

RESEARCH ARTICLE

Open Access

Molecular characterization of the *SPL* gene family in *Populus trichocarpa*

Caili Li and Shanfa Lu*

Abstract

Background: SPLs, a family of transcription factors specific to plants, play vital roles in plant growth and development through regulation of various physiological and biochemical processes. Although *Populus trichocarpa* is a model forest tree, the *PtSPL* gene family has not been systematically studied.

Results: Here we report the identification of 28 full-length *PtSPLs*, which distribute on 14 *P. trichocarpa* chromosomes. Based on the phylogenetic relationships of SPLs in *P. trichocarpa* and *Arabidopsis*, plant SPLs can be classified into 6 groups. Each group contains at least a *PtSPL* and an *AtSPL*. The N-terminal zinc finger 1 (Zn1) of SBP domain in group 6 SPLs has four cysteine residues (CCCC-type), while Zn1 of SPLs in the other groups mainly contains three cysteine and one histidine residues (C2HC-type). Comparative analyses of gene structures, conserved motifs and expression patterns of *PtSPLs* and *AtSPLs* revealed the conservation of plant SPLs within a group, whereas among groups, the *P. trichocarpa* and *Arabidopsis* SPLs were significantly different. Various conserved motifs were identified in *PtSPLs* but not found in *AtSPLs*, suggesting the diversity of plant SPLs. A total of 11 pairs of intrachromosome-duplicated *PtSPLs* were identified, suggesting the importance of gene duplication in *SPL* gene expansion in *P. trichocarpa*. In addition, 18 of the 28 *PtSPLs*, belonging to G1, G2 and G5, were found to be targets of miR156. Consistently, all of the *AtSPLs* in these groups are regulated by miR156. It suggests the conservation of miR156-mediated posttranscriptional regulation in plants.

Conclusions: A total of 28 full-length SPLs were identified from the whole genome sequence of *P. trichocarpa*. Through comprehensive analyses of gene structures, phylogenetic relationships, chromosomal locations, conserved motifs, expression patterns and miR156-mediated posttranscriptional regulation, the *PtSPL* gene family was characterized. Our results provide useful information for evolution and biological function of plant SPLs.

Background

SPL proteins constitute a diverse family of transcription factors playing vital roles in plant growth and development. SPLs are specific to plants and have a highly conserved SBP (SQUAMOSA PROMOTER BINDING PROTEIN) domain with approximately 78 amino acid residues. The domain contains three functionally important motifs, including zinc finger 1 (Zn1), zinc finger 2 (Zn2), and nuclear location signal (NLS) [1,2]. Genes encoding SPLs were first identified for SBP1 and SBP2 in *Antirrhinum majus* [3]. Lately, it has been found in various green plants, including single-celled green algae, mosses, gymnosperms, and angiosperms. The results showed that *SPLs* existed as a large gene family in plants.

For instance, the *SPL* gene family in *Arabidopsis*, rice, *Physcomitrella patens*, maize and tomato includes 16, 19, 13, 31 and 15 members, respectively [4-9].

The 16 *Arabidopsis SPLs* are termed as *AtSPL1* to *AtSPL16* [2], respectively, of which *AtSPL1*, *AtSPL7*, *AtSPL12*, *AtSPL14* and *AtSPL16* are relatively large and expressed constitutively, while the others are relatively small and highly expressed in flowers [4,10]. Ten of the 16 *AtSPLs*, including *AtSPL2-AtSPL6*, *AtSPL9-AtSPL11*, *AtSPL13* and *AtSPL15*, are regulated by miRNAs belonging to the *MIR156* family [11-17]. *AtSPL3*, *AtSPL4* and *AtSPL5* contain complementary sequences of miR156 in 3' UTR, and all of them promote vegetative phase change and flowering [10,14,18]. *AtSPL2*, *AtSPL10* and *AtSPL11* regulate morphological traits of cauline leaves and flowers [19]. Overexpression of miR156b reduces the accumulation of *AtSPL2*, *AtSPL10* and *AtSPL11* mRNA [12,14,20]. *AtSPL9* and *AtSPL15* act redundantly

* Correspondence: sflu@implad.ac.cn
Institute of Medicinal Plant Development, Chinese Academy of Medical Sciences & Peking Union Medical College, No.151, Malianwa North Road, Haidian District, Beijing 100193, China

in controlling the juvenile-to-adult growth phase transition and leaf initiation rate in *Arabidopsis* [21]. Six *AtSPLs*, including *AtSPL1*, *AtSPL7*, *AtSPL8*, *AtSPL12*, *AtSPL14* and *AtSPL16*, are not targets of miR156 in *Arabidopsis*. Among them, *AtSPL7* can bind directly to the Cu-response element (CuRE) containing a core sequence of GTAC and is a regulator of Cu homeostasis in *Arabidopsis* [22]. *AtSPL8* regulates pollen sac development [23], male fertility [24], GA biosynthesis and signaling [25]. *AtSPL14* plays significant roles in plant development and sensitivity to fumonisin B1 [26]. Among the 19 rice *SPLs*, half are predominantly expressed in various young organs [27]. OsSPLs targeted by miR156 are involved in the development of flowers in rice. *OsSPL14* regulated by miR156 also controls shoot branching in the vegetative stage [8,28,29]. In maize, *liguleless1* containing the SBP domain regulates ligule and auricle formation [30,31].

Populus trichocarpa is a model plant with whole genome sequence available [32]. A total of 352 miRNA precursors, including 12 for miR156, have been identified [33-39]. However, the regulation of miR156 in *P. trichocarpa* *PtSPLs* has not been analyzed. In our previous studies [40], 17 *PtSPLs*, which appeared to be full-length or partial sequence with at least 300 amino acids, were identified from the *Populus* genome assembly v1.1 (http://genome.jgi-psf.org/Poptr1_1/Poptr1_1.home.html). They were named *PtSPL1*–*PtSPL17*, respectively, of which *PtSPL3* and *PtSPL4* had the highest similarities with *AtSPL7* involved in Cu homeostasis [40]. In order to characterize the whole *SPL* gene family in *P. trichocarpa*, we searched the *Populus* genome assembly v1.1, v2.2 and v3.0 [32]. It resulted in the identification of 28 full-length *PtSPLs*. Gene structures, chromosomal locations, phylogenetic relationships, conserved protein motifs and expression patterns of all identified *PtSPLs* were systematically analyzed. MiR156-mediated posttranscriptional regulation of *PtSPL* genes was investigated. The results provide useful information for elucidating the biological functions of *SPLs* in *P. trichocarpa*.

Results

Identification of 28 *SPL* genes in *P. trichocarpa* genome

Analysis of the *Populus* genome assembly v1.1, v2.2 and v3.0 showed the existence of 28 full-length *SPL* genes in the *P. trichocarpa* genome (Table 1). All of the deduced *PtSPL* proteins contained the conserved SBP domain. The theoretical *pI* of deduced *PtSPL* proteins ranged from 5.87 to 9.49. The length varied between 148 and 1044 amino acids. The molecular weight (Mw) varied from 16.2 to 116.1 kDa (Additional file 1). The distribution of *pI* is similar to *AtSPLs* (Additional file 2); however, the length and Mw of *PtSPLs* are larger than *AtSPLs*.

Mapping *PtSPLs* to the *P. trichocarpa* genome showed that 28 *PtSPLs* were unevenly distributed on 14 chromosomes with four on Chr2, 3 on each of Chr1, Chr8, Chr10 and Chr14, 2 on each of Chr3, Chr11 and Chr15, and one on each of Chr4, Chr5, Chr7, Chr12, Chr16 and Chr18 (Figure 1). Relatively high densities of *PtSPLs* were observed in the top and bottom regions of Chr8, Chr10, Chr11 and Chr14, the top of Chr1, Chr4, and Chr16, and the bottom of Chr3, Chr5, Chr7, Chr12 and Chr18. Few are in the central regions of chromosomes. Moreover, 11 pair of *PtSPLs* ($K_s < 1.0$) were evolved from intrachromosomal duplication (Table 2), indicating the importance of gene duplication for *PtSPL* gene expansion.

Phylogenetic analysis of *SPLs* in *P. trichocarpa* and *Arabidopsis*

In order to investigate the evolutionary relationship between *P. trichocarpa* and *A. thaliana* *SPL* proteins, a neighbor-joining (NJ) phylogenetic tree was constructed for 28 *PtSPLs* and 16 *AtSPLs* using MEGA5.1. The reliability of branching was assessed by the bootstrap resampling method using 1,000 bootstrap replicates. Only nodes supported by bootstrap values >50% are used for further analysis. The results showed that the 44 *SPL* proteins clustered into 6 groups (named G1–G6), each of which contained at least one *AtSPL* and one *PtSPL* (Figure 2). It is consistent with the results from *SmSPLs* in *Salvia miltiorrhiza* [41]. To further confirm that there are 6 groups of *SPLs*, we also constructed a phylogenetic tree for 28 *PtSPLs*, 16 *AtSPLs*, 18 rice *OsSPLs* and 15 *SmSPLs*. As shown in Additional file 3, the 77 *SPLs* also clustered into 6 groups. The difference between the two trees constructed (Figure 2, Additional file 3) is that *PtSPL12*, *PtSPL13*, *PtSPL28* and *AtSPL6* belonging to G1 in Figure 2 are included in G2 in Additional file 3. An intron was found in the SBP domain-encoding region of all *SPL* genes from *P. trichocarpa* and *Arabidopsis* (Figure 3); however, sequence feature analysis showed that the SBP domain of *SPLs* in G6 (*AtSPL7*, *PtSPL3* and *PtSPL4*) were divergent from the other groups. The N-terminal zinc finger of G6 *SPLs* has four cysteine residues in the SBP domain, while *SPLs* in the other groups mainly contain three cysteines and one histidine, indicating the diversification of plant *SPL* evolution. On the other hand, *SPLs* within a group have similar intron number, exon-intron structure, and coding sequence length. Consistently, the length, Mw and theoretical *pI* of deduced *SPL* proteins within a group are also similar, although they are divergent among groups. It suggests the conservation of plant *SPLs* in a group. Phylogenetic analysis showed that *PtSPL3* and *PtSPL4* had high homology with *AtSPL7*, an *Arabidopsis* *SPL* with the capability of binding CuREs in the *MIR398* promoter *in vitro*

Table 1 *PtSPL* gene names and gene model IDs in the *Populus* genome assembly v1.1, v2.2 and v3.0

