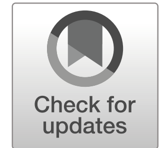
Tree Genetics & Genomes (2019) 15: 57
https://doi.org/10.1007/s11295-019-1362-5

SHORT COMMUNICATION



Regulation of *LaSCL6* expression by genomic structure, alternative splicing, and microRNA in *Larix kaempferi*

Qiao-Lu Zang[1] · Wan-Feng Li[1] · Li-Wang Qi[1]

Received: 28 January 2019 / Revised: 13 May 2019 / Accepted: 30 May 2019 / Published online: 22 June 2019
© The Author(s) 2019

Abstract
SCR-LIKE-6 (*SCL6*), a member of the *GRAS* transcription factor family, plays important roles in many aspects of plant growth and development. In a previous study, we showed that the post-transcriptional regulation of *Larix* (*La*)*SCL6* by the microRNA miR171 functions in regulating the mode of cell division and the maintenance of embryogenic potential during somatic embryogenesis in larch. To investigate its expression pattern during tree aging, which has been studied by comparative transcriptomic analysis, we re-examined the annotation and expression of every transcript and found one that was longer; we annotated it as *SCL6*, indicating that it might be a variant of *LaSCL6*. To verify this, we cloned the DNA and cDNA sequences of this transcript and found that alternative splicing indeed occurred in its expression. Moreover, both variants were detectable in embryogenic cultures and several organs. Notably, the DNA sequence of *LaSCL6* contained simple sequence repeats and a single-nucleotide polymorphism which might result in the expression of two variants with a change in their tertiary structures. Regulation of *LaSCL6* by miR171 was also found. Taken together, the expression of *LaSCL6* is controlled at several levels, and this study not only provides further information about the expression of *LaSCL6* but also offers a means to study the regulation of gene expression.

Keywords *Larix kaempferi* · *LaSCL6* · Genomic structure · Alternative splicing · microRNA171

Introduction

The microRNA171 (miR171) family is highly conserved, and its target genes are members of the *GRAS* (*GAI-RGA-SCR*) family, which are crucial for plant growth and development (Huang et al. 2015; Pysh et al. 1999). The miR171-*SCR-LIKE 6* (*SCL6*, also known as *HAM* or *LOST MERISTEMS*) module is involved in shoot branching (Llave et al. 2002; Wang et al. 2010), phase transition (Curaba et al. 2013), meristem maintenance (Fan et al. 2015), root growth and leaf formation (Hai et al. 2017), and somatic embryogenesis (Li et al. 2017a). In *Larix kaempferi*, the regulation of *LaSCL6* by miR171 participates in the control of the mode of cell division and the maintenance of embryogenic potential (Li et al. 2014). Furthermore, in the transcriptome of the stems of *L. kaempferi* at different ages (Li et al. 2017b), there is a longer transcript (3295 bp) also annotated as *SCL6*, indicating that it might be a variant of *LaSCL6* and that alternative splicing (AS) may occur after its transcription.

After the transcription of a gene, a single precursor mRNA is produced, while only mature mRNA is translated into protein. During the production of mature mRNA from the precursor mRNA, AS plays important roles, by which exons or portions of exons or non-coding regions within a precursor mRNA are differentially joined or skipped via the selection of variable donor and acceptor sites (Reddy 2007; Baralle and Giudice 2017). Therefore, AS results in a suite of divergent mRNAs produced from a single gene and expands the diversity of protein and function (Graveley 2001; Reddy 2007). In plant species, 33–60% of mRNAs are alternatively spliced (Shen et al. 2014; Zhang et al. 2010), indicating that AS plays important regulatory roles in gene expression at the post-transcriptional level. In this study, we cloned the DNA and cDNA sequences of *LaSCL6* to verify AS in its expression. Meanwhile, the regulation of its expression by genomic structure was also discussed.

