Cone-setting in spruce is regulated by conserved elements of the age-dependent flowering pathway

Shirin Akhter1*, Karl Johan Westrin2*, Nathan Zivi1,3*, Veronika Nordal1, Warren W. Kretzschmar2, Nicolas Delhomme4, Nathaniel R. Street5, Olof Emanuelsson2, Ove Nilsson4, and Jens F. Sundström1

1Department of Plant Biology, Linnean Center for Plant Biology, Uppsala BioCentre, Swedish University of Agricultural Sciences (SLU), SE-750 07 Uppsala, Sweden; 2Science for Life Laboratory, Department of Gene Technology, KTH Royal Institute of Technology, SE-171 65 Solna, Sweden; 3Skogforsk, Uppsala Science Park, Uppsala SE-751 83, Sweden; 4Department of Forest Genetics and Plant Physiology, Umeå Plant Science Centre, Swedish University of Agricultural Sciences (SLU), SE-901 83 Umeå, Sweden; 5Department of Plant Physiology, Umeå Plant Science Centre, Umeå University, SE-901 87 Umeå, Sweden

Authors for correspondence:
Olof Emanuelsson
Email: olofem@kth.se
Jens F. Sundström
Email: jens.sundstrom@slu.se

Received: 28 May 2022
Accepted: 23 August 2022

doi: 10.1111/nph.18449

Key words: cone-setting, flowering, gymnosperm, Picea abies, reproductive development, SPL-gene family, transcriptome.

Summary

- Reproductive phase change is well characterized in angiosperm model species, but less studied in gymnosperms. We utilize the early cone-setting acrocona mutant to study reproductive phase change in the conifer Picea abies (Norway spruce), a gymnosperm. The acrocona mutant frequently initiates cone-like structures, called transition shoots, in positions where wild-type P. abies always produces vegetative shoots.
- We collect acrocona and wild-type samples, and RNA-sequence their messenger RNA (mRNA) and microRNA (miRNA) fractions. We establish gene expression patterns and then use allele-specific transcript assembly to identify mutations in acrocona. We genotype a segregating population of inbred acrocona trees.
- A member of the SQUAMOSA BINDING PROTEIN-LIKE (SPL) gene family, PaSPL1, is active in reproductive meristems, whereas two putative negative regulators of PaSPL1, miRNA156 and the conifer specific miRNA529, are upregulated in vegetative and transition shoot meristems. We identify a mutation in a putative miRNA156/529 binding site of the acrocona PaSPL1 allele and show that the mutation renders the acrocona allele tolerant to these miRNAs. We show co-segregation between the early cone-setting phenotype and trees homozygous for the acrocona mutation.
- In conclusion, we demonstrate evolutionary conservation of the age-dependent flowering pathway and involvement of this pathway in regulating reproductive phase change in the conifer P. abies.

Introduction

Molecular clock-based studies, calibrated using fossil data, suggest that the gymnosperm and angiosperm lineages of extant seed plants separated c. 300 million years ago (Smith et al., 2010). Although the lineages share a common feature in the seed, their seed-bearing structures differ. Gymnosperms form seed- and pollen-bearing structures from separate shoot meristems, commonly referred to as cones (Florin, 1951) whereas angiosperms, in their ancestral state, are bisexual (Sauquet et al., 2017). Cones and flowers also differ in their branching order (Florin, 1951), where at least seed cones can be viewed as reproductive shoots analogous to angiosperm inflorescences, rather than flowers.

Comparative studies indicate that the genetic mechanisms that determine male or female organ identity are conserved between the two lineages (Rutledge et al., 1998; Tandre et al., 1998; Mouradov et al., 1999; Sundstrom et al., 1999; Winter et al., 1999). However, it is currently disputed if the mechanisms that regulate the on-set of cone-setting in gymnosperms and flowering in angiosperms are homologous (Karlgren et al., 2011; Klintenas et al., 2012; Liu et al., 2016).