Gene name	Gene ID		
	V1.1	V2.2	V3.0
<i>PtSPL1</i>	GW1.X.791.1	POPTR_0010s16370	Potri.010G154000
<i>PtSPL2</i>	FGENESH4_PM.C_LG_II000008	POPTR_0002s00440	Potri.002G002400
<i>PtSPL3</i>	ESTEXT_FGENESH4_PM.C_LG_X0096	POPTR_0010s02710	Potri.010G026200
<i>PtSPL4</i>	ESTEXT_FGENESH4_PM.C_LG_VIII0830	POPTR_0008s20160	Potri.008G197000
<i>PtSPL5</i> ^a	GRAIL3.0010027501 ^b + GRAIL3.0010027301 ^b + GRAIL3.0010027401 ^b	POPTR_0008s09810	Potri.008G098600
<i>PtSPL6</i> ^a	ESTEXT_GENEWISE1_V1.C_LG_XIV2145 ^b + GW1.XIV.2149.1 ^b	POPTR_0014s10960	Potri.014G114300
<i>PtSPL7</i> ^a	GRAIL3.0050015101 ^b + GW1.8978.5.1 ^b + GW1.II.489.1 ^b	POPTR_0002s18970	Potri.002G188700
<i>PtSPL8</i>	FGENESH4_PG.C_LG_II001303	POPTR_0002s14330	Potri.002G142400
<i>PtSPL9</i> ^a	EUGENE3.00051637 ^b + EUGENE3.00051638 ^b	POPTR_0005s28010	Potri.005G258700
<i>PtSPL11</i>	GRAIL3.0047015901 ^b	POPTR_0003s17120	Potri.003G172600 ^b
<i>PtSPL12</i>	GRAIL3.0010026901	POPTR_0008s09750	Potri.008G097900
<i>PtSPL13</i>	FGENESH4_PG.C_LG_X001404	POPTR_0010s16400	Potri.010G154300
<i>PtSPL14</i>	ESTEXT_GENEWISE1_V1.C_LG_XV2187	POPTR_0015s11100	Potri.015G098900
<i>PtSPL15</i>	EUGENE3.00120942	POPTR_0012s10260	Potri.012G100700
<i>PtSPL16</i>	ESTEXT_GENEWISE1_V1.C_1240186	POPTR_0011s05480	Potri.011G055900
<i>PtSPL17</i> ^a	EUGENE3.00160416	POPTR_0016s04880 ^b + POPTR_0016s04890 ^b	Potri.016G048500 ^c
<i>PtSPL18</i>	GW1.I.7783.1 ^b	POPTR_0001s13630	Potri.001G058600
<i>PtSPL19</i>	GW1.I.7690.1 ^b	POPTR_0001s13890	Potri.001G055900
<i>PtSPL20</i>	GW1.107.39.1 ^b	POPTR_0001s40870	Potri.001G398200
<i>PtSPL21</i>	GW1.II.3778.1 ^b	POPTR_0002s14320	Potri.002G142200
<i>PtSPL22</i>	GW1.III.2396.1 ^b	POPTR_0003s16780	Potri.003G169400
<i>PtSPL23</i>	GW1.IV.3037.1 ^b	POPTR_0004s04630	Potri.004G046700
<i>PtSPL24</i>	GW1.VII.548.1 ^b	POPTR_0007s01030	Potri.007G138800
<i>PtSPL25</i>	GW1.XI.3794.1 ^b	POPTR_0011s11770	Potri.011G116800 ^d
<i>PtSPL26</i>	GW1.40.81.1 ^b	POPTR_0014s05680	Potri.014G057700
<i>PtSPL27</i>	GW1.40.76.1 ^b	POPTR_0014s05690	Potri.014G057800
<i>PtSPL28</i>	GW1.129.152.1 ^b	POPTR_0015s07140	Potri.015G060400
<i>PtSPL29</i>	GW1.164.76.1 ^b	POPTR_0018s14680	Potri.018G149900

^aGenes are split into 2 or 3 gene models in v1.1 or v2.2;

^bGene models with partial sequence;

^cThe gene model includes 7 additional amino acids at the N-terminal compared with the model in v1.1.

^dThe gene model includes 49 additional amino acids at the N-terminal compared with the model in v2.2.

and involved in response to copper deficiency in *Arabidopsis* [22]. It is consistent with our previous results for *PtSPLs* [40]. Based on the phylogenetic tree, *PtSPL3* and *AtSPL7* are very likely to be orthologous proteins (Figure 2). Additionally, 5 pairs of *AtSPLs* and 11 pairs of *PtSPLs* seem to be paralogous proteins (Figure 2). It includes *AtSPL9/15*, *AtSPL10/11*, *PtSPL8/27*, *PtSPL12/13* and *PtSPL11/19* belonging to G1, *PtSPL18/22* and *PtSPL14/15* from G2, *PtSPL21/26* belonging to G3, *AtSPL14/16*, *AtSPL1/12*, *PtSPL2/9*, *PtSPL1/5* and *PtSPL6/7* included in G4, and *AtSPL3/4*, *PtSPL16/23* and *PtSPL20/25* clustering in G5. About 62.5% of the 16 *AtSPLs* and

78.5% of the 28 *PtSPLs* exist as paralogous pairs. It suggested that the expansion of *SPL* genes occurred after separation of paralogous genes. The results from paralogous pair identification were consistent with segmental duplications in the *P. trichocarpa* genome (<http://chibba.agtec.uga.edu/duplication/>) [32], suggesting the origination of paralogous *PtSPLs* from segmental duplication. Prediction of potential age of tandem duplication events using synonymous substitutions (Ks) values showed that the segmental duplication events for *PtSPLs* appeared to occur in 9–21 mya (Table 2). It is consistent with the age of *P. trichocarpa* genome duplication events [32].

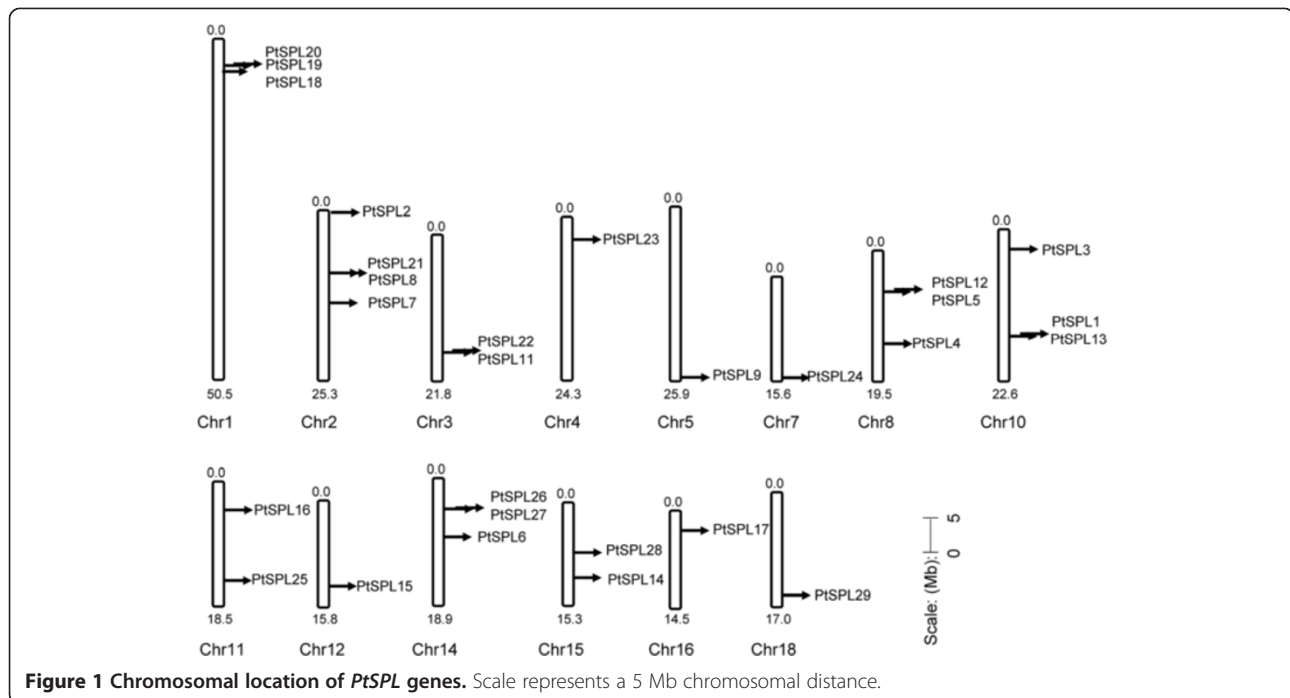


Figure 1 Chromosomal location of *PtSPL* genes. Scale represents a 5 Mb chromosomal distance.

Comparative analysis of *PtSPL* and *AtSPL* gene structures

Gene structure analysis showed that the number of introns in the coding region of 28 *PtSPL* genes varied from 1 to 10. The number of *PtSPLs* with 1, 2, 3, 4, 9 and 10 introns is five, ten, four, one, six, and two, respectively (Figure 3, Additional file 1). Similarly, the intron number of 16 *AtSPLs* varies between 1 and 9 (Additional file 2). The pattern of intron distribution in *PtSPLs* is quite similar to *AtSPLs* with the majority to be 2 and 9 introns, followed by 1 and 3 (Figure 3, Additional files 1 and 2) [41]. In addition, the position of intron in the SBP domain is highly conserved. It locates in the codon for the 48th amino acid of SBP domain (Additional

file 4). These results suggest the conservation of exon-intron structures between *PtSPLs* and *AtSPLs*.

The length of introns varies significantly among *PtSPL* genes, such as those in G1, G2 and G5 (Figure 3). We analyzed the internal exons and introns of *PtSPLs* and *AtSPLs*. The results showed that the exons of *PtSPLs* had a size from 43 to 884 bp with an average of 314 bp, which is slightly greater than 293 bp of the average length of *AtSPL* exons. Approximately 59% of *PtSPL* exons and 63% of *AtSPL* exons have a size below 300 bp and 71% and 70% of exons are between 60 and 160 bp in *PtSPLs* and *AtSPLs*, respectively (Figure 4). Although the size distribution of *PtSPLs* exons is similar to *AtSPL* exons, intron size distribution is more variable, ranging from 30 bp to 3.0 kb. There are 6 *PtSPL* introns (5%) with sizes >1.5 kb; however, no such introns exist in *AtSPLs*. About 55% of *PtSPLs* have sizes below 300 bp and 56% of introns are between 60 and 160 bp; however, the majority of *AtSPLs* (94%) have sizes below 300 bp. The average size of *PtSPL* introns is 476 bp, which is much greater than 120 bp of *AtSPLs*. These results suggest the difference of exon and intron size distribution between *PtSPLs* and *AtSPLs*.