Communicated by G. G. Vendramin

✉ Wan-Feng Li
liwf@caf.ac.cn

✉ Li-Wang Qi
lwqi@caf.ac.cn

[1] State Key Laboratory of Tree Genetics and Breeding, Key Laboratory of Tree Breeding and Cultivation of the State Forestry Administration, Research Institute of Forestry, Chinese Academy of Forestry, Beijing 100091, People's Republic of China

Materials and methods

Plant materials

Mature and immature seeds of *L. kaempferi* were collected from a Dagujia seed orchard (42° 22′ N, 124° 51′ E), Liaoning Province, in Northeast China. The mature seeds were grown in a pot for three months; then, roots, stems, needles, and hypocotyls were collected from several seedlings and pooled, frozen in liquid nitrogen, and stored at −80 °C until DNA and RNA extraction. The endosperm was exfoliated individually from ten mature seeds for DNA extraction after immersion in water for one day. The embryonal-suspensor mass (ESM) was generated from immature seeds, and cell cultures were used for total RNA extraction and gene expression analysis. The ESMs were cultured on differentiation medium, and cultures were collected at different stages of somatic embryogenesis (0, 7, 14, 21, 28, 35, and 42 days) according to Zhang et al. (2012).

Cloning of genomic DNA and full-length cDNA

In our published transcriptome (Li et al. 2017b), a transcript was annotated as *SCL6* and longer than *LaSCL6* (GenBank accession no. JX280920). After alignment analysis, the primers 5′-TCAAGCCAACGCCAAAGC-3′ and 5′-AAGAAGCGAAGAAGCAGACG-3′ were used to clone the genomic DNA and full-length cDNA (Fig. 1) (Li et al. 2014).

The RNA and genomic DNA were isolated with the RNAiso Plus reagent kit (TaKaRa, Japan) and the CTAB plant genome DNA rapid extraction kit (Aidlab Biotech, China) according to the manufacturers' protocols, respectively. The quality and quantity of the DNA and RNA were measured using 1% agarose gel electrophoresis. The RNA was reverse transcribed into cDNA with the *TransScript*® II one-step gDNA removal and cDNA Synthesis SuperMix kit (TransGen, China). The genomic DNA and full-length cDNA were amplified with Platinum® Taq DNA polymerase (Invitrogen, USA). The PCR products were purified with the gel extraction kit (Tiangen, China), ligated into the *pEASY*®-T1 simple cloning vector (TransGen, China), and sequenced.

Sequence analysis

The structure was analyzed using Gene Structure Display Server 2.0 (Hu et al. 2015). Multiple sequence alignments were performed with ClustalX (Thompson et al. 1997). The tertiary structure was predicted by SWISS-MODEL (https://swissmodel.expasy.org/).

Simple sequence repeat analysis

The simple sequence repeat (SSR) analysis was conducted using SSRIT (Simple Sequence Repeat Identification Tool) (https://www. gramene.org/db/markers/ssrtool) with the maximum motif length group being a pentamer. To measure the frequency of SSRs, the primers 5′-AGCGAGGTCAAGAAAGAAGAGC-3′ and 5′-CAGATCCACCGCTAGTAT-3′ were used to amplify the DNA extracted from endosperms with the TIANcombi DNA Lyse & Det PCR kit (Tiangen, China).

Semi-quantitative reverse transcription polymerase chain reaction

Semi-quantitative reverse transcription polymerase chain reaction (RT-PCR) was used to assess the expression of different variants. *L. kaempferi* translation elongation factor-1 alpha 1 (*LaEF1A1*) (GenBank accession no. JX157845) served as the internal control (Li et al. 2014). The primers 5′-GCACAGTTGAAGCACCAGATT-3′ and 5′-CGCATAACAAACCACCAGAAA-3′ were designed to amplify the long transcript, and the primers 5′-TCTG

