNA was extracted and then reverse transcripted into cDNA as we described previously [40]. All the primer sequences used for qRT-PCR were listed in Table S1 and *CpGAPDH* was used as the internal control [40]. The relative expression levels of 15 *CpSPL*s were calculated according to the method described by Livak and Schmittgen (2001) [41]. All the experiments included three biological and three technical replicates.

Vector construction and hairy root transformation

The complete open reading frames of *CpSPL2* and *CpSPL10* were amplified through PCR using the specific primer pairs *CpSPL2*-F/R and *CpSPL10*-F/R (Table S1), respectively, with the following PCR conditions: 98 °C, 3 min; 30 cycles of 98 °C, 10 s, 58 °C, 30 s, 72 °C, 45 s; 72 °C, 5 min. Then the products were digested with *Pac* I and *Asc* I and ligated into pMDC85 to generate overexpression (OE) vectors pMDC85-*CpSPL2* and pMDC85-*CpSPL10*.

Transgenic hairy roots overexpressing *CpSPL2* or *CpSPL10* were obtained by *Agrobacterium*-mediated method according to the protocol established in our lab and selected on the MS medium containing 2 mg/L hygromycin [39]. In parallel, pMDC85 was introduced into *C. pilosula* as the empty vector control (EV). Every transgenic line was excised and sub-cultured separately as we described previously [39]. Five independent *CpSPL2*-OE lines, seven *CpSPL10*-OE lines, and four EV lines were obtained, and then confirmed by genomic DNA PCR using primers *hptII*-F/R (Table S1) for hygromycin phosphotransferase II gene (*hptII*), followed by expression analysis of *CpSPL2* or *CpSPL10* by qRT-PCR.

Determination of lobetyolin and total saponins

Transgenic hairy roots sub-cultured for one month were used for determination of lobetyolin and total saponins.

To determine the concentration of lobetyolin, we ground the dried hairy roots into powder, followed by extracted three times with 10 mL methanol by sonication (50–150 W) in an ultrasonic bath (Kunshan Instrument Co., Ltd., China) for 30 min, 20 min, and 15 min, respectively. The extracts were put together and the methanol solution was evaporated, followed by dissolved with methanol to 5 mL volumetric flask. After filtration with 0.22 μm microporous membrane, the solution was used for HPLC analysis on a Shimadzu LC-20A instrument (Shimadzu, Japan) equipped with an Agilent 5 TC-C_{18} column (250 mm × 4.6 mm, 5 μm). The mobile phase consisted of ultrapure water (A) and methanol (B) and the gradient condition was 0–5 min, 20–40% B; 5–10 min, 40–70% B; 10–12 min, 79–90% B; 12–25 min, 90% B. The separation was performed at 30 °C, with the flow rate of 1.0 mL/min and UV detector wavelength at 220 nm.

The concentration of total saponins in transgenic hairy roots was determined as we described previously [39].

Statistical analysis

All the experiments and date presented here involved at least three biological repeats. SPSS version 20.0 software (SPSS Inc., Chicago, IL, USA) was used for statistical evaluation. The error bars indicate standard deviation. Significant difference of the mean values was set at $P < 0.05$.

Results

Genome-wide identification and sequence feature analysis of CpSPLs

To identify possible *SPL* genes in *C. pilosula* genome sequences, we employed the SBP domain (PF03110) to search the databases by HMMER. A total of 15 *SPLs* containing complete SBP domain were identified based on the genome sequence of

C. pilosula and their cDNA sequences were listed in (Table S2). Consulting the homologous AtSPLs in Arabidopsis, 15 *CpSPLs* were named from *CpSPL1* to *CpSPL15*. The deduced CpSPLs exhibited great variations in terms of their molecular weight (MW), ranging from 17.98 KDa (*CpSPL5*) to 119.80 KDa (*CpSPL14*). Similarly, the lengths of the CDS were found to be varied in the *CpSPLs*, from 480 bp (*CpSPL5*) to 3276 bp (*CpSPL14*). The detailed information, including the gene length, intron number, protein length, predicted MW, and theoretical isoelectric point (pI), were listed in Table 1.

Table 1
The information of 15 *SPL* genes in *Codonopsis pilosula*

Gene name	No. intron	Gene length (bp)	CDS length (bp)	Protein siza(aa)	Mw (Da)	pI	Atomic composition	GRAVY	Instability Index
CpSPL1	10	10066	3036	1011	112293.93	7.10	$C_{4918}H_{7801}N_{1421}O_{1515}S_{38}$	-0.395	54.03
CpSPL2	4	3264	879	292	32767.81	8.74	$C_{1414}H_{2225}N_{415}O_{447}S_{18}$	-0.534	67.85
CpSPL3	2	6640	1095	364	40962.56	8.70	$C_{1778}H_{2735}N_{545}O_{544}S_{16}$	-0.708	58.47
CpSPL4	1	3004	618	205	23029.53	6.20	$C_{964}H_{1560}N_{310}O_{324}S_{11}$	-0.993	68.72
CpSPL5	1	4163	480	159	17978.16	9.26	$C_{753}H_{1218}N_{252}O_{239}S_{11}$	-1.005	44.17
CpSPL6	3	6537	1602	533	58148.49	8.55	$C_{2492}H_{3928}N_{746}O_{817}S_{23}$	-0.662	47.64
CpSPL7	10	13540	2391	796	89393.21	6.52	$C_{3932}H_{6209}N_{1099}O_{1182}S_{50}$	-0.36	58.97
CpSPL8	3	2322	957	318	35386.95	8.40	$C_{1531}H_{2332}N_{454}O_{490}S_{14}$	-0.777	59.08
CpSPL9	2	19029	1140	379	40795.13	8.38	$C_{1767}H_{2718}N_{530}O_{558}S_{15}$	-0.713	63.79
CpSPL10	4	7056	1383	460	50220.85	8.40	$C_{2185}H_{3399}N_{631}O_{698}S_{17}$	-0.577	52.76
CpSPL11	4	3304	1077	358	40255.95	8.69	$C_{1731}H_{2717}N_{535}O_{541}S_{18}$	-0.717	54.19
CpSPL12	3	3095	969	322	35688.83	9.11	$C_{1524}H_{2415}N_{463}O_{493}S_{18}$	-0.641	60.11
CpSPL13	2	3364	1176	391	43540.24	6.52	$C_{1874}H_{2908}N_{552}O_{608}S_{20}$	-0.709	61.52
CpSPL14	10	6316	3276	1091	119799.88	8.47	$C_{5200}H_{8251}N_{1549}O_{1616}S_{46}$	-0.479	51.64
CpSPL15	2	10641	1140	379	40765.11	8.38	$C_{1766}H_{2716}N_{530}O_{557}S_{15}$	-0.712	63.05

Prediction of CpSPLs **targeted by** miR156

miR156, one of the most conserved miRNA families, plays very important roles in the process of plant growth and development by direct cleavage of *SPL* transcripts [42]. In the model plant Arabidopsis, 10 of 16 *AtSPLs* are direct targets of AtmiR156 [42]. In rice, 11 of 19 *OsSPL* genes are targeted by OsmiR156 [9]. We predicted *CpSPL*s targeted by AtmiR156 using the on line plant target prediction tool. The prediction result indicated that a total of 10 *CpSPLs* were the targets of AtmiR156, eight of which (including *CpSPL3, CpSPL6, CpSPL9-13*, and *CpSPL15*) were targeted in the coding regions, while two (*CpSPL4* and *CpSPL5*) in 3' UTR regions (Fig. 1).

Phylogenetic analysis of CpSPLs

We constructed a phylogenetic tree of 15 CpSPLs and 16 AtSPLs using MEGA X with ML method. As shown in Fig. 2, 31 SPLs from two species were classfied into seven groups, named from G1 to G7, and each group consisted of at least one SPL from *C. pilosula* and one from *A. thaliana*. There are only two members in G3 (*AtSPL8* and *CpSPL8*) and G7 (*AtSPL7* and *CpSPL7*). G5 is the largest group with three *AtSPLs* (*AtSPL6/9/15*) and four *CpSPLs* (*CpSPL3/6/9/15*). It was interesting that all the members in

G1, G3, and G7 belong to non-miR156-targeted *SPLs*, including six *AtSPLs* (*AtSPL1/7/8/12/14/16*) and four *CpSPLs* (*CpSPL1/7/8/14*). Except for *CpSPL2*, All the members in G2, G4, G5, and G6 were miR156-targeted *SPLs*. In *Arabidopsis*, *AtSPL2/10/11*, three members closely related, regulate root regeneration by inhibiting auxin biosynthesis [43]. Phylogenetic tree clustered *CpSPL2*, *CpSPL10*, *AtSPL2*, *AtSPL10*, and *AtSPL11* in G4, indicating *CpSPL2* and *CpSPL10* are probably involved in root growth.