Table 2 Estimated age of the duplication events for *PtSPL* paralogous genes

Paralogous genes	Ks	Estimated time (mya)
<i>PtSPL19</i> (Chr1)/ <i>PtSPL11</i> (Chr3)	0.16	9
<i>PtSPL21</i> (Chr2)/ <i>PtSPL26</i> (Chr14)	0.31	17
<i>PtSPL8</i> (Chr2)/ <i>PtSPL27</i> (Chr14)	0.29	16
<i>PtSPL2</i> (Chr2)/ <i>PtSPL9</i> (Chr5)	0.27	15
<i>PtSPL12</i> (Chr8)/ <i>PtSPL13</i> (Chr10)	0.24	13
<i>PtSPL5</i> (Chr8) / <i>PtSPL1</i> (Chr10)	0.22	12
<i>PtSPL18</i> (Chr1)/ <i>PtSPL22</i> (Chr3)	0.17	9
<i>PtSPL23</i> (Chr 4)/ <i>PtSPL16</i> (Chr11)	0.22	12
<i>PtSPL20</i> (Chr1)/ <i>PtSPL25</i> (Chr11)	0.39	21
<i>PtSPL7</i> (Chr2)/ <i>PtSPL6</i> (Chr14)	0.25	13
<i>PtSPL15</i> (Chr12)/ <i>PtSPL14</i> (Chr15)	0.30	16

Identification of 25 conserved motifs

Conserved domains of *PtSPLs* were analyzed using Pfam (<http://pfam.sanger.ac.uk>) and by BLAST analysis of protein sequences against the Conserved Domain Database (CDD, <http://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi>). The results showed that all of the 28 *PtSPLs* and 16 *AtSPLs* contained a SBP domain with about 78 amino acid residues

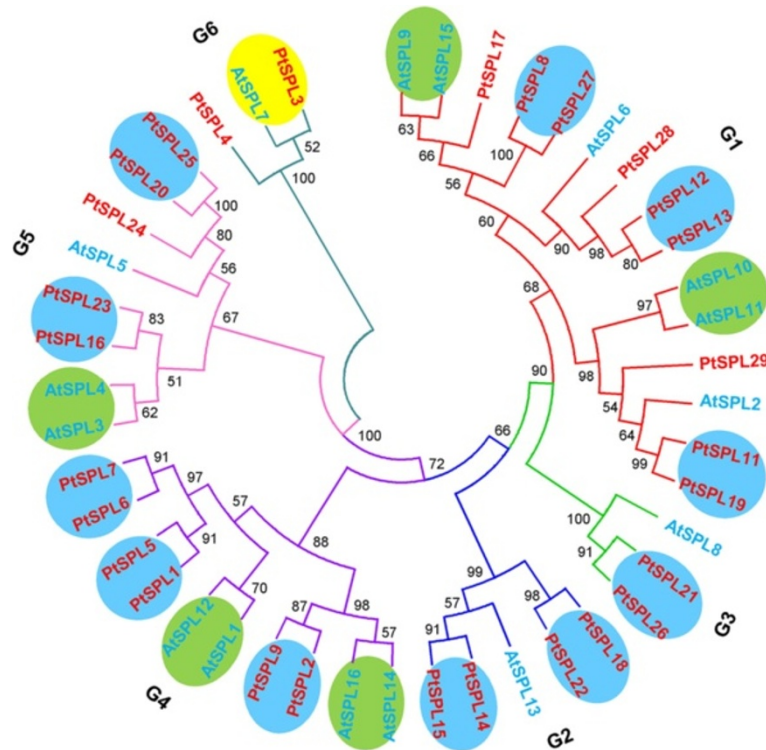
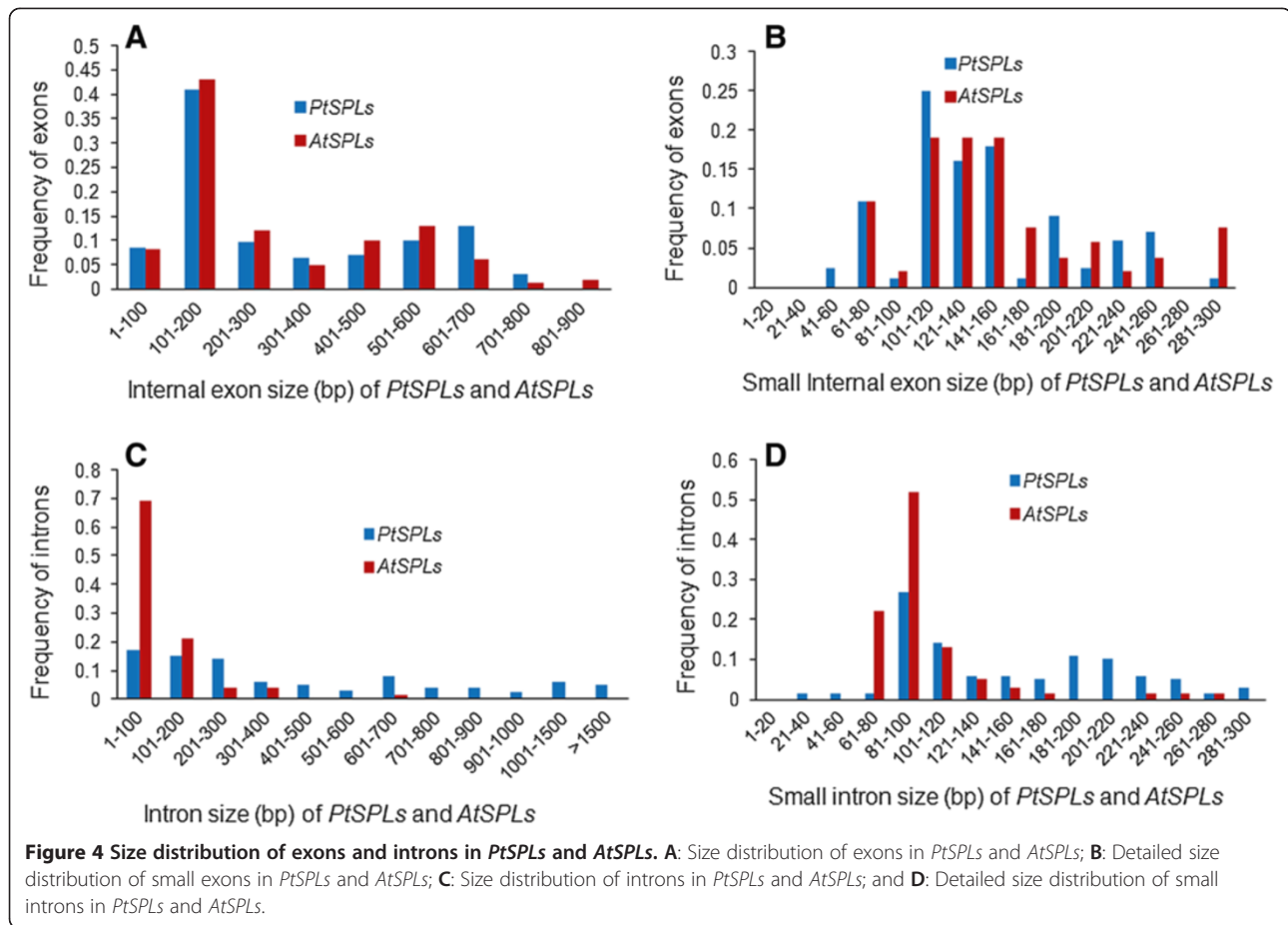


Figure 2 Neighbor-joining (NJ) phylogenetic tree for 44 SPLs in *P. trichocarpa* and *Arabidopsis*. The groups of homologous genes identified and bootstrap values are shown. The reliability of branching was assessed by the bootstrap re-sampling method using 1,000 bootstrap replicates. Bootstrap values are shown below nodes.



Figure 3 Exon-intron structures of *PtSPLs*. Introns are represented by lines. Exons are indicated by green boxes. The SBP domains are shown in red boxes. Intron phases are shown by 0, 1 and 2.



in length (Figure 5). It is not surprising given that the SBP domain was used for *PtSPL* identification. Sequence analysis of SBP domains revealed that the conserved zinc-binding sites, Zn1 and Zn2, also existed in the SBP domain of PtSPLs (Figure 5). Zn1 is Cys3His-type (CCCH-type) in G1–G5 SPLs (Figure 5A); however, the His residue in Zn1 is replaced by a Cys residue in G6, which results in the signature sequence of G6 SPLs to be CCCC (Figure 5B). Unlike Zn1, the signature sequence (C2HC) of Zn2 is highly conserved in all SPLs analyzed. In addition to Zn1 and Zn2, the SBP domain contains a conserved nuclear location signal (NLS) in the C-terminus of SBP domains (Figure 5). The conservation of SBP domains between PtSPLs and AtSPLs indicates that the domain organization has been established in ancient plants. Moreover, six PtSPLs (PtSPL1, PtSPL2, PtSPL5, PtSPL6, PtSPL7 and PtSPL9) belonging to G4 contain an ANK or Ank-2 domain with three or four ankyrin repeats (Additional file 5), which are involved in protein-protein interaction [42]. It is consistent with previous results from AtSPLs and SmSPLs [41].

In addition to the conserved domains, other conserved motifs could also be important for the function of SPLs [27,43]. We searched conserved motifs using MEME and applied an e-value cut off of $1e^{-10}$

to the recognition. It resulted in the identification of 25 motifs for 28 PtSPLs (Figure 6, Table 3). The majority of motifs identified are conserved between PtSPLs and AtSPLs [41], while three, including motifs 11, 19 and 23, are specific to PtSPLs. It indicates the conservation and diversity of PtSPLs and AtSPLs. The number of motifs in each SPL varies from 1 to 16 (Figure 6). Motif 1 is actually the SBP domain. Consistently, it exists in all SPLs analyzed. Motif 14 existed in G1 and G2 SPLs contains the target gene sequence of miR156, indicating the posttranscriptional regulation of G1 and G2 SPLs by miR156. In addition to motifs 1 and 14, several motifs widely exist in two SPL groups, such as motif 12 found in G1 and G2, motifs 2, 4, 5, 6, 15 and 16 existing in G4 and G6 (Figure 6), indicating the importance of these motifs. We also found several motifs to be group-unique, such as motif 24 specifically existing in G6 SPLs and motifs 7, 9, 10 and 18 specific to G4 (Figure 6). These group-unique motifs could be important for specific roles of SPLs in the group. Moreover, PtSPLs and AtSPLs [41] within a group share similar motif (s), indicating they probably play similar roles in plant growth and development.