Angiosperm flowering is regulated by several independent pathways that act in parallel, and converge on common floral integrators (O’Maoileidigh et al., 2014). The pathways are often referred to as the Age-dependent pathway, the Day-Length pathway, the Hormonal pathway, and the Vernalization pathway (Blazquez & Wei gel, 2000). The transition from vegetative growth to flowering occurs once in annual plants but can occur repeatedly in perennials (Albani & Coupland, 2010). The repeated flowering of the perennial herb Arabis alpina can be explained by the regulation of transcription factor proteins belonging to the SQUAMOSA BINDING PROTEIN-LIKE (SPL) family (Hyun et al., 2019). SQUAMOSA BINDING PROTEIN-LIKE proteins act as activators of flowering through the regulation of flower meristem identity genes (Wang

*These authors contributed equally to this work.
already in their second growth period (Uddenberg et al., 2013). After the first cone-setting, acrocona trees also initiate cones frequently, almost every year. This frequent cone-setting phenotype is semi-dominant and can to a degree also be observed in adult heterozygous acrocona mutants. In addition, heterozygous acrocona mutants commonly form cone-like structures, called transition shoots, on leading vegetative branches (Carlsbecker et al., 2013; Uddenberg et al., 2013).

Massively parallel DNA sequencing has been employed to study different aspects of reproductive development in conifers by us (Uddenberg et al., 2013; Giacomello et al., 2017) and others (Niu et al., 2014, 2016; Futamura et al., 2019). Previously, we have studied inbred siblings of young acrocona trees (Uddenberg et al., 2013) and identified the MADS-box gene DEFICIENS AGAMOUS LIKE 19 (DAL19) as being upregulated in needle samples of early cone-setting shoots. Later we have shown that distinct DAL19 isoforms are expressed in male and female cones, and in vegetative shoots (Akhter et al., 2018).

In the present study, we take advantage of the transition shoots and the numerous female cones that regularly form on adult heterozygous acrocona trees, and during cone-years also in the upper one-third of adult wild-type P. abies trees. We use massively parallel DNA sequencing to analyse both the messenger RNA (mRNA) and microRNA (miRNA) fractions of early meristems and transition shoot primordia from acrocona and compare those to corresponding samples from wild-type vegetative shoots and female cones. We hypothesize that candidate genes active in these early meristems are important not only for the acrocona phenotype but also for the regulation of cone-setting in wild-type P. abies. In line with this hypothesis, we identify candidates for a cone-setting regulatory circuit consisting of a conifer SPL-gene family member and two miRNAs. Furthermore, by genotyping a segregating sibling population of inbred acrocona trees, we provide evidence for a functional link between a mutation in a candidate gene, PaSPL1, and the early cone-setting acrocona phenotype.

Materials and Methods

Plant materials and morphological conditions

Plant material was collected from an acrocona tree located in Uppsala, Sweden and from a wild-type Norway spruce (P. abies (L.) H. Karst.) at the Rörby seed orchard (latitude 59°54′290′′N) near Uppsala, Sweden. Both trees were estimated to be at least 50 yr. Samples representing two developmental stages were collected from both genotypes. In the first developmental stage, the samples consisted of meristematic tissue. Samples in the second developmental stage harboured bud primordia with differentiating lateral organs. The acrocona samples consisted of transition shoots collected from apical positions on leading branches and female cones collected from apical positions on lateral branches. The acrocona samples used in RNA-sequencing (RNA-Seq) experiments were collected at two dates in 2016, 1 August and 18 October. Whereas the corresponding wild-type samples consisted of vegetative shoots collected from apical positions on leading branches and female cones collected from apical positions on lateral branches. Wild-type samples were collected in 2016, on 1 August, 16 September, and 25 October. Independent control samples of female cones and vegetative shoots were also collected from four additional wild-type genotypes on 8 October 2013. All plant materials used for RNA preparations were snap-frozen in liquid nitrogen and stored at −70°C. For a summary of the sample information and a detailed description of the sampling procedure, see Table S1.

RNA preparation

Tissue homogenization, extraction, CHISAM (chloroform/isoamylalcohol, 24:1) purification and isopropanol precipitation were performed as described by Azavedo et al. (2003). Harvested RNA pellets were dissolved in 350 μl RLT buffer (Qiagen RNeasy Kit; Qiagen, Carlsbad, CA, USA). Separate miRNA-enriched fractions (<200 nt) and total RNA fractions were purified from each RNA sample using the RNeasy MinElute Cleanup Kit (74204; Qiagen) following manufacturer’s instructions. RNA integrity was assessed via Bioanalyzer (Agilent Technologies, Santa Clara, CA, USA) and NanoDrop ND-1000 Spectrophotometer (ThermoFisher Scientific, Waltham, MA, USA). All RNA samples used for sequencing and subsequent molecular analyses had an RNA Integrity Number (RIN) between seven and nine.