Gene structure and conserved motif analysis

To clarify the structural diversities of 15 *CpSPL*s, we performed gene exon/intron structure analysis. The result displayed that the number of introns had a high variation and ranged from one to ten (Fig. 3). Interestingly, we found that most *CpSPLs* in the same group share similar structure. For instance, *CpSPL1* and *CpSPL14*, belonging to G1, have ten introns, respectively. *CpSPL4* and *CpSPL5*, members in G6, have only one intron, respectively (Fig. 3).

To explore the conserved motifs, 15 *CpSPLs* were subjected to analysis with MEME program. Among the 12 conserved motifs identified (Fig. 4, Table S3), motif 1, motif 2, and motif 3 existed in all the 15 CpSPLs and formed the conserved SBP domain. Similar motif composition existed in the same group. For example, *CpSPL2* and *CpSPL10* in G4 all consisted of five conserved motifs (motif 1/2/3/10/11). The motif composition in *CpSPL9* was completely consistent with that in *CpSPL15*, suggesting that *CpSPL9* and *CpSPL15* probably have similar and redundant functions in plant development.

Cis-acting elements analysis of CpSPLs promoter regions

We analyzed the cis-acting elements of 15 *CpSPLs* promoter regions and light responsive elements (including G-box, GATA-motif, GTGGC-motif, AE-box, TCT-motif, and chs-CMA2a), hormone responsive elements [such as gibberellin (GARE-motif), MeJA (CGTCA- and TGACG-motif), and abscisic acid (ABA) (ABRE)], stress responsive elements [such as drought (MBS), low-temperature (LTR), and anaerobic induction (ARE)], and CAT box related to meristem expression were found in their promoter regions (Table S4). Among these cis-elements, MeJA-responsive elements existed in the promoter regions of almost all the *CpSPLs* except for *CpSPL9* and *CpSPL13*, and ABA-responsive element (ABRE) existed in the promoter regions of 10 *CpSPLs* (including *CpSPL1*, *CpSPL6-10*, *CpSPL12*, and *CpSPL14-15*).

Spatiotemporal expression analysis of CpSPL genes

We investigated the expression patters of 15 *CpSPLs* in the leaves, stems, and roots from one-, two-, and three-month-old seedlings, and the flower and calyx from the plants at the flowering stage by qRT-PCR assay. The results showed that most *CpSPLs* expressed in almost all the tissues (Fig. 5). Compared with other genes, the expression level of *CpSPL7* was more constant in all the tissues tested. *CpSPL8* showed highest level in calyx. *CpSPL5* was expressed at relatively higher levels in leaf and calyx. The expression levels of *CpSPL3, CpSPL8, CpSPL10, CpSPL12*, and *CpSPL13* in the stems gradually decreased with the maturation of the seedlings. *CpSPL1* and *CpSPL14*, two members in G1, showed similar expression patterns and their expression levels in the root increased gradually with the maturation of the seedlings. In addition, the expression patterns of *CpSPL9* and *CpSPL15* were highly similar, with higher levels in flowers and 3-month-old roots. In summary, spatiotemporal expression analysis results indicated that *CpSPL* genes exhibited various expression patterns, which provide preliminary information for understanding their potential functions in the development of *C. pilosula*.

Expression profiles of CpSPLs under various conditions

To assess the expression profiles of 15 *CpSPL* genes under various treatments (NaCl, MeJA, and ABA), a histogram was generated using the relative expression level (Fig. 6). When treated with NaCl, the transcript levels of eight *CpSPLs* (*CpSPL1, CpSPL2, CpSPL4, CpSPL6, CpSPL7, CpSPL11, CpSPL14* and *CpSPL15*) and four *CpSPLs* (*CpSPL5, CpSPL8, CpSPL10*, and *CpSPL13*) were significantly up-regulated and down-regulated, respectively. Among those, *CpSPL2, CpSPL6*, and *CpSPL11* increased to 6.02, 5.66, and 7.94 times than the control, respectively, while *CpSPL5* and *CpSPL8* decreased to 20.00 and 12.50 times than the control, respectively (Fig. 6A). For MeJA treatment, the transcript levels of *CpSPL4*, *CpSPL6*, *CpSPL14* and *CpSPL15* significantly increased, with the highest change folds in *CpSPL15* (3.03 folds). The transcript levels of five *CpSPL*

genes (*CpSPL3, CpSPL5, CpSPL8*, *CpSPL10*, and *CpSPL11*) significantly decreased, with 14.28, and 10.00 change folds in *CpSPL5* and *CpSPL8*, respectively (Fig. 6B). Under ABA treatment, eight *CpSPLs* (*CpSPL1, CpSPL4, CpSPL6, CpSPL7, CpSPL9, CpSPL12*, *CpSPL14* and *CpSPL15*) responded positively to the treatment, while three genes (*CpSPL3*, *CpSPL5* and *CpSPL8*) responded negatively to ABA treatment. Among those genes, *CpSPL15* and *CpSPL8* exhibited highest upregulation and downregulation, respectively (Fig. 6C).

Overexpression of CpSPL2 **or** CpSPL10 **promotes the growth of** C. pilosula **hairy root**

To investigate the function of *CpSPL2* and *CpSPL10* in root development, we generated *CpSPL2*-overexpressing or *CpSPL10*-overexpressing transgenic hairy roots. The expression level of *CpSPL2* or *CpSPL10* in the transgenics was examined by qRT-PCR (Fig. 7A, B). Two independent *CpSPL2*-overexpressing lines (*CpSPL2*-OE3 and *CpSPL2*-OE5) and *CpSPL10*-overexpressing lines (*CpSPL10*-OE2 and *CpSPL2*-OE3) with dramatically elevated *CpSPL2* or *CpSPL10* expression were selected for further analysis. In comparison to the control, the hairy roots overexpressing *CpSPL2* or *CpSPL10* grew faster (Fig. 7C, D, and E). When the transgenic hairy roots with the length about 1.0 cm were cultured for one month, the biomass of *CpSPL2*-OE3, *CpSPL2*-OE5, *CpSPL10*-OE2, and *CpSPL10*- OE3 was 2.19, 1.98, 3.15, and 2.83 times that of the control (EV2), respectively (Fig. 7F). Our results indicated that both *CpSPL2* and *CpSPL10* promote the growth of hairy roots.

Overexpression of CpSPL2 **or** CpSPL10 **promotes accumulation of lobetyolin and total saponins in** C. pilosula **hairy root**

To evaluate the impact of *CpSPL2* or *CpSPL10* on active ingredients, HPLC and UV spectrophotometer were used to determine the concentrations of lobetyolin and total saponins in those transgenic lines, respectively. It was surprising that the production of both lobetyolin and total saponins were greatly increased in *CpSPL2*-OE or *CpSPL10*-OE lines. The concentration of lobetyolin in *CpSPL2*-OE3, *CpSPL2*-OE5, *CpSPL10*-OE2, and *CpSPL10*-OE3 was 6.43, 6.25, 6.29, and 7.03 times that of the control (EV2), respectively (Fig. 8A). The concentration of total saponins in *CpSPL2*-OE3, *CpSPL2*-OE5, *CpSPL10*-OE2, and *CpSPL10*-OE3 was 3.18, 2.72, 1.81, and 1.94 times that of the control (EV2), respectively (Fig. 8B). In summary, *CpSPL2* and *CpSPL10* promote not only the growth of hairy roots but also accumulations of lobetyolin and total saponins.