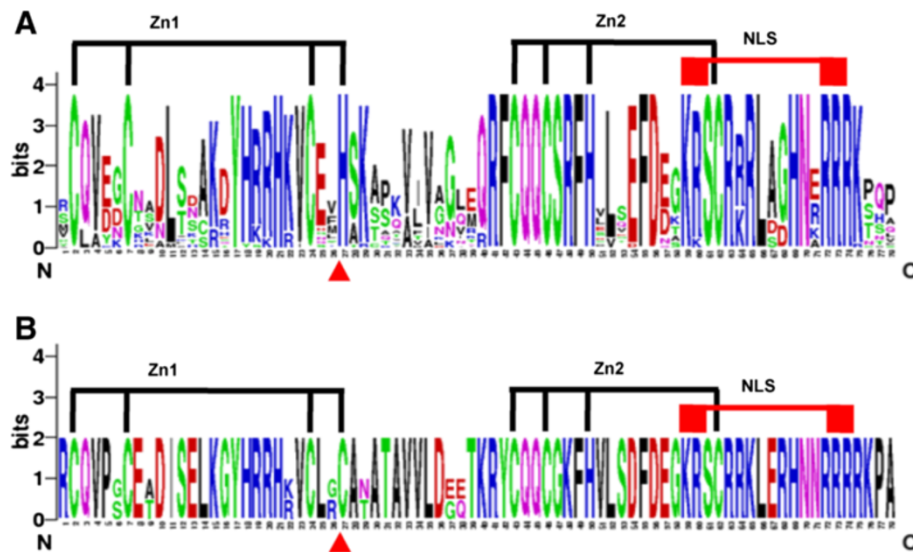


Figure 5 Sequence logo of the SBP domain of PtSPLs. **A:** Sequence logo of the SBP domain of PtSPLs in G1–G5; **B:** Sequence logo of the SBP domain of PtSPLs in G6. Two conserved Zn-finger structures and the NLS are indicated.

Expression patterns of SPLs in *P. trichocarpa*

The expression pattern of a gene is often correlated with its function. In order to preliminarily elucidate the roles of *PtSPLs* in *P. trichocarpa* development, we first searched PopGenIE for gene expression data from microarray analysis [44]. Except for *PtSPL17*, the expression levels of 27 *PtSPLs* in roots, stems, young leaves and mature leaves were obtained (Figure 7). Next, we examined the relative expression levels of 28 *PtSPLs* in young leaves, mature leaves, young stems, young roots and tissues from developing secondary xylem and phloem from the 4th–6th and 12th–25th internodes of one-year-old *P. trichocarpa* plants using the quantitative real-time RT-PCR method (Figure 8). The results showed that qRT-PCR data was generally consistent with microarray data for relative expression of *PtSPLs* in roots, stems, young leaves and mature leaves (Figures 7 and 8). Although all *PtSPLs* were expressed in at least one of the tissues examined, differential expression was observed. Many putative paralogous genes, such as *PtSPL18/22* in G2, *PtSPL21/26* in G3, *PtSPL2/9*, *PtSPL1/5* and *PtSPL6/7* in G4 and *PtSPL16/23* belonging to G5, show similar expression patterns, suggesting redundant roles of these *PtSPL* gene pairs. However, the expression patterns of few gene pairs, including *PtSPL12/13* in G1, and *PtSPL14/15* belonging to G2 are distinct. It indicates these *PtSPLs* may play different roles in *P. trichocarpa* development, although they are paralogous genes.

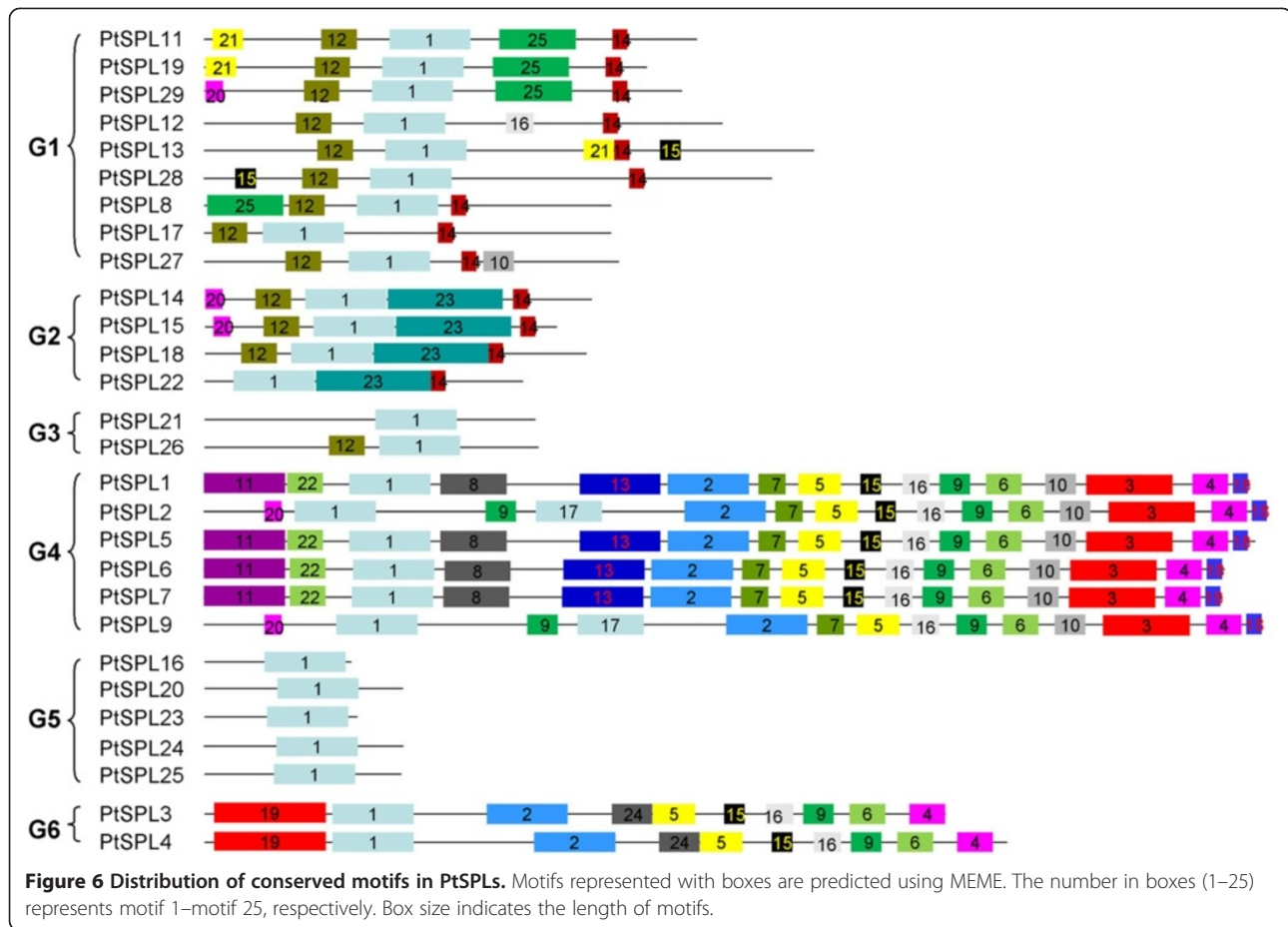
MiR156-mediated posttranscriptional regulation of *PtSPLs*

It has been shown that 10 *AtSPLs* are regulated by miR156 [11]. The complementary sites of miR156 are in the coding regions or 3' UTRs of *AtSPLs*. In order to

know miR156-mediated posttranscriptional regulation of *PtSPLs*, we searched coding regions and 3' UTRs of all *PtSPLs* for targets of *P. trichocarpa* miR156a–miR156j on the psRNATarget server using default parameters [45]. The results showed that 18 *PtSPLs* were potential targets of miR156 (Figures 9 and 10). MiR156-targeting sites in 13 *PtSPLs* belonging to G1 and G2 locate in the last exon and encode the conserved peptide ALSLLS. The target sites for other 5 *PtSPLs* belonging to G5 locate in the 3' UTRs close to the stop codons (Figure 10). Consistently, *AtSPLs* clustering in G1, G2 and G5 are targets of miR156 in *Arabidopsis*. It suggests that miR156-mediated posttranscriptional regulation of SPLs is conserved in *P. trichocarpa* and *Arabidopsis*.

Discussion

SPLs are plant-specific transcription factors containing a highly conserved SBP (SQUAMOSA PROMOTER BINDING PROTEIN) domain. It can specifically bind to the promoters of floral meristem identity gene *SQUAMOSA* and its orthologous genes and plays important regulatory roles in plant growth and development [46–49]. The genes encoding *SPLs* have been identified from various plant species, such as *Arabidopsis* [2,10,23,26], maize [30], *Antirrhinum majus* [3], rice [50], silver birch [51], and *S. miltiorrhiza* [41]. *SPL* genes exist as a large gene family in plants. The number of *SPLs* in *Arabidopsis*, rice, *P. patens*, maize and tomato is 16, 19, 13, 31 and 15, respectively [4–9]. Availability of the whole genome sequence allows us to perform genome-wide identification of *SPLs* in *P. trichocarpa*. Analysis of three versions of the annotated *P. trichocarpa* genome showed the existence of 28 full-length



PtSPLs, which distribute on 14 chromosomes. It is the first attempt to analyze the *PtSPL* gene family. The results provide a basis for elucidating the functions of *SPLs* in *P. trichocarpa*, a model forest tree.

The number of *SPL* genes in *P. trichocarpa* is much greater than that in *Arabidopsis*, rice, *P. patens* and tomato, although it is similar to the number of maize *SPLs* [4-9]. Sequence homologous analysis suggests that gene duplication plays an important role in *SPL* gene expansion in *P. trichocarpa*. A total of 11 pairs of intrachromosome-duplicated *PtSPLs* were identified in this study. All of them clustered together in the phylogenetic tree (Figure 2). It is consistent with previous findings for generation and maintenance of gene families in other organisms, such as mouse, human and *Arabidopsis* [52,53]. Actually, gene duplication has been reported for many plant transcription factor gene families, such as *MYB*, *AP2*, *MADS* and so on [54-56] and duplicated *SPL* gene pairs have been identified in *Arabidopsis* (*AtSPL10/11*, *AtSPL4/5* and *AtSPL1/12*) and rice (*OsSPL2/19*, *OsSPL3/12*, *OsSPL4/11*, *OsSPL5/10* and *OsSPL16/18*) [57-61]. However, the number of homologous *PtSPL* gene pairs is obviously greater than that in *Arabidopsis* and rice, indicating that more segment duplication events happened