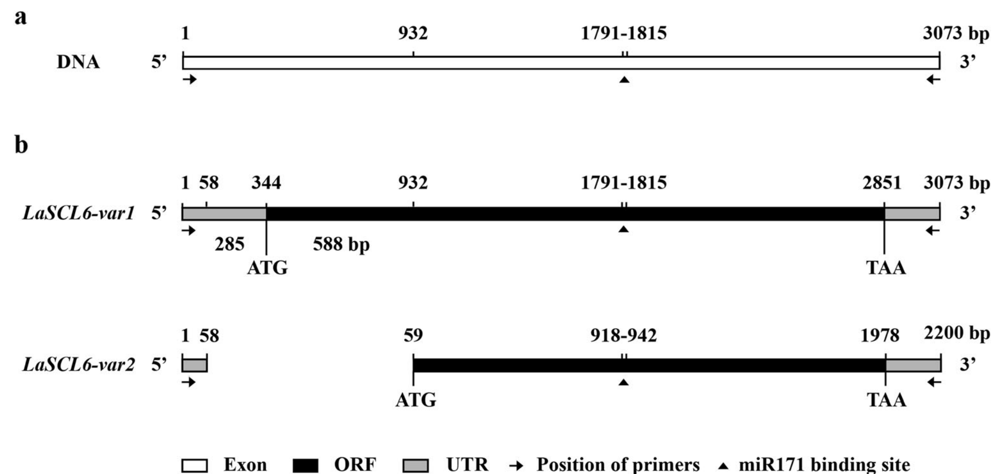


Fig. 1 Identification and characterization of two splice variants of *LaSCL6*. **a** Schematic representation of the genomic structure of *LaSCL6*. **b** Schematic representation of cDNA sequence of the variants *LaSCL6-var1* and *LaSCL6-var2*. White indicates exon; black indicates open reading frame (ORF); gray indicates the untranslated region (UTR); arrows indicate the position of the same primers used to amplify the genomic DNA and full-length cDNA sequences of *LaSCL6*; arrowheads indicate the binding site for miR171

GTGGTAATGGAAGA-3′ and 5′-CAGATCCACCGCTAGTAT-3′ were designed to amplify the short transcript (Fig. 4a). The PCR for *LaEF1A1* was limited to 28 cycles in order to stay within the linear range of amplification and obtain a more accurate picture of the relative levels of gene expression. The short transcript was amplified with 40 cycles, and the long transcript with 35 cycles, to be sure of their relative expression levels.

Quantitative reverse transcription polymerase chain reaction

RNAs extracted from roots, stems, needles, hypocotyls, and cell cultures were reverse transcribed and used to assay the expression patterns of *LaSCL6* with the same primers as previously described (Li et al. 2014). The quantitative reverse transcription polymerase chain reaction (qRT-PCR) was performed with three biological replicates, and the data are shown as mean ± SD. Statistical analysis was performed with SPSS 19.0 using analysis of variance.

Results and discussion

Identification and characterization of splice variants of *LaSCL6*

Only one DNA sequence with 3073 bp in length was amplified (Fig. 1a), indicating that the two transcripts were derived from the same gene and AS indeed occurred during the expression of *LaSCL6*. Here, the longer transcript (3073 bp) was named *LaSCL6-variant1* (*LaSCL6-var1*), and the shorter one (2200 bp), which was cloned in our previous work (Li et al. 2014), was named *LaSCL6-variant2* (*LaSCL6-var2*) (Fig. 1b). Genetic structural analysis showed that *LaSCL6* had one exon and no intron, which might be transcribed into *LaSCL6-var1* without splicing (Fig. 1a). Compared with *LaSCL6-var1*, 285 bp of the 5′ untranslated region and 588 bp of the open reading frame were spliced in *LaSCL6-var2* (Fig. 1b). Thus, *LaSCL6-var1* encoded a polypeptide of 835 amino acids (aa), while *LaSCL6-var2* encoded 639 aa, 196 aa shorter than *LaSCL6-var1*. The results showed that in the DNA sequence of *LaSCL6*, there was a single exon that was fully conserved in *LaSCL6-var1*, while due to the existence of *LaSCL6-var2*, we concluded that AS occurred in the regulation of *LaSCL6* expression. Various mechanisms are involved in AS (Graveley 2001; Keren et al. 2010); in the DNA sequence of *LaSCL6*, there existed an alternative exon (Graveley 2001), which was spliced after its transcription to produce the *LaSCL6-var2* transcript.

GRAS proteins contain a less conserved, variable N-terminal region, a conserved C-terminal region, and five characteristic domains located in the C-terminal region: LHRI, VHIID, LHRII, PFYRE, and SAW (Hakoshima 2018; Pysh et al. 1999). All these conserved domains were found in LaSCL6-var1 and LaSCL6-var2; the shortened 196 aa were in the N-terminal region and did not affect these five GRAS domains (Fig. 2). Three-dimensional structural analysis

Fig. 2 Alignment of LaSCL6-var1, LaSCL6-var2, and PtHAM1 amino acid sequences. Identical residues are indicated by asterisks. The GenBank accession numbers are MK501379 for *LaSCL6-var1*, JX280920 for *LaSCL6-var2*, and KJ711054 for *PtHAM1*