Discussion

Identification of SPL genes in C. pilosula

SPLs are plant-specific TFs and characterized by a highly conserved SBP domain [5, 6]. They play critical roles in regulating diverse aspects of plant growth and development, including vegetative phase change, plant architecture, anthocyanin accumulation, lateral root growth, etc [18–24]. Since its first discovery in *A. majus* [3], the SPL gene family from various plants has been isolated and identified. For instance, there are 16 SPL gene family members in *Arabidopsis thaliana* [8], 19 in *Oryza sativa* [9], 15 in *Solanum lycopersicum* [11], 15 in *Salvia miltiorrhiza* [13], and 15 in *Ricinus communis* [17]. However, little information is known about SPL gene family in *C. pilosula*, a famous species with important medical and edible values. Here, we identified 15 *CpSPL* genes in *C. pilosula* genome. Since most *SPL* genes are targets of miR156 [42], we predicted miR156-targeted *CpSPL* genes by psRNATarget. Prediction results showed that ten *CpSPL* genes were targeted by miR156 (Fig. 1), indicating the miR156-SPL module is universal in plants.

We constructed the phylogenetic tree of 16 *AtSPLs* from *A. thaliana* and 15 *CpSPLs* from *C. pilosula*. 31 *SPL* genes were divided into seven groups and each group had at least one *CpSPL* and one *AtSPL* (Fig. 2). All the non-miR156-targeted *AtSPLs* and *CpSPLs* were grouped into G1, G3, and G7. CpSPL family members in the same group showed similar gene structure and motif composition (Fig. 3; Fig. 4), which was consistent with previous report [17]. In *Arabidopsis*, members in the same group often have the same or similar function. For instance, *AtSPL3*, *AtSPL4*, and *AtSPL5*, clustered in G6, synergistically induce flowering under long-day photoperiod [44]. Most recently, *AtSPL2*, *AtSPL10*, and AtSPL11, members in G4, have been reported to inhibit root regeneration by dampening auxin biosynthesis [43]. We speculate that *CpSPLs* in the same group maybe have the same function, such as *CpSPL2* and *CpSPL10* in G4, *CpSPL4* and *CpSPL5* in G6 and so on.

CpSPL **genes expression patterns in** C. pilosula

Gene expression patterns, to a large extent, will provide valuable information for its potential function [45]. In this study, the spatiotemporal expression patters of 15 *CpSPLs* in the leaves, stems, and roots from one-, two-, and three-month-old seedlings, and the flower and calyx from the plants at the flowering stage were detected by qRT-PCR (Fig. 5). The results showed that *CpSPL1* and *CpSPL14* in G1 exhibited similar expression patterns, and the expression patterns of paralogous *CpSPL9* and *CpSPL15* in G5 showed high similarity. Our results were consistent with previous conclusion that paralogous *SPL* genes in the same group often showed similar expression profiles [46, 47]. *AtSPL9*, *AtSPL10*, and *AtSPL15* contribute to the vegetative to reproductive phase transition [8]. Here, *CpSPL9*, *CpSPL10*, and *CpSPL15* expressed predominantly in the flower, suggesting they might function in the development of flower in *C. pilosula*. The expression level of non-minR156-targeted *CpSPL7* was more constant in all the tissues tested, which was consistent with the result of *SmSPL7* from *Salvia miltiorrhiza* [13].

Some *SPL* genes have been proved to be involved in abiotic stress. For example, in *Arabidopsis*, *AtSPL1* and *AtSPL12* function redundantly in thermotolerance and overexpression of *AtSPL1* or *AtSPL12* increased plant thermotolerance [48]. In alfalfa, silencing *MsSPL13* enhanced tolerance to drought and heat stress (40 °C) [26, 49], and down-regulation of *MsSPL8* led to enhanced salt and drought tolerance [50]. In the present study, we investigated the expression levels of 15 *CpSPLs* under various stress conditions, including NaCl, MeJA, or ABA treatment. We found that the expression levels of most *CpSPL* genes significantly changed under NaCl, MeJA, and ABA treatment (Fig. 6). Among those genes with significant change, *CpSPL4*, *CpSPL6*, *CpSPL14*, and *CpSPL15* positively response to all the treatments, while *CpSPL5*, *CpSPL8*, and *CpSPL10* negatively response to all the treatments. Compared with other genes, *CpSPL5* and *CpSPL8* showed higher fold change under different treatments. We speculate that those two genes are potential candidates involved in abiotic stress.

Functional study of the CpSPL2 and CpSPL10 genes

Since the medicinal and edible part of *C. pilosula* is the root, increasing root yield is one of the main goals of breeding for this species. In *Arabidopsis*, *AtSPL2*, *AtSPL10*, and *AtSPL11*, inhibit root regeneration by dampening auxin biosynthesis [43]. The miR156-targeted *SPL10* is involved in regulating not only lateral root growth but also primary root growth [21, 23]. Recently, it was reported that overexpression of *FvSPL10*, a *SPL* gene from *Fragaria vesca*, resulted in increased organs size, including longer root, larger floral organ and seeds [30]. We speculated that *CpSPL2* and *CpSPL10*, two members clustered in the same group with *AtSPL2/10/11* (Fig. 2), were probably involved in the regulation of root development. To investigate the function of *CpSPL2* and *CpSPL10*, we generated transgenic hairy roots overexpressing *CpSPL2* or *CpSPL10*. Compared with the control, transgenic lines overexpressing *CpSPL2* or *CpSPL10* grew faster and the biomass of *CpSPL2*-OE3, *CpSPL2*-OE5, *CpSPL10*-OE2, and *CpSPL10*-OE3 was 2.19, 1.98, 3.15, and 2.83 times that of the control when the transgenic hairy roots with the length about 1.0 cm were cultured for one month (Fig. 7). Our results indicated that overexpression of *CpSPL2* or *CpSPL10* significantly promote the growth of hairy root.

Furthermore, we determined the concentration of lobetyolin and total saponins in those transgenic lines. Unexpectedly, we found that overexpressing *CpSPL2* or *CpSPL10* dramatically promoted the accumulation of lobetyolin and total saponins in the hairy roots (Fig. 8). Among 16 *AtSPLs* in *Arabidopsis*, *AtSPL9* is the only one that has been reported to regulate biosynthesis of secondary metabolites [24, 25]. *AtSPL9* negatively regulates anthocyanin accumulation by preventing the formation of MBW complex [24], and it positively regulates the formation of (E)-β-caryophyllene by binding to the promoter of sesquiterpene synthase gene *TPS21* and activates its expression [25]. Our results indicated that *CpSPL2* and *CpSPL10* are potential candidates for genetic improvement of *C. pilosula* because they can significantly promote not only the growth of hairy roots, but also accumulation of lobetyolin and total saponins. The molecular mechanism that *CpSPL2* and *CpSPL10* function in hairy roots needs to be addressed in the future.

Conclusions

In this study, we identified 15 *CpSPL* genes, which were supported by confirmation of the SBP domain, based on the genome data of *C. pilosula*. Ten of 15 *CpSPLs* were predicted to be directly targeted by miR156, including *CpSPL3-6, CpSPL9-13*, and *CpSPL15*. All *CpSPLs* were clustered into seven groups and members in the same group share similar gene structure and conserved motif composition. The spatiotemporal expression analysis of 15 *CpSPLs* showed that *CpSPL* gene family had

various expression patterns. The expression levels of most *CpSPL*s significantly changed under NaCl, MeJA, or ABA treatment, and *CpSPL5* and *CpSPL8* showed higher change folds under different treatments. Overexpression of *CpSPL2* or *CpSPL10* significantly promoted not only the growth of hairy roots, but also the accumulation of lobetyolin and total saponins. *CpSPL2* and *CpSPL10* are potential candidates for genetic improvement of *C. pilosula*. These results established a foundation for further investigation of *CpSPLs* and provided novel insights into their biological functions.

Abbreviations

Not applicable.

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Availability of data and materials

All data generated or analysed during this study are included in this published article and its supplementary information files.

Competing Interests

There is no conflict of interest among authors.

Funding

This work was supported by the Major Project of Shaanxi Province, China (Grant No.2017ZDXM-SF-005) and The Youth Innovation Team of Shaanxi Universities. The funders did not design the experiment or draft and revise the manuscript.

Authors' contributions

JY performed the experiments, analyzed the data, and wrote the draft manuscript. ZG prepared the materials and assisted in the experiments. YC, RC and WW assisted in manuscript preparation and helped modify some Figs. XY and XC conceived and coordinated the overall study and revised the manuscript. All authors read and approved the final version of the manuscript.

Acknowledgements

Not applicable.