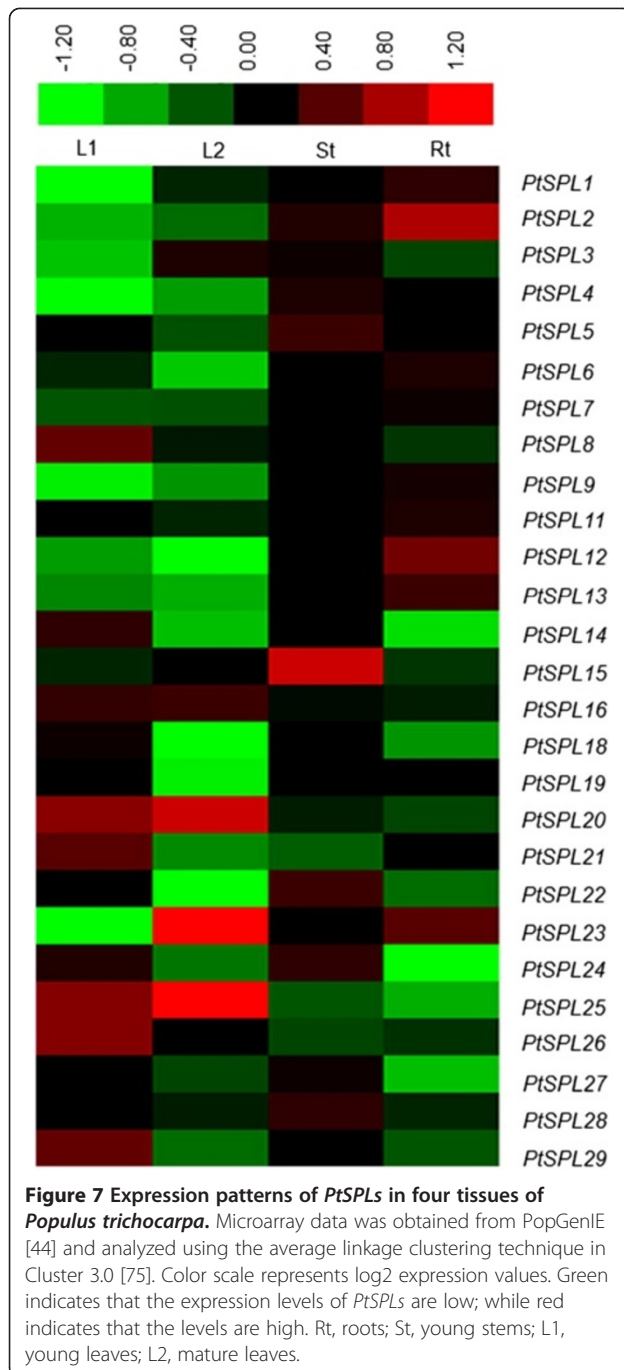
in *Populus* and most *SPL* genes in *Arabidopsis* and *Populus* expanded in a species-specific manner [62-64].

Comparative analysis of *P. trichocarpa PtSPLs* and *Arabidopsis AtSPLs* revealed many conserved sequence features. For instance, all of the deduced proteins contain the highly conserved SBP domain with about 78 amino acid residues. The intron position and intron phase in the SBP-domain-encoding regions are also conserved among all *SPL* genes in *P. trichocarpa* and *Arabidopsis*, indicating that plant *SPL* genes originate from a common ancestor. Based on the neighbor-joining (NJ) phylogenetic tree constructed using MEGA 5.1., 44 *SPL* proteins from *P. trichocarpa* and *Arabidopsis* were found to cluster into 6 groups. Each group includes at least a *PtSPL* and one *AtSPL*. The intron number and intron phase are similar for *PtSPLs* and *AtSPLs* within a group. The results suggest the conservation between *P. trichocarpa PtSPLs* and *Arabidopsis AtSPLs*.

It has been shown that *AtSPLs* play significant regulatory roles in a variety of developmental processes in *Arabidopsis*. For instance, morphological traits of cauline leaves and flowers are regulated by *AtSPL2*, *AtSPL10* and *AtSPL11* [19]. Juvenile-to-adult growth phase transition and leaf initiation rate are controlled by the

Table 3 E-value and consensus sequences of 25 motifs identified in PtSPLs

Motif	E-value	Consensus sequence
1	1.0e-2549	CQVEGCNADLSSAKDYHRRHKVCEVHSAKPKVIVAGLEQRFCQCCSRFHLLSEFDEGKRSCRRRLAGHNERRRKPOPD
2	3.3e-516	SDQPSSSSSSGDAQCRTGRIVFKLFDKDPNDFPGLRRTQILDWLSHSTDMESYIRPGCIILTIYLAMPEAAWEELCCDLG
3	6.2e-387	LFRPDVAGPAGLTPHIAACKDGSVDLALTEDPGEVGISAWKNARDATGFTPYARLRGHHSYIHLVQRKLADKRNGQVSWI
4	3.6e-174	ASRSLLYRPAMLSMVAIAAVCVALLFKSCEPELVY
5	9.0e-136	VEAGEETEFWKGRNLYQPGTRLLCAVEGKYLVOETTQALMD
6	7.4e-135	FPLRRFKLLEFSMDRDWCAWRKLLDMLVEGNVCRD
7	5.6e-100	FWRTGWFYVRVQNQLAFHKNGQWLDTSL
8	3.0e-085	GGSMNDDQGYLLTSILSNLHSNRSDQTKDQDLLSHLLRSLASHAGEHNGRNLFGLLQGPRGL
9	1.8e-095	EGMPSKEQALDFLNEIGWLLHRSDLKSRL
10	4.8e-088	SSLEALSEMGLLHRAVRRNSRKMVELLLR
11	7.4e-085	MEARFGGESHIFYAPVPSDLKAVGKRGLEWDLNDWKWDGDLFIASPLNPVSDCRSRQFFPTGPGLEKAGGNNNSSSCS
12	3.2e-077	STSLGASxSSGESLLGLKLGKRIYFEDAxGxNNxK
13	3.5e-070	FSIPNNFAAKSEEPEATAGQIKLNFDLNDIYDDSDGIEDIERSHAPVNAGMGSDCPLMVQQDSHKSSPPHTSGNSDS
14	1.3e-052	ASDSDCALSLSSQS
15	8.9e-057	NFSCSxPNLLGRGFIEVD
16	9.8e-064	PFIADADVCEIRILEQEFD
17	8.7e-053	GERISSCNEPSESDSDSQGDSPNLPLQLFSSSPENESRPKVASSRKYFSSASSNPIEDRSPSSSP
18	1.8e-039	PFRWELLDYGT
19	5.1e-035	QHDGDMEIHLPPITTDWDWDGILDFAVDDQFPLSFDTPGDLTQPIDNPTPEIESQQLEAPVPDRVRKRDPRLTCSNFLAGIVPCACPEMDELLLEEAALPGKKRVRVARAG
20	2.8e-033	DDWNLKAWDWDGDEFEA
21	8.5e-032	MDCNGKPHLQWDWENLIMFNATTENSKK
22	2.7e-030	DEDNLGDEKGRELEKRRRVFIDDDLND
23	3.3e-028	VNSARIFSNQGTRYLHFGSSQIFSTSAMNAAWTGAACAERDPMNLSQSSMNFDRGNLFPGLSPNYKEGKQFPFLQGTSTIPGDSIHLNANSTLGNLQKMFSDGLNR
24	3.4e-027	KGRMRVYLNMFNMFVTKDGHSMKVNKVGHAPRLHYVHPTC
25	5.9e-025	DERQQMSHAWDKAPLVHARPANLNTWEGTSISKFTITKDYIAKPAEIGGNDGQFHLPGFDLTNGIATQHHSKSN



redundant action of *AtSPL9* and *AtSPL15* [21]. Pollen sac development, male fertility and GA biosynthesis and signaling are regulated by *AtSPL8*, a member of G3 [23-25]. Cu homeostasis in *Arabidopsis* is regulated by the member of group 6, *AtSPL7* [22]. In this study, we found that many motifs were unique to or mainly existed in a group of SPLs. It is consistent with the redundant roles of AtSPLs in a group and indicates that the members of PtSPLs in the same group may play similar roles as their *Arabidopsis* counterparts. The

function of SPLs in different groups could be functionally distinct. On the other hand, three PtSPL-specific motifs, including motifs 11, 19 and 23, were identified, suggesting that some PtSPLs may play species-specific roles. Consistently, most of paralogous *PtSPL* gene pairs in the same group show similar expression patterns, whereas a few of them exhibit differential patterns. The results indicate subfunctionalisation and neofunctionalisation of SPLs within a plant species and among different species.

MiR156-mediated posttranscriptional regulation is important for the function of a subset of *SPLs* [11,41,65]. Target prediction showed that all *PtSPLs* in groups 1, 2 and 5 were regulated by miR156. The complementary sites of miR156 locate in the coding region of G1 and G2 *SPLs*, whereas it locates in 3' UTR of G5 *SPLs*. It is consistent with the results from *Arabidopsis SPLs* and suggests the conservation of miR156-mediated posttranscriptional regulation in plants.

Conclusion

In this study, a total of 28 full-length SPLs were identified from the whole genome sequence of *P. trichocarpa*. Through a comprehensive analysis of gene structures, phylogenetic relationships, chromosomal locations, conserved motifs, expression patterns and miR156-mediated posttranscriptional regulation, the *PtSPL* gene family was characterized and compared with *SPLs* in *Arabidopsis*. The results showed that 28 PtSPLs and 16 AtSPLs clustered into 6 groups. Many PtSPLs and AtSPLs within a group are highly conserved in sequence features, gene structures, motifs, expression patterns and posttranscriptional regulation, suggesting the conservation of plant SPLs within a group. However, significant differences were observed for *SPLs* among groups. In addition, various motifs were identified in PtSPLs but not in AtSPLs. It suggests the diversity of plant SPLs. The results provide useful information for elucidating the functions of *SPLs* in *P. trichocarpa*.

Methods

Identification of *PtSPL* genes

The nucleotide sequences and deduced amino acid sequences of 16 known *SPL* genes in *Arabidopsis* [2,4] were obtained from the TAIR database (<http://www.arabidopsis.org>) (Additional file 2). The SBP domain of *AtSPLs* was identified using Pfam (<http://pfam.sanger.ac.uk>). BLAST search of *PtSPLs* against *Populus trichocarpa* v1.1, v2.2 and v3.0 was carried out using *AtSPL* SBP as the query sequences [32] (http://genome.jgi-psf.org/Poptr1_1/Poptr1_1.home.html, <http://www.phytozome.net/poplar.php#B>). An e-value cut off of $1e^{-5}$ was applied to the recognition. We also searched the databases for SBP using the keywords search tool on the web servers. Protein