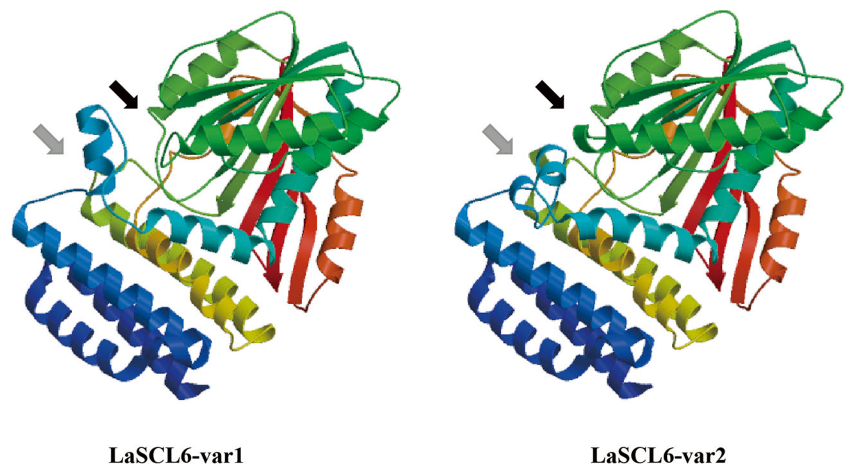


Fig. 3 Prediction of tertiary structures of LaSCL6-var1 and LaSCL6-var2. Differences in the tertiary structure of LaSCL6-var1 and LaSCL6-var2 are indicated by the black and gray arrows

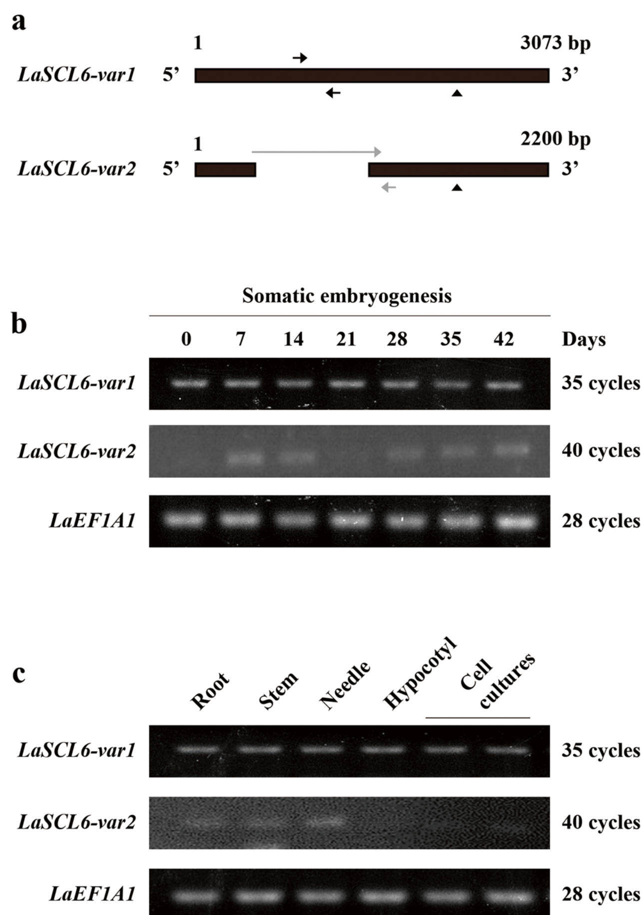


Fig. 4 Expression profiles of *LaSCL6-var1* and *LaSCL6-var2* in different samples assayed by semi-quantitative RT-PCR. *L. kaempferi* translation elongation factor-1 alpha 1 (*LaEF1A1*) was used as the internal control. **a** Schematic of primer design. Primers indicated by black arrows, which are located in the spliced part, were used to measure the expression of *LaSCL6-var1*; primers indicated by gray arrows, one of which spans the splice site, were used to measure the expression of *LaSCL6-var2*. Arrowheads indicate the binding site for miR171. **b** Expression profiles of *LaSCL6-var1* and *LaSCL6-var2* during somatic embryogenesis. **c** Expression profiles of *LaSCL6-var1* and *LaSCL6-var2* in roots, stems, needles, hypocotyls, and cell cultures

showed that the same aa sequence displayed different structures (Fig. 3), demonstrating that the 196 amino acids change the three-dimensional structure of the protein, and this might result in diversification of the function of the *LaSCL6* variants. Further work is needed to test this hypothesis.