References

1. Luscombe NM, Austin SE, Bermanand HM, Thornton JM. An overview of the structures of protein-DNA complexes. Genome Biol. 2000;1(1):REVIEWS001.
2. Jin J, Tian F, Yang DC, Meng YQ, Kong L, Luo J, Gao G. PlantTFDB 4.0: toward a central hub for transcription factors and regulatory interactions in plants. Nucleic Acids Res. 2017;45(D1):D1040–D1045.
3. 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.
4. Yamasaki K, Kigawa T, Inoue M, Tateno M, Yamasaki T, Yabuki T, Aoki M, Seki E, Matsuda T, Nunokawa E, Ishizuka Y, Terada T, Shirouzu M, Osanai T, Tanaka A, Seki M, Shinozaki K, Yokoyama S. A novel zinc-binding motif revealed by solution

structures of DNA-binding domains of *Arabidopsis* SBP-family transcription factors. J Mol Biol., 2004;337(1):49–63.

5. 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 Bio. 2005;352(3):585–596.
6. 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):1–11.
7. Preston JC, Hileman LC. Functional evolution in the plant SQUAMOSAPROMOTER BINDING PROTEIN-LIKE (SPL) gene family. Front Plant Sci. 2013; 4:80.
8. Xu M, Hu T, Zhao J, Park MY, Earley KW, Wu G, Yang L, Poethig RS. Developmental functions of miR156-regulated *SQUAMOSA PROMOTER BINDING PROTEIN-LIKE (SPL)* genes in *Arabidopsis thaliana*. PLoS Genet. 2016;2(8):e1006263.
9. 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;142:280–293.
10. 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. Dev Genes Evol. 2017;227(2):101–119.
11. 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):1171–1184.
12. Li J, Hou H, Li X, Xiang J, Yin X, Gao H, Zheng Y, Bassett CL, Wang X. Genome-wide identification and analysis of the SBP-box family genes in apple (*Malus* × *domestica*). Plant Physiol Biochem. 2013;70:100–114.
13. Zhang L, Wu B, Zhao D, Li C, Shao F, Lu S. Genome wide analysis and molecular dissection of the SPL gene family in *Salvia miltiorrhiza*. J Integr Plant Biol. 2014;56:38–50.
14. Hou H, Li J, Gao M, Singer SD, Wang H, Mao L, Fei Z, Wang X. Genomic organization, phylogenetic comparison and differential expression of the SBP-Box family genes in grape. PLoS ONE. 2013; 8: e59358.
15. Pan F, Wang Y, Liu H, Wu M, Chu W, Chen D, Xiang Y. Genome-wide identification and expression analysis of SBP-like transcription factor genes in Moso bamboo (*Phyllostachys edulis*). BMC Genomics. 2017;18(1):486.
16. Zhang HX, Jin JH, He YM, Lu BY, Li DW, Chai WG, Khan A, Gong ZH. Genome-wide identification and analysis of the SBP-box family genes under *Phytophthora capsici* stress in pepper (*Capsicum annuum*). Front Plant Sci. 2016;*7*:504.
17. Zhang SD, Ling LZ. Genome-wide identification and evolutionary analysis of the SBP-box gene family in castor bean. PLoS ONE. 2014;9(1):e86688.
18. Gandikota M, Birkenbihl RP, Höhmann S, Cardon GH, Saedler H, Huijser P. The miRNA156/157 recognition element in the 3' UTR of the *Arabidopsis* SBP box gene *SPL3* prevents early flowering by translational inhibition in seedlings. Plant J. 2010;49(4):683–693.
19. Lal S, Pacis LB, Smith HM. Regulation of the S*QUAMOSA PROMOTER-BINDING PROTEIN-LIKE* genes/microRNA156 module by the homeodomain proteins PENNYWISE and POUND-FOOLISH in *Arabidopsis*. Mol Plant. 2011;4:1123–1132.
20. Hyun Y, Richter R, Vincent C, Martinez-Gallegos R, Porri A, Coupland G. Multi-layered regulation of *SPL15* and cooperation with SOC1 integrate endogenous flowering pathways at the *Arabidopsis* shoot meristem. Dev Cell. 2015;37(3):254–266.
21. Gao R, Wang Y, Gruber MY, Hannoufa A. miR156/SPL10 modulates lateral root development, branching and leaf morphology in *Arabidopsis* by silencing AGAMOUS-LIKE 79. Front Plant Sci. 2017;8:2226.
22. Ye BB, Zhang K, Wang JW. The role of miR156 in rejuvenation in *Arabidopsis thaliana*. J Integr Plant Biol. 2019;62(5):550–555.
23. Barrera-Rojas CH, Rocha GH, Polverari L, Pinheiro Brito DA, Batista DS, Notini MM, Cruz AC, Morea EG, Sabatini S, Otoni WC, Nogueira FT. miR156-targeted *SPL10* controls root meristem activity and root-derived *de novo* shoot regeneration via cytokinin responses. J Exp Bot. 2020;71(3):934–950.
24. Gou JY, Felippes FF, Liu CJ, Weigel D, Wang JW. Negative regulation of anthocyanin biosynthesis in *Arabidopsis* by a miR156-targeted SPL transcription factor. Plant Cell. 2011;23(4):1512–1522.
25. Yu ZX, Wang LJ, Zhao B, Shan CM, Zhang YH, Chen DF, Chen XY. Progressive regulation of sesquiterpene biosynthesis in *Arabidopsis* and Patchouli (*Pogostemon cablin*) by the miR156-targeted SPL transcription factors. Mol Plant. 2015;8:98–

110.

26. Feyissa BA, Arshad M, Gruber MY, Kohalmi SE, Hannoufa A. The interplay between miR156/SPL13 and DFR/WD40-1 regulate drought tolerance in alfalfa. BMC Plant Biol. 2019;19:434.
27. Mermod M, Takusagawa M, Kurata T, Kamiya T, Fujiwara T, Shikanai T. *SQUAMOSA promoter-binding protein-like 7* mediates copper deficiency response in the presence of high nitrogen in *Arabidopsis thaliana*. Plant Cell Rep. 2019;38: 835–
28. Miura K, Ikeda M, Matsubara A, Song XJ, Ito M, Asano K, Matsuoka M, Kitano H, Ashikari M. *OsSPL14* promotes panicle branching and higher grain productivity in rice. Genet. 2010;42:545–549.
29. Wang SK, Wu K, Yuan QB, Liu XY, Liu ZB, Lin XY, Zeng RZ, Zhu HT, Dong GJ, Qian Q, Zhang G, Fu X. Control of grain size, shape and quality by *OsSPL16* in rice. Genet. 2012;44:950–954.
30. Xiong J, Bai Y, Ma C, Zhu H, Zheng D, Cheng Z. Molecular cloning and characterization of *SQUAMOSA*-promoter binding protein-like gene *FvSPL10* from Woodland Strawberry (*Fragaria vesca*). Plants (Basel). 2019;8(9):342.
31. Committee for the Pharmacopoeia of the People's Republic of China, Pharmacopoeia of the People's Republic of China, Part I. China Medical Science Press: Beijing; 2020;293-294.
32. Li FJ, Wang ZC, Yang K. Summary of recent research on *Codonopsis pilosula*. Technol Inf. 2008;35:422–440.
33. Li D, Li ZL. The research status of that *Codonopsis pilosula* polysaccharide is as an immune adjuvant. Guide of China Medicine. 2013;11:56–57.
34. He JY, Ma N, Zhu S, Komatsu K, Li ZY, Fu WM. The genus *Codonopsis* (Campanulaceae): a review of phytochemistry, bioactivity and quality control. J Nat Med. 2015;69(1):1–
35. Yoon IS, Cho SS. Effects of lobetyolin on xanthine oxidase activity in vitro and in vivo: weak and mixed inhibition. Nat Prod Res. 2019;29:1–
36. He W, Tao W, Zhang F, Jie Q, He Y, Zhu W, Tan J, Shen W, Li L, Yang Y, Cheng H, Sun D. Lobetyolin induces apoptosis of colon cancer cells by inhibiting glutamine metabolism. J Cell Mol Med. 2020;24(6):3359–
37. Gao JP, Wang D, Cao LY, Sun HF. Transcriptome sequencing of *Codonopsis pilosula* and identification of candidate genes involved in polysaccharide biosynthesis. PLoS ONE. 2015;10:e0117342.
38. Ji JJ, Feng Q, Sun HF, Zhang XJ, Li XX, Li JK, Gao JP. Response of bioactive metabolite and biosynthesis related genes to methyl jasmonate elicitation in *Codonopsis pilosula*. Molecules. 2019;24:533.
39. Yang J, Yang X, Li B, Lu X, Kang J, Cao X. Establishment of in vitro culture system of *Codonopsis pilosula* transgenic hairy root. 3 Biotech. 2020;10:137.
40. Yang J, Yang XZ, Kuang Z, Li B, Lu XY, Cao XY, Kang JF. Selection of suitable reference genes for qRT-PCR expression analysis of *Codonopsis pilosula* under different experimental conditions. Molecular biology reports. 2020;47(6): 4169–
41. Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-\Delta\Delta CT}$ Methods. 2001;25:402–408.
42. Axtell MJ, Bowman JL. Evolution of plant microRNAs and their targets. Trends Plant Sci. 2008;13:343–349.
43. Ye BB, Shang GD, Pan Y, Xu ZG, Zhou CM, Mao YB, Bao N, Sun L, Xu T, Wang JW. AP2/ERF transcription factors integrate age and wound signals for root regeneration. Plant Cell. 2020;32:226–
44. Jung JH, Lee HJ, Ryu JY, Park CM. SPL3/4/5 integrate developmental aging and photoperiodic signals into the FT-FD module in *Arabidopsis* Mol Plant. 2016;9(12):1647–1659.
45. Zhou Q, Zhang S, Chen F, Liu B, Wu L, Li F, Zhang J, Bao M, Liu G. Genome-wide identification and characterization of the SBP-box gene family in *Petunia*. BMC Genomics. 2018;19(2):193–211.
46. Li C, Lu S. Molecular characterization of the SPL gene family in *Populus trichocarpa*. BMC Plant Biol. 2014;14:131.
47. Zeng RF, Zhou JJ, Liu SR, Gan ZM, Zhang JZ, Hu CG. Genome-wide identification and characterization of *SQUAMOSA*-promoter-binding protein (SBP) genes involved in the flowering development of *Citrus Clementina*. Biomolecules. 2019;9(2):66.