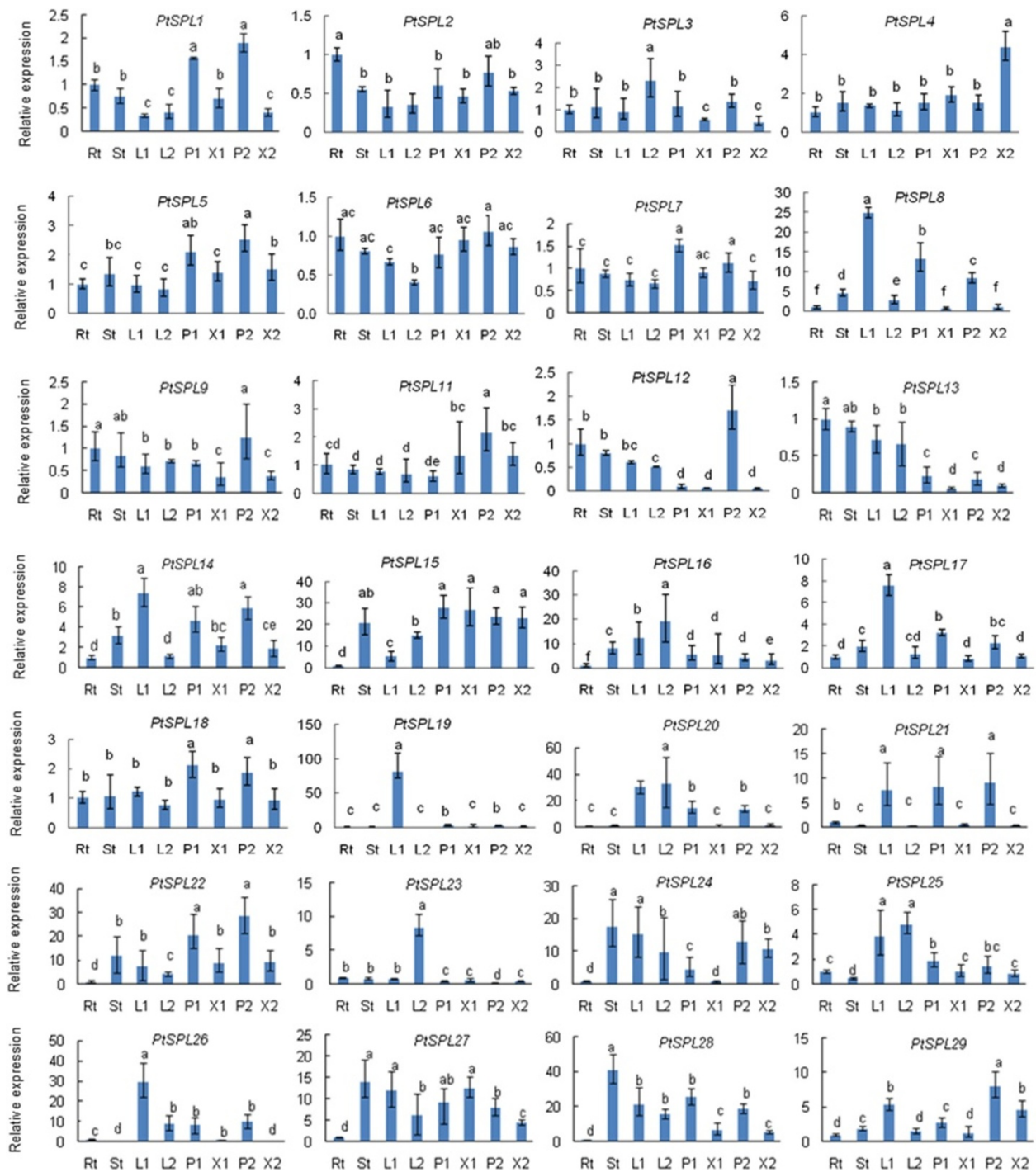


Figure 8 Analysis of *PtSPL* gene expression in eight tissues of *Populus trichocarpa* using the qRT-PCR method. Fold changes of transcript levels in young root (Rt), young stems (St), young leaves (L1), mature leaves (L2), developing secondary xylem (X1), developing secondary phloem (P1), developed secondary xylem (X2) and developed secondary phloem (P2) of *Populus* plants are shown. Transcript levels in roots were arbitrarily set to 1 and the levels in other tissues were given relative to this. Error bars represent standard deviations of mean value from three biological replicates. ANOVA (analysis of variance) was calculated using SPSS. P < 0.05 was considered statistically significant.

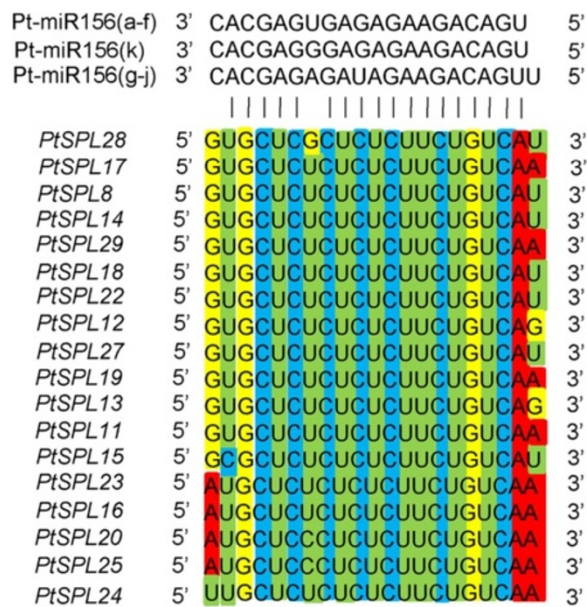


Figure 9 Sequence alignment of *P. trichocarpa* miR156a–miR156j with their complementary sequence in coding regions and 3' UTRs of 18 *PtSPLs*.

sequences retrieved from *Populus trichocarpa* v1.1, v2.2 and v3.0 were then aligned and combined based on sequence identities.

Chromosome location and sequence feature analyses

Chromosome locations of *PtSPL* genes were determined by BLAST analysis of *PtSPLs* against *Populus trichocarpa* v3.0 (<http://www.phytozome.net/poplar.php#B>). Paralogous gene pairs were analyzed on the Plant Genome Duplication Database (PGDD) server (<http://chibba.agtec.uga.edu/duplication/index/locus>) with display range for 100 kb. The approximate date of the duplication events was calculated using $T = Ks/2\lambda$ by assuming clock-like rates (λ) in *Populus* for 9.0×10^{-9} [32,57,66]. Synonymous substitutions (Ks) values of paralogous gene pairs were calculated using DnaSP [67]. The theoretical isoelectric point (pI) and molecular weight (Mw) were predicted using the Compute pI/Mw tool on the ExPASy server (http://web.expasy.org/compute_pi/) [68]. The intron/exon structure of *SPL* genes was predicted with the Gene Structure Display Server (<http://gsds.cbi.pku.edu.cn/chinese.php>) [69].

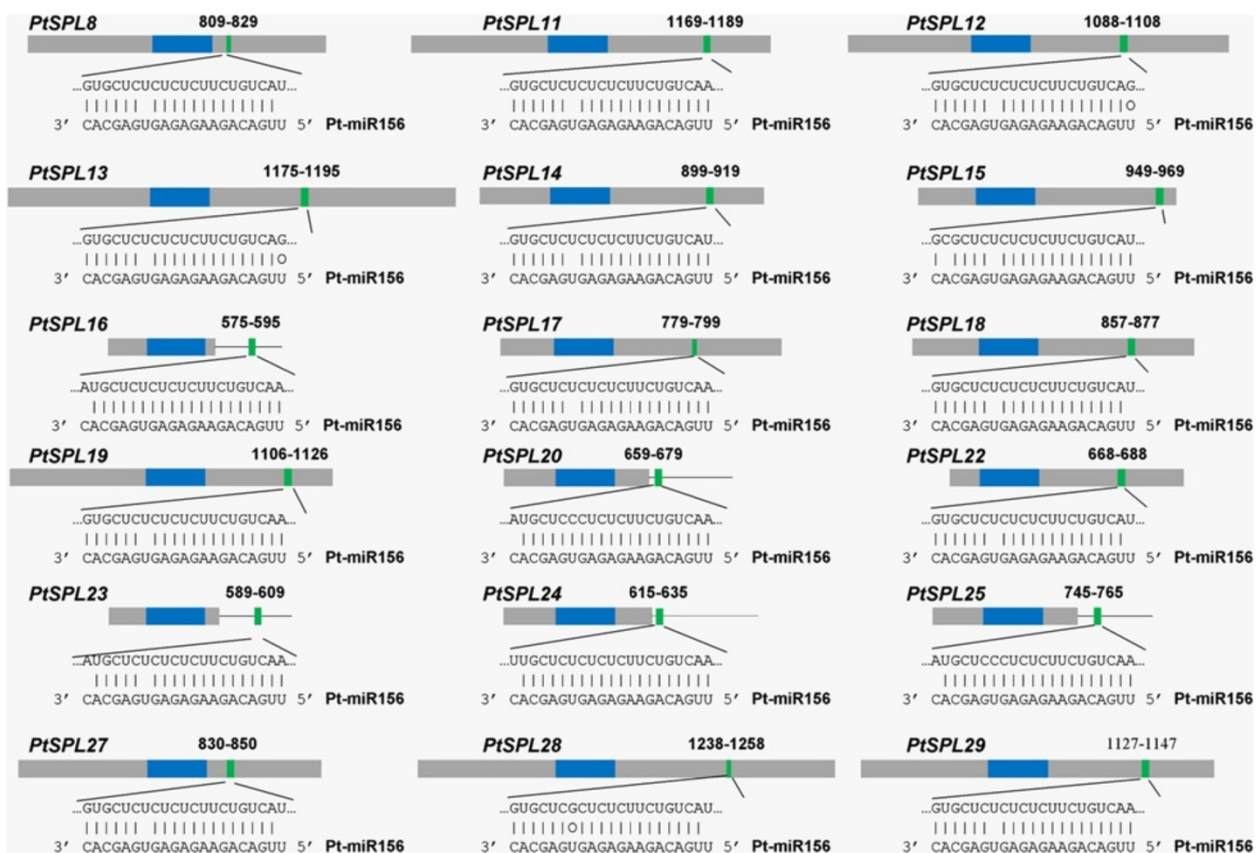


Figure 10 *PtSPLs* targeted by miR156. Heavy grey lines represent ORFs. The lines flanking ORFs represent 3'-UTR. The blue lines represent SBP domain. miRNA complementary sites (green) with the nucleotide positions of *PtSPL* cDNAs are indicated. The RNA sequence of each complementary site from 5' to 3' and the predicted miRNA sequence from 3' to 5' are shown in the expanded regions.

Phylogenetic construction and motif analysis

Phylogenetic trees were constructed using the neighbor-joining (NJ) method in MEGA5.1. Branching reliability was assessed by the bootstrap re-sampling method using 1,000 bootstrap replicates. Only nodes supported by bootstrap values greater than 50% were analyzed. Conserved domains of PtSPLs were identified using Pfam (<http://pfam.sanger.ac.uk>) and by BLAST analysis of protein sequences against the Conserved Domain Database (CDD, <http://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi>) with the expected e-value threshold of 1.0 and the maximum size of hits to be 500 amino acids [70]. The 78 amino acids of SBP domain were aligned using clustalW. Sequence logos were generated using the weblogo platform (<http://weblogo.berkeley.edu/>). Potential protein motifs were predicted using the MEME package (<http://meme.sdsc.edu/meme/>) with the following parameters applied. It includes the distribution of motifs: zero and one per sequence, maximum number of motifs to find: 25, minimum width of motif: 8, and maximum width of motif: 150. An e-value cut off of $1e^{-10}$ was applied to the recognition.