The occurrence of AS in a gene makes its function diverse, and variants are expressed differentially depending on developmental stages and environmental conditions (Airoldi et al. 2015). Here, semi-quantitative RT-PCR analysis showed that the two variants of *LaSCL6* were detectable in many samples (Fig. 4b, c). *LaSCL6-var1* was readily detectable, and almost no differences in its transcript level were found, while *LaSCL6-var2* was detectable in some samples (Fig. 4b, c), indicating that the two variants are differentially expressed during the growth and development of *Larix*.

Effects of SSRs and a single-nucleotide polymorphism on *LaSCL6* expression

During DNA cloning and sequence analysis, SSRs were found in the genomic sequence of *LaSCL6*. So we analyzed

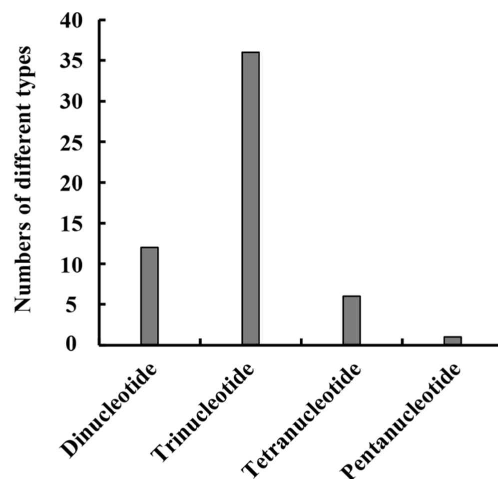


Fig. 5 Simple sequence repeats in the sequence of *LaSCL6*

the SSRs in this DNA sequence (3073 bp). We found four repeat types of SSRs when the minimum number of tandem repeats was set at two, and trinucleotide repeats were the most numerous (Fig. 5); when it was set at five, only two SSRs were identified: GCA and CCA, and the location of GCA repeats was in the spliced 588-bp open reading frame (Fig. 6a). To avoid heterozygosity and further verify the occurrence of the GCA repeats, 10 haploid endosperms were used to amplify the GCA repeat region. The results showed that seven GCA repeats occurred in six endosperms and nine GCA repeats occurred in four endosperms (Fig. 6b), indicating that there was polymorphism in this locus.

Exonic single-nucleotide polymorphisms (SNPs) have direct effects on the properties of proteins, while SNPs within introns and untranslated regions affect the expression and splicing of mRNA (Wang et al. 2014). When the needles from one seedling were used for DNA cloning, an SNP (C-T) was detected in the region of GCA repeats, and four sequences had C and two had T in this locus (Fig. 6c). Notably, due to this SNP, the CAG codon was changed to TAG, resulting in the termination of translation, and this might be another reason for the production of LaSCL6-var2 that is 639 aa in length.

SSRs and SNPs are routinely used in molecular genetics studies to reflect genetic diversity (García et al. 2018; Lind and Gailing 2013). Our results not only provide markers for molecular genetics studies but also provide an example of how the markers affect the structure and expression of a gene.