48. Chao LM, Liu YQ, Chen DY, Xue XY, Mao YB, Chen XY. *Arabidopsis* transcription factors *SPL1* and *SPL12* confer plant thermotolerance at reproductive stage. Mol Plant. 2017;10:735–
49. Matthews C, Arshad M, Hannoufa A. Alfalfa response to heat stress is modulated by microRNA156. Physiol Plant. 2019;165(4):830–
50. Gou J, Debnath S, Sun L, Flanagan A, Tang Y, Jiang Q, Wen J, Wang ZY. From model to crop: functional characterization of *SPL 8* in *truncatula* led to genetic improvement of biomass yield and abiotic stress tolerance in alfalfa. Plant Biotechnol J. 2018;16(4):951–962.

Appendix A. Supplemental Files

The supplemental tables were not provided with this version of the manuscript.

Table S1. Primer sequences used in the study

Table S2. cDNA sequences of *15 SPLs* in *Codonopsis pilosula*

Table S3. The conserved motifs of SPL in *Codonopsis pilosula*

Table S4. Putative cis-acting elements present in *SPL* promoters of *Codonopsis pilosula*

Figures

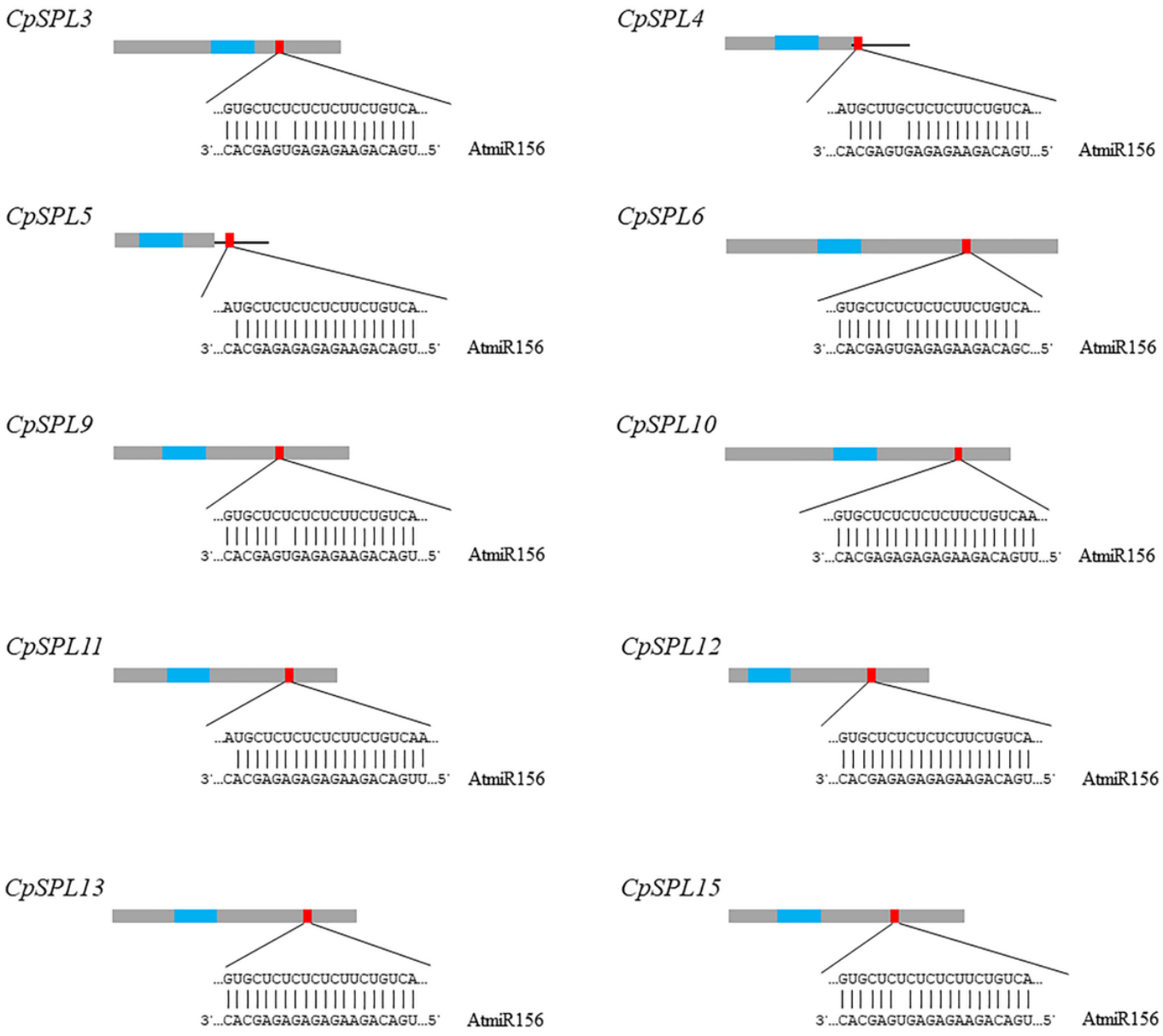

Figure 1

Binding sites of AtmiR156 in CpSPL genes. The CpSPLs regulated by miR156. The gray boxes represent the CDS of CpSPL genes. The blue boxes represent the conserved SBP domain. The red boxes represent the miR156 target site.