Quantitative real-time reverse transcription-PCR (qRT-PCR)

P. trichocarpa plants were grown in an artificial climate chamber for about one year. Young leaves (2nd–3rd from the top), mature leaves (12th from the top), young stems (1st–3rd from the top), young roots, tissues of developing secondary xylem and phloem from the 4th–6th and 12th–25th internodes from the top of *P. trichocarpa* plants were collected. Three biological repeats were carried out. Total RNA was extracted using the plant total RNA extraction kit (Aidlab, China). Genomic DNA contamination was eliminated by pre-treating total RNA with RNase-free DNase (Promega, USA). RNA integrity was analyzed on a 1.2% agarose gel and its quantity was determined using a NanoDrop 2000C Spectrophotometer (Thermo Scientific, USA). Total RNA was reverse-transcribed by Superscript III Reverse Transcriptase (Invitrogen, USA). qRT-PCRs were carried out in triplicate for each tissue sample using gene-specific primers (Additional file 6) as described previously [71]. The program used for qRT-PCR is as follows: pre-denaturation at 95°C for 30s, 40 cycles of amplification at 95°C for 5 s, 60°C for 18 s and 72°C for 15 s. The length of amplicons was between 80 bp and 250 bp. *Actin* was used as a reference gene as described previously [72]. Dissociation curve was used to assess amplification specificity. Relative abundance of transcripts was analyzed using the comparative Ct method [73]. The arithmetic formula, $2^{-\Delta\Delta Cq}$, was used to achieve results for relative quantification. Cq represents the threshold cycle. Standardization of gene expression data from three biological replicates was performed as described [74]. For statistical analysis, ANOVA (analysis of variance)

was calculated using SPSS (Version 19.0, IBM, USA). $P < 0.05$ was considered statistically significant.

Microarray data analysis

Microarray data of PtSPLs was obtained by the ePlant-tissue expression tool at PopGenIE (<http://www.popgenie.org/>). The data was gene-wise normalized and then analyzed using the average linkage clustering technique in Cluster 3.0 [75].

Prediction of PtSPLs targeted by miR156

The sequences of *P. trichocarpa* miR156a–miR156j were obtained from miRBase [36] (<http://www.mirbase.org/>). PtSPLs targeted by miR156 were predicted by searching the coding regions and 3' UTRs of all PtSPLs for complementary sequences of *P. trichocarpa* miR156a–miR156j on the psRNATarget server using default parameters [45] (<http://plantgrn.noble.org/psRNATarget/?function=3>).

Availability of supporting data

The data sets supporting the results of this article are included within the article and its additional files.

Additional files

Additional file 1: Sequence features of PtSPLs in *P. trichocarpa*. Protein length, intron number, pI and molecular weight of SPLs in *P. trichocarpa* are shown.

Additional file 2: Sequence features of AtSPLs in *A. thaliana*. Gene IDs, protein length, intron number, pI and molecular weight of SPLs in *A. thaliana* are shown.

Additional file 3: Neighbor-joining (NJ) phylogenetic tree constructed for 77 SPLs from *P. trichocarpa*, *Arabidopsis*, rice and *S. miltiorrhiza*. The groups of homologous genes identified and bootstrap values are shown. The reliability of branching was assessed by the bootstrap re-sampling method using 1,000 bootstrap replicates. Bootstrap values are shown below nodes.

Additional file 4: Intron distribution on SBP domains of *Populus* and *Arabidopsis*. Intron distribution on SBP domains of *Populus* and *Arabidopsis* are shown.

Additional file 5: Alignment of the ANK/ANK-2 domain. The ANK/ANK-2 domain is indicated by solid lines.

Additional file 6: Primers used for qRT-PCR analysis of PtSPL genes. Complete set of primers used for qRT-PCR.

Abbreviations

CuRE: Cu-response element; Mw: The molecular weight; NJ: Neighbor-joining; NLS: Nuclear location signal; pI: Isoelectric point; qRT-PCR: Quantitative realtime reverse transcription-PCR; SBP: SQUAMOSA PROMOTER BINDING PROTEIN; Zn1: Zinc finger 1; Zn2: Zinc finger 1.

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

CL contributed to bioinformatics and qRT-PCR analyses and participated in writing the manuscript. SL designed the experiment, performed bioinformatics analysis and wrote the manuscript. Both authors have read and approved the version of manuscript.

Acknowledgements

This work was supported by grants from the National Key Basic Research Program of China (973 program) (2012CB114502 to S.L.) and the Program for Xiehe Scholars in Chinese Academy of Medical Sciences & Peking Union Medical College (to S.L.).

Received: 30 March 2014 Accepted: 6 May 2014

Published: 15 May 2014

References

1. 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**:49–63.
2. 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**:585–596.
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**:7–16.
4. Cardon G, Hohmann S, Klein J, Nettessheim K, Saedler H, Huijser P: **Molecular characterisation of the *Arabidopsis* SBP-box genes.** *Gene* 1999, **237**:91–104.
5. Arazi T, Talmor-Neiman M, Stav R, Riese M, Huijser P, Baulcombe DC: **Cloning and characterization of micro-RNAs from moss.** *Plant J* 2005, **43**:837–848.
6. Hultquist JF, Dorweiler JE: **Feminized tassels of maize *mop1* and *ts1* mutants exhibit altered levels of miR156 and specific SBP-box genes.** *Planta* 2008, **229**:99–113.
7. Riese M, Zobell O, Saedler H, Huijser P: **SBP-domain transcription factors as possible effectors of cryptochrome-mediated blue light signalling in the moss *Physcomitrella patens*.** *Planta* 2008, **227**:505–515.
8. Miura K, Ikeda M, Matsubara A, Song X, Ito M, Asano K, Matsuoka M, Kitano H, Ashikari M: **OsSPL14 promotes panicle branching and higher grain productivity in rice.** *Nat Genet* 2010, **42**:545–549.
9. 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**:1171–1184.
10. Cardon GH, Hohmann S, Nettessheim K, Saedler H, Huijser P: **Functional analysis of the *Arabidopsis thaliana* SBP-box gene SPL3: a novel gene involved in the floral transition.** *Plant J* 1997, **12**:367–377.
11. Rhoades MW, Reinhart BJ, Lim LP, Burge CB, Bartel B, Bartel DP: **Prediction of plant microRNA targets.** *Cell* 2002, **110**:513–520.
12. Kasschau KD, Xie Z, Allen E, Llave C, Chapman EJ, Krizan KA, Carrington JC: **P1/HC-Pro a viral suppressor of RNA silencing interferes with *Arabidopsis* development and miRNA function.** *Dev Cell* 2003, **4**:205–217.
13. Chen J, Li WX, Xie D, Peng JR, Ding SW: **Viral virulence protein suppresses RNA silencing-mediated defense but upregulates the role of microRNA in host gene expression.** *Plant Cell* 2004, **16**:1302–1313.
14. Schwab R, Palatnik JF, Riester M, Schommer C, Schmid M, Weigel D: **Specific effects of microRNAs on the plant transcriptome.** *Dev Cell* 2005, **8**:517–527.
15. Wu G, Poethig RS: **Temporal regulation of shoot development in *Arabidopsis thaliana* by miR156 and its target SPL3.** *Development* 2006, **133**:3539–3547.
16. Wang JW, Czech B, Weigel D: **miR156-regulated SPL transcription factors define an endogenous flowering pathway in *Arabidopsis thaliana*.** *Cell* 2009, **138**:738–749.
17. Yu N, Cai WJ, Wang S, Shan CM, Wang LJ, Chen XY: **Temporal control of trichome distribution by microRNA156-targeted SPL genes in *Arabidopsis thaliana*.** *Plant Cell* 2010, **22**:2322–2335.
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* 2007, **49**:683–693.
19. 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**:2133–2145.
20. Vazquez F, Gasciolli V, Crété P, Vaucheret H: **The nuclear dsRNA binding protein HYL1 is required for microRNA accumulation and plant development but not posttranscriptional transgene silencing.** *Curr Biol* 2004, **14**:346–351.
21. Schwarz S, Grande AV, Bujdosó N, Saedler H, Huijser P: **The microRNA regulated SBP-box genes SPL9 and SPL15 control shoot maturation in *Arabidopsis*.** *Plant Mol Biol* 2008, **67**:183–195.
22. Yamasaki H, Hayashi M, Fukazawa M, Kobayashi Y, Shikanai T: **SQUAMOSA promoter binding protein-like7 is a central regulator for copper homeostasis in *Arabidopsis*.** *Plant Cell* 2009, **21**:347–361.
23. Unte US, Sorensen AM, Pesaresi P, Gandikota M, Leister D, Saedler H, Huijser P: **SPL8 an SBP-Box gene that affects pollen sac development in *Arabidopsis*.** *Plant Cell* 2003, **15**:1009–1019.
24. Xing S, Salinas M, Höhmann S, Berndtgen R, Huijser P: **miR156-targeted and nontargeted SBP-box transcription factors act in concert to secure male fertility in *Arabidopsis*.** *Plant Cell* 2010, **22**:3935–3950.
25. Zhang Y, Schwarz S, Saedler H, Huijser P: **SPL8 a local regulator in a subset of gibberellins-mediated developmental processes in *Arabidopsis*.** *Plant Mol Biol* 2007, **63**:429–439.
26. 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**:744–754.
27. Xie K, Wu C, Xiong L: **Genomic organization differential expression and interaction of SQUAMOSA promoter-binding-like transcription factors and micro-RNA156 in rice.** *Plant Physiol* 2006, **142**:280–293.
28. Jiao Y, Wang Y, Xue D, Wang J, Yan M, Liu G, Dong G, Zeng D, Liu Z, Zhu X, Qian Q, Li J: **Regulation of OsSPL14 by OsmiR156 defines ideal plant architecture in rice.** *Nat Genet* 2010, **42**:541–544.
29. Wang S, Wu K, Yuan Q, Liu X, Liu Z, Lin X, Zeng R, Zhu H, Dong G, Qian Q, Zhang G, Fu X: **Control of grain size shape and quality by OsSPL16 in rice.** *Nat Genet* 2012, **44**:950–955.
30. Becraft PW, Bongard-Pierce DK, Sylvester AW, Poethig RS, Freeling M: **The *liguleless-1* gene acts tissue specifically in maize leaf development.** *Dev Biol* 1990, **141**:220–232.
31. Moreno MA, Harper LC, Krueger RW, Dellaporta SL, Freeling M: ***liguleless1* encodes a nuclear-localized protein required for induction of ligules and auricles during maize leaf organogenesis.** *Genes Dev* 1997, **11**:616–628.
32. Tuskan GA, DiFazio S, Jansson S, Bohlmann J, Grigoriev I, Hellsten U, Putnam N, Ralph S, Rombauts S, Salamov A, Schein J, Sterck L, Aerts A, Bhalerao RR, Bhalerao RP, Blaudez D, Boerjan W, Brun A, Brunner A, Busov V, Campbell M, Carlson J, Chalot M, Chapman J, Chen GL, Cooper D, Coutinho PM, Couturier J, Covert S, Cronk Q, et al: **The genome of black cottonwood *Populus trichocarpa* (Torr. & Gray).** *Science* 2006, **313**:1596–1604.
33. Lu S, Sun YH, Shi R, Clark C, Li L, Chang VL: **Novel and mechanical stress-responsive microRNAs in *Populus trichocarpa* that are absent from *Arabidopsis*.** *Plant Cell* 2005, **17**:2186–2203.
34. Tuskan GA, DiFazio SP, Teichmann T: **Poplar genomics is getting popular: the impact of the poplar genome project on tree research.** *Plant Biol* 2004, **6**:2–4.
35. Lu S, Sun YH, Chiang VL: **Stress-responsive microRNAs in *Populus*.** *Plant J* 2008, **55**:131–151.
36. Kozomara A, Griffiths-Jones S: **miRBase: integrating microRNA annotation and deep-sequencing data.** *Nucleic Acids Res* 2011, **39**:D152–D157.
37. Puzey JR, Karger A, Axtell M, Kramer EM: **Deep annotation of *Populus trichocarpa* microRNAs from diverse tissue sets.** *PLoS One* 2012, **7**:e33034.
38. Lu S, Li Q, Wei H, Chang MJ, Tunlaya-Anukit S, Kim H, Liu J, Song J, Sun YH, Yuan L, Yeh TF, Peszlen I, Ralph J, Sederoff RR, Chiang VL: **PTR-miR397a is a negative regulator of laccase genes affecting lignin content in *Populus trichocarpa*.** *Proc Natl Acad Sci USA* 2013, **110**:10848–10853.
39. Shuai P, Liang D, Zhang Z, Yin W, Xia X: **Identification of drought-responsive and novel *Populus trichocarpa* microRNAs by high-throughput sequencing and their targets using degradome analysis.** *BMC Genomics* 2013, **14**:233.
40. Lu S, Yang C, Chiang VL: **Conservation and diversity of microRNA-associated copper-regulatory networks in *Populus trichocarpa*.** *J Integr Plant Biol* 2011, **53**:879–891.
41. 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* 2013, **56**:38–50.
42. Mosavi LK, Minor DL Jr, Peng ZY: **Consensus-derived structural determinants of the ankyrin repeat motif.** *Proc Natl Acad Sci USA* 2002, **99**:16029–16034.