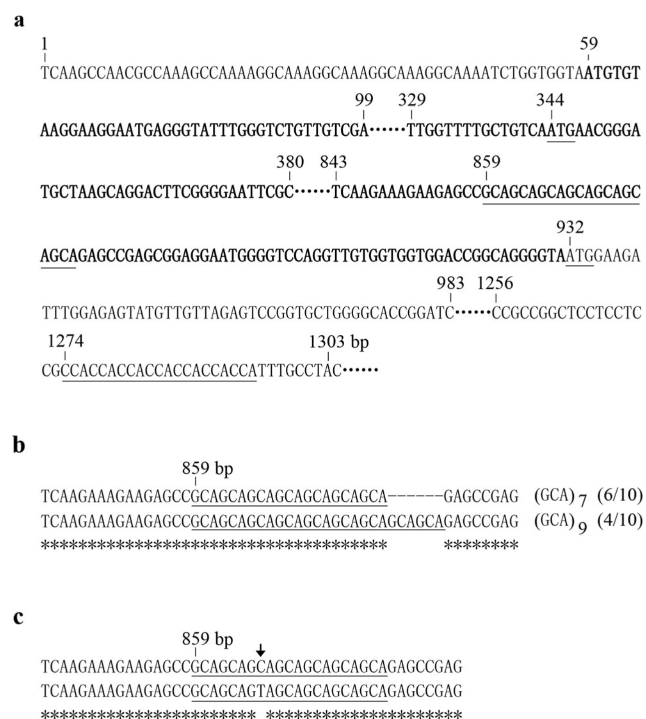


Fig. 6 DNA sequence analysis of *LaSCL6*. **a** The nucleotides in bold were spliced in the *LaSCL6-var2* transcript; the initiation codon ATG of *LaSCL6-var1* was located at 344 bp, and that of *LaSCL6-var2* at 932 bp. **b** Frequencies of GCA repeats in 10 endosperms. **c** Occurrence of an SNP (C-T) in the region of GCA repeats. Identical residues are indicated by asterisks

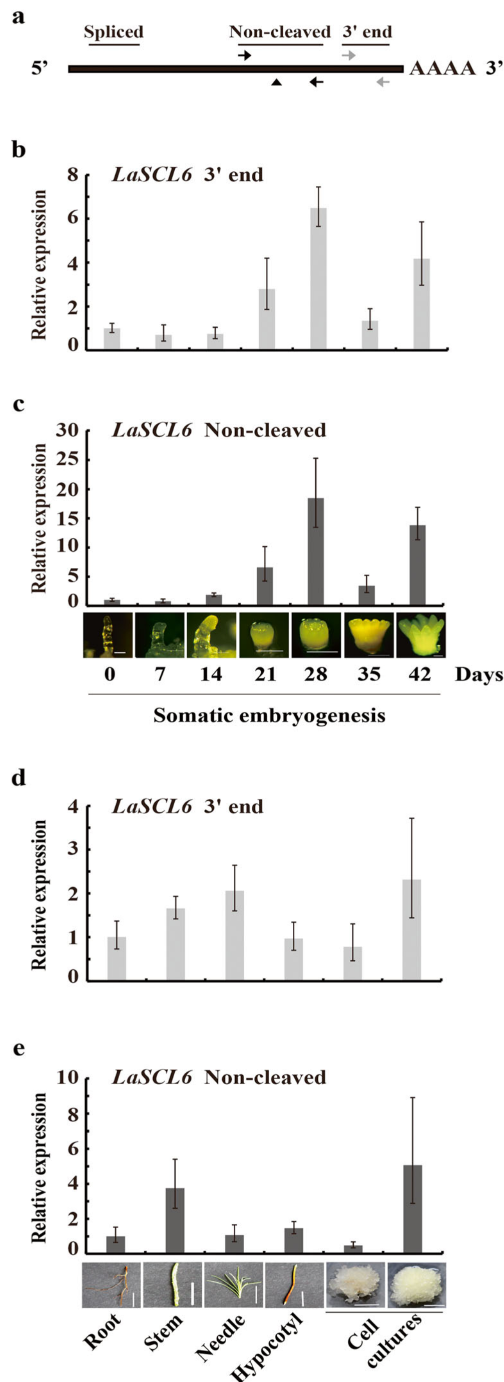


Fig. 7 Expression patterns of *LaSCL6* assayed by qRT-PCR using *LaEF1A1* as the internal control. **a** Schematic of the primer design. Primers indicated by black arrows, which span the miR171 cleavage sites, were used to measure the non-cleaved mRNAs (non-cleaved), and primers indicated by gray arrows were used to measure both non-cleaved and cleaved mRNAs (3′ end) (Li et al. 2014). The arrowhead indicates the binding site for miR171. **b**, **c** Expression patterns of *LaSCL6* during somatic embryogenesis. **d**, **e** Expression patterns of *LaSCL6* in roots, stems, needles, hypocotyls, and cell cultures

Regulation of *LaSCL6* by microRNA171 at the post-transcriptional level

The binding site for miR171 was not spliced in the two variants of *LaSCL6* (Fig. 7a), which meant that they might still be regulated by miR171. But it was difficult to determine which variant was regulated by miR171. The regulation of *LaSCL6* by miR171 could occur before or after AS and be independent of AS. With the method used in our previous study (Li et al. 2014), we also detected the expression patterns of the initial (both non-cleaved and cleaved transcripts) and the full-length *LaSCL6* transcripts (non-cleaved transcripts) to determine the activity of miR171 in other materials.