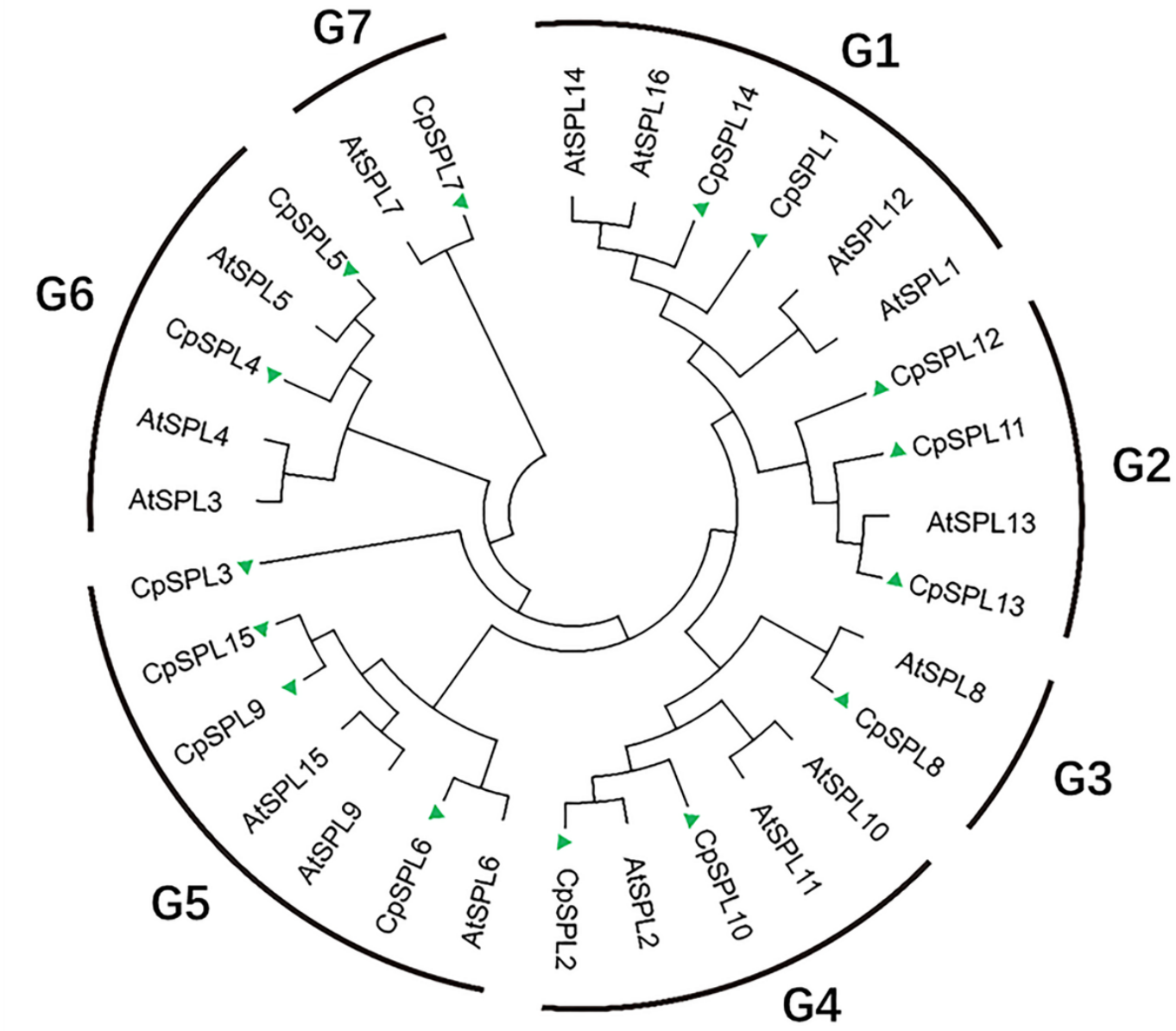

Figure 2

An ML phylogenetic tree of the SPLs from Codonopsis pilosula and Arabidopsis thaliana.

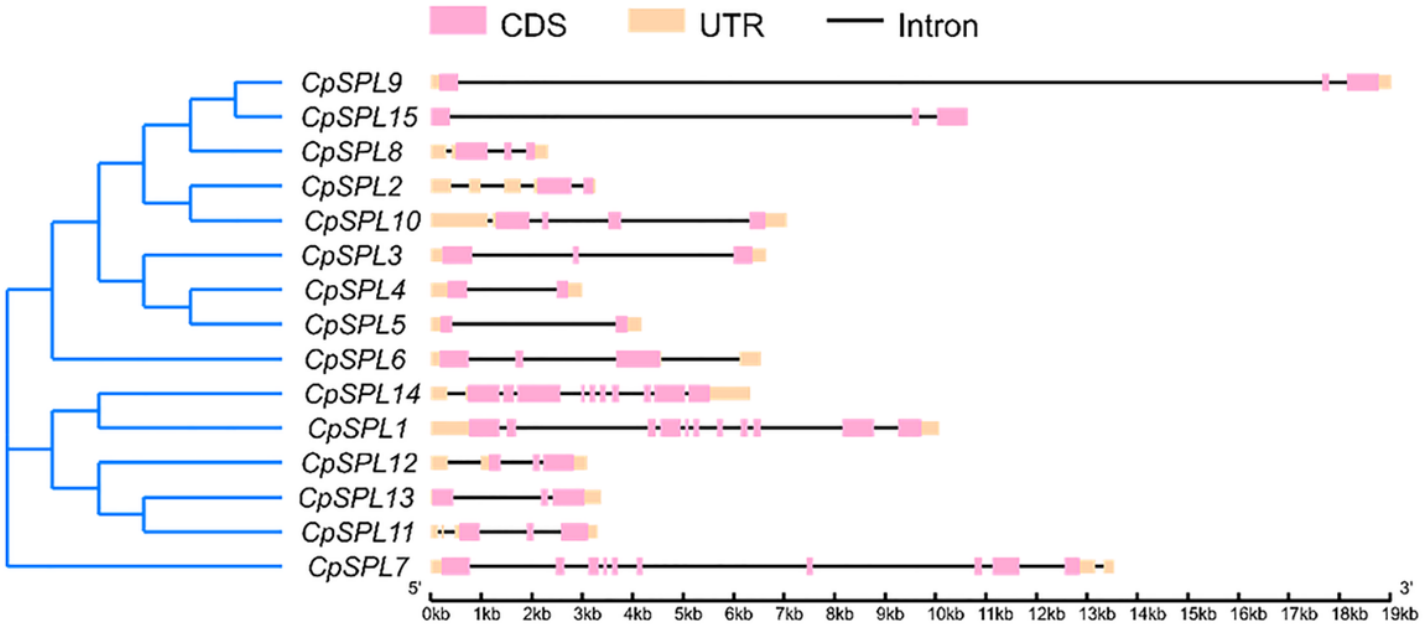

Figure 3

Exon-intron organization structures of 15 SPL genes in Codonopsis pilosula. Exons are represented by pink rectangles, introns are represented by black lines, UTRs are represented by orange rectangles.