43. Guo AY, Zhu QH, Gu X, Ge X, Yang J, Luo J: **Genome-wide identification and evolutionary analysis of the plant specific SBP-box transcription factor family.** *Gene* 2008, **418**:1–8.
44. Sjödin A, Street NR, Sandberg G, Gustafsson P, Jansson S: **The Populus Genome Integrative Explorer (PopGenIE): a new resource for exploring the Populus genome.** *New phytol* 2009, **182**:1013–1025.
45. Dai X, Zhao PX: **psRNATarget: a Plant Small RNA Target Analysis Server.** *Nucleic Acids Res* 2011, **39**:W155–W159.
46. Huijser P, Klein J, Lonig W, Meijer H, Saedler H, Sommer H: **Bracteomania an inflorescence anomaly is caused by the loss of function of the MADS-box gene SQUAMOSA in *Antirrhinum majus*.** *EMBO J* 1992, **11**:1239–1249.
47. Saedler H, Becker A, Winter KU, Kirchner C, Theissen G: **MADS-box genes are involved in floral development and evolution.** *Acta Biochim Pol* 2001, **48**:351–358.
48. Fornara F, Parenicova L, Falasca G, Pelucchi N, Masiero S, Ciannamea S, Lopez-Dee Z, Altamura MM, Colombo L, Kater MM: **Functional characterization of OsMADS18 a member of the AP1/SQUA subfamily of MADS box genes.** *Plant Physiol* 2004, **135**:2207–2219.
49. Robles P, Pelaz S: **Flower and fruit development in *Arabidopsis thaliana*.** *Int J Dev Biol* 2005, **49**:633–643.
50. Shao CX, Takeda Y, Hatano S, Matsuoka M, Hirano HY: **Rice genes encoding the SBP domain protein which is a new type of transcription factor controlling plant development.** *Rice Genet Newsl* 1999, **16**:114.
51. Lännenpää M, Jänönen I, Hölttä-Vuori M, Gardemeister M, Porali I, Sopanen T: **A new SBP-box gene *BpSPL1* in silver birch (*Betula pendula*).** *Physiol Plant* 2004, **120**:491–500.
52. Kent WJ, Baertsch R, Hinrichs A, Miller W, Haussler D: **Evolution's cauldron: duplication deletion and rearrangement in the mouse and human genomes.** *Proc Natl Acad Sci USA* 2003, **100**:11484–11489.
53. Cannon SB, Mitra A, Baumgarten A, Young ND, May G: **The roles of segmental and tandem gene duplication in the evolution of large gene families in *Arabidopsis thaliana*.** *BMC Plant Biol* 2004, **4**:10.
54. Du H, Yang SS, Liang Z, Feng BR, Liu L, Huang YB, Tang YX: **Genome-wide analysis of the MYB transcription factor superfamily in soybean.** *BMC Plant Biol* 2012, **12**:106.
55. Zahn LM, Kong H, Leebens-Mack JH, Kim S, Soltis PS, Landherr LL, Soltis DE, Depamphilis CW, Ma H: **The evolution of the SEPALLATA subfamily of MADS-box genes: a preangiosperm origin with multiple duplications throughout angiosperm history.** *Genetics* 2005, **169**:2209–2223.
56. Shigyo M, Hasebe M, Ito M: **Molecular evolution of the AP2 subfamily.** *Gene* 2006, **366**:256–265.
57. Blanc G, Wolfe KH: **Widespread paleopolyploidy in model plant species inferred from age distributions of duplicate genes.** *Plant Cell* 2004, **16**:1667–1678.
58. Bowers JE, Chapman BA, Rong J, Paterson AH: **Unravelling angiosperm genome evolution by phylogenetic analysis of chromosomal duplication events.** *Nature* 2003, **422**:433–438.
59. Paterson AH, Bowers JE, Chapman BA: **Ancient polyploidization predating divergence of the cereals and its consequences for comparative genomics.** *Proc Natl Acad Sci USA* 2004, **101**:9903–9908.
60. Wang X, Shi X, Hao B, Ge S, Luo J: **Duplication and DNA segmental loss in the rice genome: implications for diploidization.** *New Phytol* 2005, **165**:937–946.
61. 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**:1–11.
62. Bai J, Pennill LA, Ning J, Lee SW, Ramalingam J, Webb CA, Zhao B, Sun Q, Nelson JC, Leach JE, Hulbert SH: **Diversity in nucleotide binding site-leucine-rich repeat genes in cereals.** *Genome Res* 2002, **12**:1871–1884.
63. Zhang S, Chen C, Li L, Meng L, Singh J, Jiang N, Deng XW, He ZH, Lemaux PG: **Evolutionary expansion gene structure and expression of the rice wall-associated kinase gene family.** *Plant Physiol* 2005, **139**:1107–1124.
64. Jain M, Tyagi AK, Khurana JP: **Genome-wide analysis evolutionary expansion and expression of early auxin-responsive SAUR gene family in rice (*Oryza sativa*).** *Genomics* 2006, **88**:360–371.
65. Wu G, Park MY, Conway SR, Wang JW, Weigel D, Poethig RS: **The sequential action of miR156 and miR172 regulates developmental timing in *Arabidopsis*.** *Cell* 2009, **138**:750–759.
66. Shiu SH, Karolowski WM, Pan R, Tzeng YH, Mayer KF, Li WH: **Comparative analysis of the receptor-like kinase family in *Arabidopsis* and rice.** *Plant Cell* 2004, **16**:1220–1234.
67. Librado P, Rozas J: **DnaSP v5: a software for comprehensive analysis of DNA polymorphism data.** *Bioinformatics* 2009, **25**:1451–1452.
68. Bjellqvist B, Basse B, Olsen E, Celis JE: **Reference points for comparisons of two-dimensional maps of proteins from different human cell types defined in a pH scale where isoelectric points correlate with polypeptide compositions.** *Electrophoresis* 1994, **15**:529–539.
69. Guo AY, Zhu QH, Chen X, Luo JC: **GSDS: a gene structure display server.** *Yi Chuan* 2007, **29**:1023–1026.
70. Marchler-Bauer A, Lu S, Anderson JB, Chitsaz F, Derbyshire MK, DeWeese-Scott C, Fong JH, Geer LY, Geer RC, Gonzales NR, Gwadz M, Hurwitz DJ, Jackson JD, Ke Z, Lanczycki CJ, Lu F, Marchler GH, Mullokandov M, Omelchenko MV, Robertson CL, Song JS, Thanki N, Yamashita RA, Zhang D, Zhang N, Zheng C, Bryant SH: **CDD: A Conserved Domain Database for the functional annotation of proteins.** *Nucleic Acids Res* 2011, **39**:D225–D229.
71. Ma Y, Yuan L, Wu B, Li X, Chen S, Lu S: **Genome-wide identification and characterization of novel genes involved in terpenoid biosynthesis in *Salvia miltiorrhiza*.** *J Exp Bot* 2012, **63**:2809–2823.
72. Su X, Fan B, Yuan L, Cui X, Lu S: **Selection and validation of reference genes for quantitative RT-PCR analysis of gene expression in *Populus trichocarpa*.** *Chin Bull Bot* 2013, **48**:507–518.
73. Livak KJ, Schmittgen TD: **Analysis of relative gene expression data using real-time quantitative PCR and the 2^{-Delta Delta C(T)} method.** *Methods* 2001, **25**:402–408.
74. Willems E, Leyns L, Vandesompele J: **Standardization of real-time PCR gene expression data from independent biological replicates.** *Anal Biochem* 2008, **379**:127–129.
75. Eisen MB, Spellman PT, Brown PO, Botstein D: **Cluster analysis and display of genome-wide expression patterns.** *Proc Natl Acad Sci USA* 1998, **95**:14863–14868.

doi:10.1186/1471-2229-14-131

Cite this article as: Li and Lu: Molecular characterization of the SPL gene family in *Populus trichocarpa*. *BMC Plant Biology* 2014 **14**:131.

Submit your next manuscript to BioMed Central and take full advantage of:

- Convenient online submission
- Thorough peer review
- No space constraints or color figure charges
- Immediate publication on acceptance
- Inclusion in PubMed, CAS, Scopus and Google Scholar
- Research which is freely available for redistribution

Submit your manuscript at
www.biomedcentral.com/submit