During somatic embryogenesis, the level of initial transcripts in cotyledon embryos was 2–6 times higher than that in single embryos (Fig. 7b), indicating that the expression of *LaSCL6* is regulated at the transcriptional level, while the level of full-length transcripts in cotyledon embryos was 6–18 times higher than that in single embryos (Fig. 7c), indicating that more full-length transcripts are cleaved in the development of single embryos, and the regulation of *LaSCL6* by miR171 at the post-transcriptional level occurs during somatic embryogenesis. In other materials, both kinds of the *LaSCL6* transcript were detected; notably, the level of initial transcripts in stems and needles showed almost no difference (Fig. 7d), but the level of full-length transcripts in stems was higher than that in needles (Fig. 7e), suggesting that in needles, more full-length transcripts are cleaved and the activity of miR171 is stronger.

Gene expression in eukaryotes is complex and regulated at many levels, including genomic structure, transcription, post-transcription, translation, and post-translation. In the case of *LaSCL6*, we concluded that at the genomic level, the occurrence of SSRs and an SNP in the exon affects its expression and protein structure, and AS and miRNA regulate its expression at the post-transcriptional level. These results not only provide more information about the expression of *LaSCL6* and its possible functions in plant growth and development but also offer a means to study the regulation of gene expression.

Acknowledgments The authors thank Dr. IC Bruce (Peking University) for the critical reading of the manuscript.

Author contributions Qiao-Lu Zang carried out the study, analyzed the data, and wrote the manuscript. Wan-Feng Li designed the study and revised the manuscript. Li-Wang Qi provided suggestions on the experimental design and analyses. All authors read and approved the manuscript.

Funding information This work was supported by the National Natural Science Foundation of China (31770714 and 31330017), the Basic Research Fund of Research Institute of Forestry (RIF2014-07), and the National Transgenic Major Program (2018ZX08020-003).

Data archiving statement The full-length cDNA sequences have been submitted to GenBank with the accession number MK501379 for *LaSCL6 variant 1*.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

Airoldi CA, McKay M, Davies B (2015) *MAF2* is regulated by temperature-dependent splicing and represses flowering at low temperatures in parallel with *FLM*. PLoS One 10:e0126516. https://doi.org/10.1371/journal.pone.0126516

Baralle FE, Giudice J (2017) Alternative splicing as a regulator of development and tissue identity. Nat Rev Mol Cell Biol 18:437–451. https://doi.org/10.1038/nrm.2017.27

Curaba J, Talbot M, Li Z, Helliwell C (2013) Over-expression of microRNA171 affects phase transitions and floral meristem determinancy in barley. BMC Plant Biol 13:6–16. https://doi.org/10.1186/1471-2229-13-6

Fan T, Li X, Yang W, Xia K, Ouyang J, Zhang M (2015) Rice osa-miR171c mediates phase change from vegetative to reproductive development and shoot apical meristem maintenance by repressing four *OsHAM* transcription factors. PLoS One 10:e0125833. https://doi.org/10.1371/journal.pone.0125833

García C, Guichoux E, Hampe A (2018) A comparative analysis between SNPs and SSRs to investigate genetic variation in a juniper species (*Juniperus phoenicea* ssp. *turbinata*). Tree Genet Genomes 14:1–9. https://doi.org/10.1007/s11295-018-1301-x

Graveley BR (2001) Alternative splicing: increasing diversity in the proteomic world. Trends Genet 17:100–107. https://doi.org/10.1016/S0168-9525(00)02176-4

Hai BZ, Qiu ZB, He YY, Yuan MM, Li F (2017) Characterization and primary functional analysis of *Pinus densata* miR171. Biol Plant 62:318–324. https://doi.org/10.1007/s10535-018-0774-7

Hakoshima T (2018) Structural basis of the specific interactions of GRAS family proteins. FEBS Lett 592:489–501. https://doi.org/10.1002/1873-3468.12987

Hu B, Jin J, Guo A-Y, Zhang H, Luo J, Gao G (2015) GSDS 2.0: an upgraded gene feature visualization server. Bioinformatics 31:1296–1297. https://doi.org/10.1093/bioinformatics/btu817