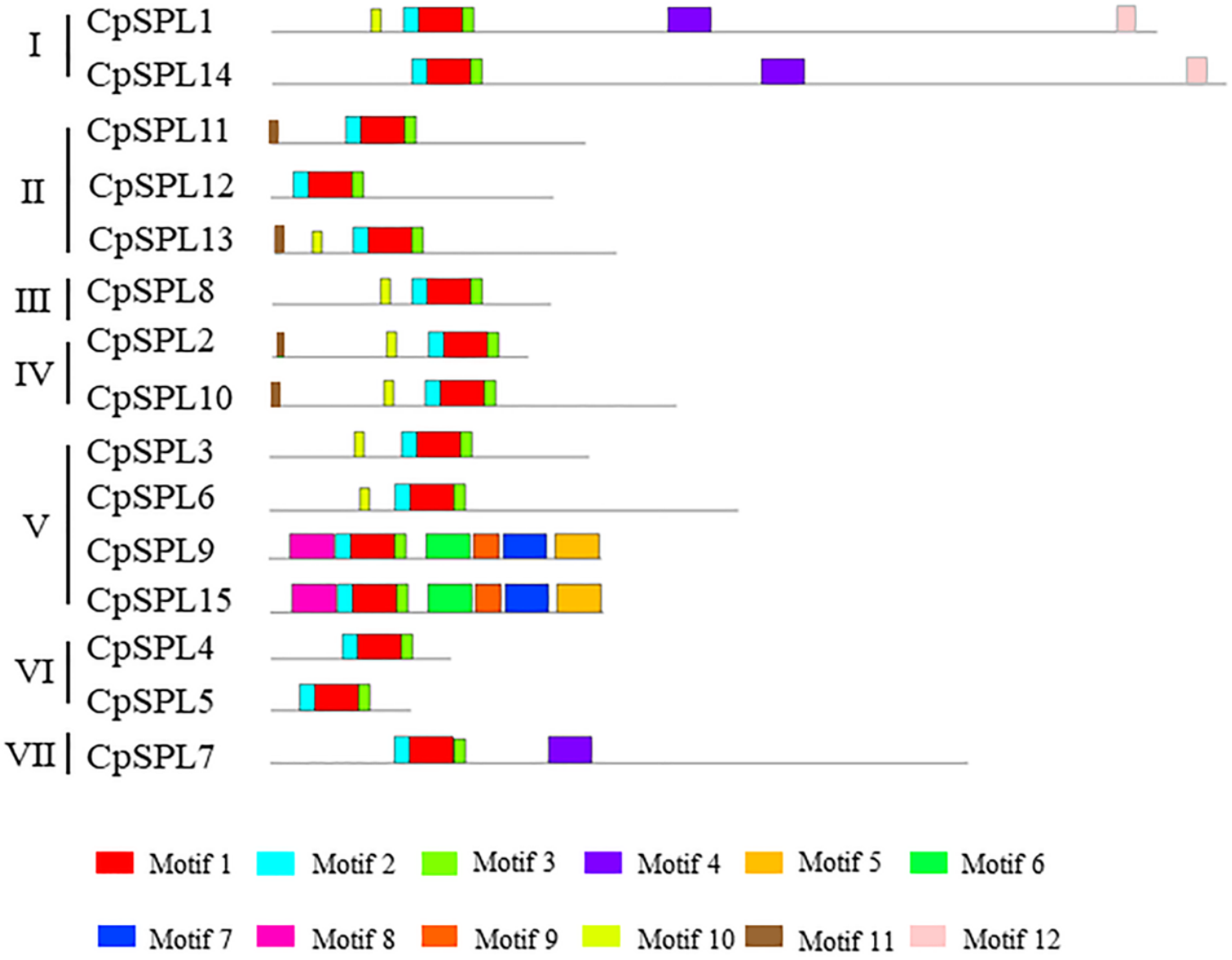

Figure 4

The analysis of conserved motifs of SPLs in Codonopsis pilosula.

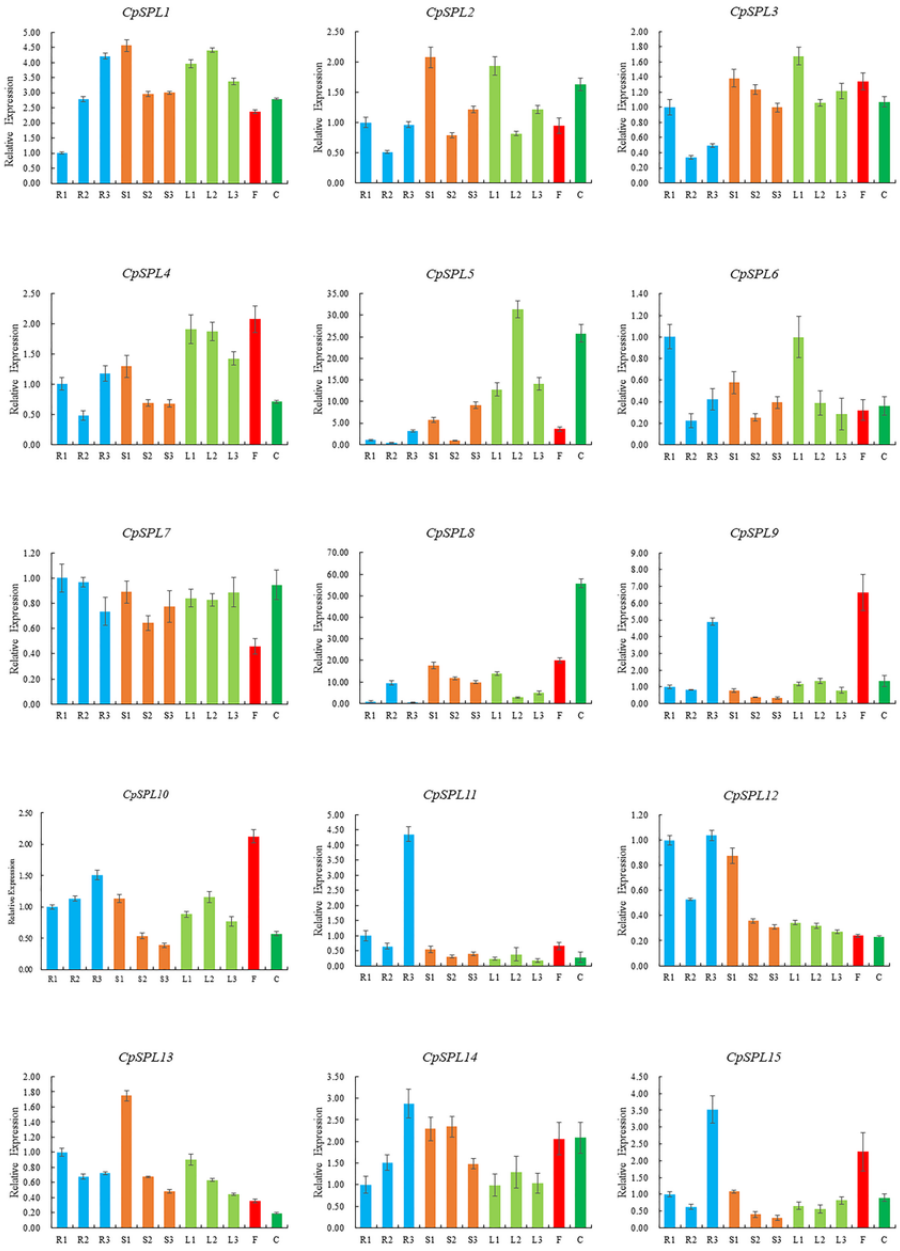

Figure 5

Spatiotemporal expression analysis of 15 SPL genes in Codonopsis pilosula. R1, R2, and R3 represent roots from one-, two-, and three-month-old seedlings, respectively; S1, S2, and S3 represent stems from one-, two-, and three-month-old seedlings, respectively; L1, L2, and L3 represent leaves from one-, two-, and three-month-old seedlings, respectively; F: flower; C: calyx.

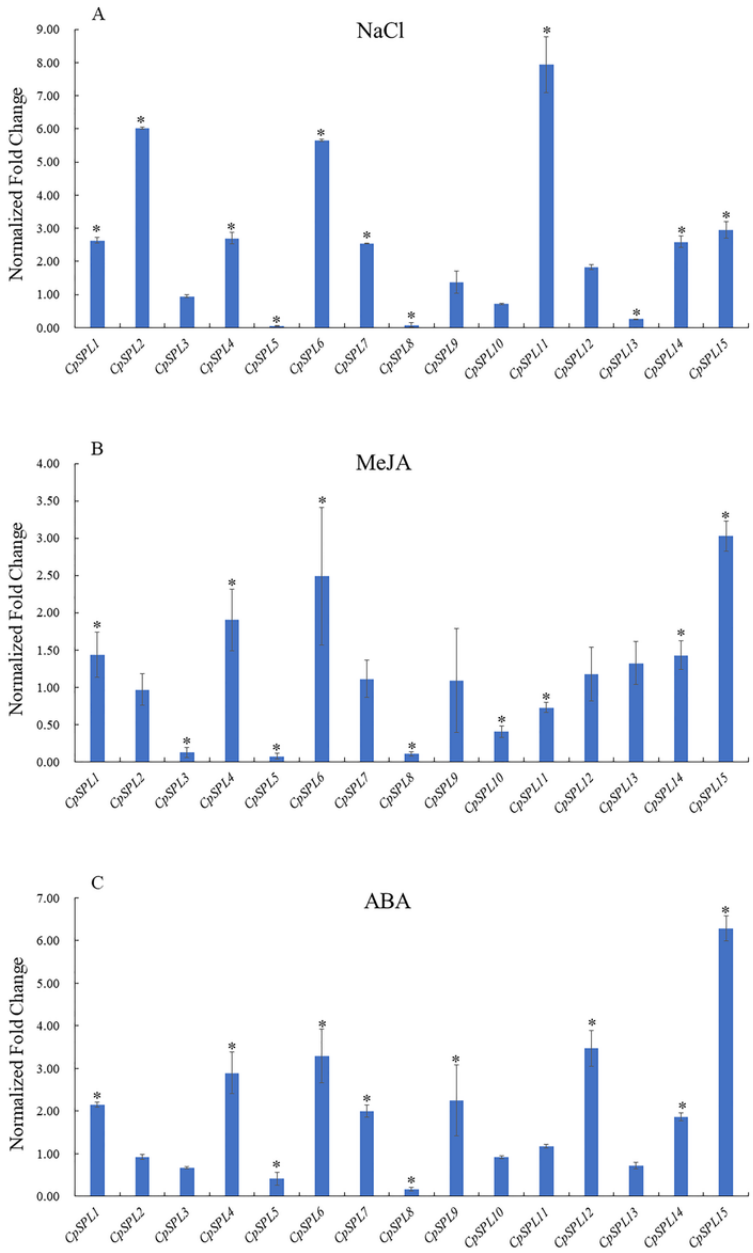

Figure 6

Relative expression levels of 15 SPL genes in Codonopsis pilosula under NaCl, methyl jasmonate (MeJA), and abscisic acid (ABA) treatment. Change multiples are relative to the expression level of each gene treated with ddH2O (set to 1). * means the value is significantly different from the control at $p<0.05$.

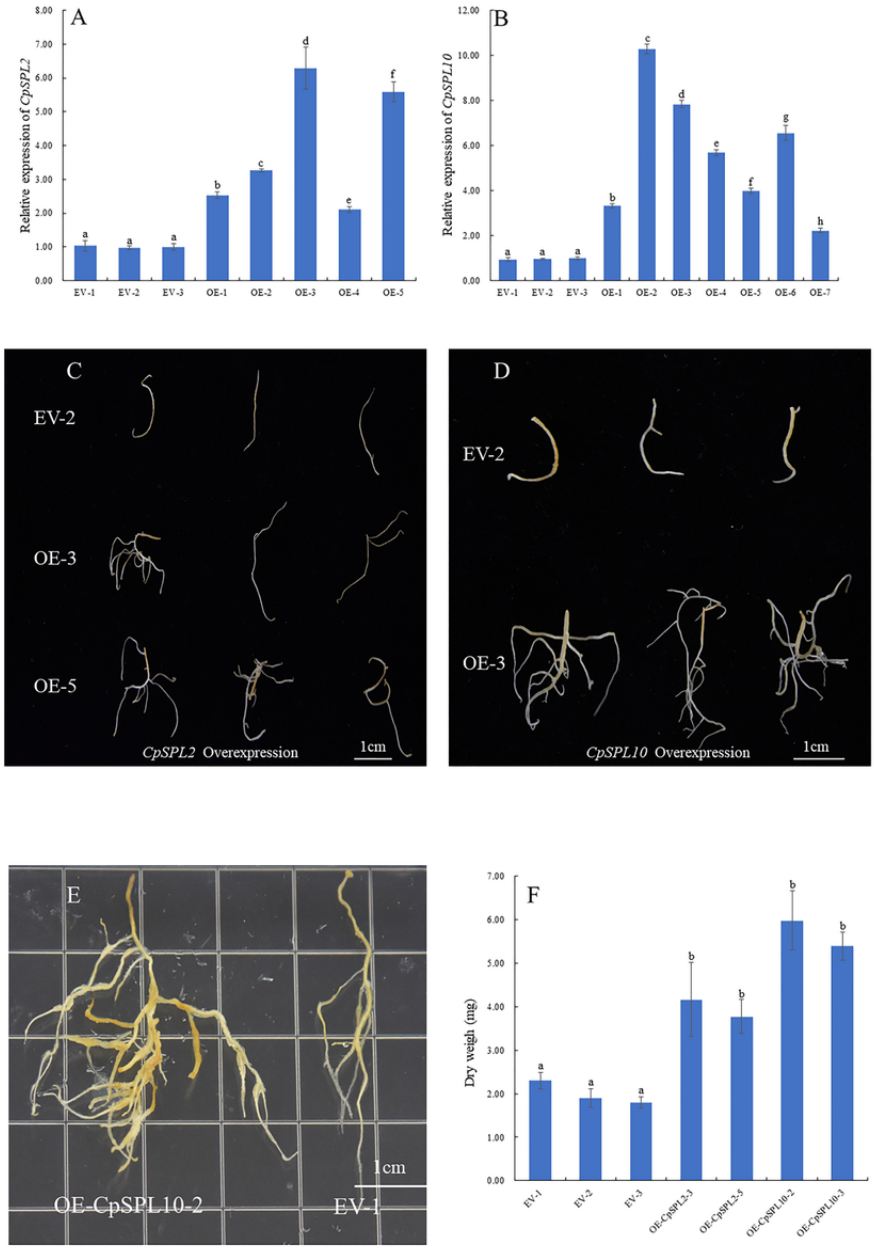

Figure 7

Identification and growth phenotype of CpSPL2 or CpSPL10-overexpressing (OE) Codonopsis pilosula transgenic hairy roots. (A) Expression levels of CpSPL2 in CpSPL2-OE lines. (B) Expression levels of CpSPL10 in CpSPL10-OE lines. (C) Growth phenotype of empty vector (EV) and CpSPL2-OE lines when 1 cm hairy roots were cultured for one month. (D) Growth phenotype of EV and CpSPL10-OE lines when 1 cm hairy roots were cultured for one month. (E) Growth phenotype of EV and CpSPL10-OE lines when 1 cm hairy roots were cultured for two months. (F) The biomass of EV, CpSPL2-OE and CpSPL10-OE lines when 1 cm hairy roots were cultured for one months. One-way ANOVA (followed by Tukey's comparisons) tested for significant differences among means (indicated by different letters at $p<0.05$).

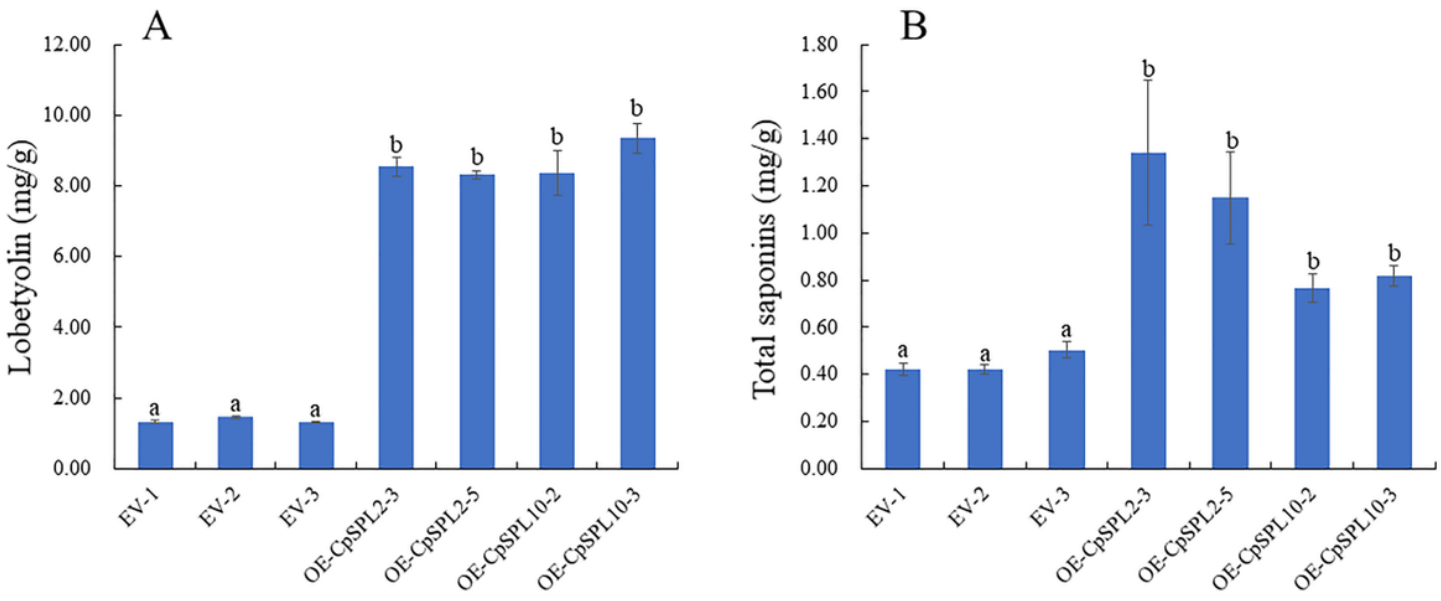

Figure 8

Concentration of lobetyolin and total saponins contents in Codonopsis pilosula transgenic hairy roots. One-way ANOVA (followed by Tukey's comparisons) tested for significant differences among means (indicated by different letters at $p<0.05$).