Huang W, Xian Z, Kang X, Tang N, Li Z (2015) Genome-wide identification, phylogeny and expression analysis of GRAS gene family in tomato. BMC Plant Biol 15:209–227. https://doi.org/10.1186/s12870-015-0590-6

Keren H, Lev-Maor G, Ast G (2010) Alternative splicing and evolution: diversification, exon definition and function. Nat Rev Genet 11:345–355. https://doi.org/10.1038/nrg2776

Li H, Zhang J, Yang Y, Jia N, Wang C, Sun H (2017a) miR171 and its target gene *SCL6* contribute to embryogenic callus induction and torpedo-shaped embryo formation during somatic embryogenesis in two lily species. Plant Cell Tissue Organ Cult 130:591–600. https://doi.org/10.1007/s11240-017-1249-9

Li WF, Yang WH, Zhang SG, Han SY, Qi LW (2017b) Transcriptome analysis provides insights into wood formation during larch tree aging. Tree Genet Genomes 13:19–31. https://doi.org/10.1007/s11295-017-1106-3

Li WF, Zhang SG, Han SY, Wu T, Zhang JH, Qi LW (2014) The post-transcriptional regulation of *LaSCL6* by miR171 during

maintenance of embryogenic potential in *Larix kaempferi* (Lamb.) Carr. Tree Genet Genomes 10:223–229. https://doi.org/10.1007/s11295-013-0668-y

Lind JF, Gailing O (2013) Genetic structure of *Quercus rubra* L. and *Quercus ellipsoidalis* E. J. Hill populations at gene-based EST-SSR and nuclear SSR markers. Tree Genet Genomes 9:707–722. https://doi.org/10.1007/s11295-012-0586-4

Llave C, Xie ZX, Kasschau KD, Carrington JC (2002) Cleavage of *Scarecrow-like* mRNA targets directed by a class of *Arabidopsis* miRNA. Science 297:2053–2056. https://doi.org/10.1126/science.1076311

Pysh LD, Wysocka-Diller JW, Camilleri C, Bouchez D, Benfey PN (1999) The GRAS gene family in Arabidopsis sequence characterization and basic expression analysis of the *SCARECROW-LIKE* genes. Plant J 18:111–119. https://doi.org/10.1046/j.1365-313X.1999.00431.x

Reddy AS (2007) Alternative splicing of pre-messenger RNAs in plants in the genomic era. Annu Rev Plant Biol 58:267–294. https://doi.org/10.1146/annurev.arplant.58.032806.103754

Shen Y, Zhou Z, Wang Z, Li W, Fang C, Wu M, Ma Y, Liu T, Kong LA, Peng DL, Tian Z (2014) Global dissection of alternative splicing in paleopolyploid soybean. Plant Cell 26:996–1008. https://doi.org/10.1105/tpc.114.122739

Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG (1997) The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res 25:4876–4882. https://doi.org/10.1093/nar/25.24.4876

Wang L, Mai YX, Zhang YC, Luo Q, Yang HQ (2010) MicroRNA171c-targeted *SCL6-II*, *SCL6-III*, and *SCL6-IV* genes regulate shoot branching in *Arabidopsis*. Mol Plant 3:794–806. https://doi.org/10.1093/mp/ssq042

Wang X, Zhong J, Gao Y, Ju Z, Huang J (2014) A SNP in intron 8 of *CD46* causes a novel transcript associated with mastitis in Holsteins. BMC Genomics 15:630–641. https://doi.org/10.1186/1471-2164-15-630

Zhang G, Guo G, Hu X, Zhang Y, Li Q, Li R, Zhuang R, Lu Z, He Z, Fang X, Chen L, Tian W, Tao Y, Kristiansen K, Zhang X, Li S, Yang H, Wang J, Wang J (2010) Deep RNA sequencing at single base-pair resolution reveals high complexity of the rice transcriptome. Genome Res 20:646–654. https://doi.org/10.1101/gr.100677.109

Zhang J, Zhang S, Han S, Wu T, Li X, Li W, Qi L (2012) Genome-wide identification of microRNAs in larch and stage-specific modulation of 11 conserved microRNAs and their targets during somatic embryogenesis. Planta 236:647–657. https://doi.org/10.1007/s00425-012-1643-9

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.