Library preparation and RNA-sequencing (mRNA)

Sequencing libraries were prepared from 500 ng total RNA using the TruSeq stranded mRNA library preparation kit (RS-122-2101/2102; Illumina Inc., San Diego, CA, USA) including polyA selection. The library preparation was performed according to the manufacturer’s protocol (#5031047). A 2 × 125 bp short-read paired-end RNA-Seq of all bud samples was performed using a HiSeq2500 with v4-sequencing chemistry by The SNP & SEQ Technology Platform in Uppsala, Sweden.

Pre-processing of RNA-sequencing data: quality control, gene quantification (mRNA)

The data pre-processing was performed following the guidelines described in http://www.epigenesys.eu/en/protocols/bio-informatics/1283-guidelines-for-rna-seq-data-analysis. Briefly, the quality of
the raw sequence data was assessed using FASTQC (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/), v.0.11.4. Residual ribosomal RNA (rRNA) contamination was assessed and filtered using SortMrRNA (v.2.1; Kopylova & No, 2012). Data were then filtered to remove adapters and trimmed for quality using Trimomatic (v.0.36; Bolger et al., 2014). After both filtering steps, FASTQC was run again to ensure that no technical artefacts were introduced. Read counts were obtained using KALLISTO (v.0.43.0; Bray et al., 2016) using the *P. abies* v.1.0 complementary DNA (cDNA) sequences as a reference (retrieved from the PlantGenIE resource (Sundell et al., 2015)). The KALLISTO abundance values were imported into R (v.3.4.0; R_Core_Team, 2013) using the BIOCONDUCTOR module in GENEIOUS (v.3.2.6 (Huelsenbeck & Ronquist, 2001). Phylogenetic analysis

Annotated *SPL* genes from *P. abies*, *Arabidopsis thaliana* and *Populus trichocarpa* were included in the analysis. For each gene, the coding sequences were translationally aligned using the MAFFT module in GENEIOUS (GENEIOUS v.10.2.3; Biomatters Ltd, Auckland, New Zealand) and the resulting alignments were curated using the BMGE software with default settings (Criscuolo & Gribaldo, 2010). Phylogenetic analysis was carried out using MrBayes v.3.2.6 (Huelsenbeck & Ronquist, 2001).

Library preparation and RNA-sequencing (miRNA)

Sequencing libraries were prepared from the fraction of small (<200 nt) RNAs resulting from the RNA preparation using the TruSeq small RNA library preparation kit (RS-200-0012; Illumina Inc.) according to the manufacturer’s protocol. A 2 × 50 bp short-read paired-end RNA-Seq of all bud samples were performed using a NovaSeq SP-100 by the SNP & SEQ Technology Platform in Uppsala, Sweden.

Pre-processing of RNA-sequencing data: quality control, gene quantification (miRNA)

Small RNA fraction RNA-Seq reads were pre-processed by means of quality pruning and adapter trimming using FASTP (Chen et al., 2018) with the default settings, resulting in a set of high-quality reads. We used the miRNA database miRBase release 22.1 to match high-quality reads with known miRNAs (Kozomara et al., 2019) originating from *P. abies*, *A. thaliana* and *Populus trichocarpa*. We used KALLISTO to estimate miRNA abundance levels (both estimated counts and transcript per million (TPM)) by creating an index of 19-mers and transcript per million (TPM) by creating an index of 19-mers and 100 bootstrap samples during the actual quantification (Bray et al., 2016). We used sleuth to perform differential expression analysis of miRNAs (Pimentel et al., 2017). In sleuth, we used the likelihood ratio test (LRT) to detect differential expression and excluded miRNAs with a q-value larger than 0.05 from further analyses.

Transcriptome reconstruction

We used ClustRast (Westrin et al., 2022) with default settings (and using the built-in approach to generate the so-called guiding
contigs it requires) to generate a de novo assembly of the entire transcriptome. We included all the samples from 1 August in the assembly and aligned the de novo assembled transcripts to the *P. abies* reference genome (*P. abies* v.1.0) using MINIMAP2 (Li, 2018), with the preset option ‘splice:hq’.

Since several of the de novo assembled transcripts mapped sequentially to several ConGenIE scaffolds (MAs) we used this information to connect the exon sequences of genes that mapped to multiple scaffolds in the current genome assembly. Each reference sequence was counted for only once.

**Allele specific assembly**

To identify single nucleotide polymorphisms (SNPs) in candidate genes, we performed separate de novo assemblies using linked De Bruijn graphs (Turner et al., 2018) combined with kallisto (Bray et al., 2016), as outlined in Akhter et al. (2018). The method has since been named Abeona and is available at: https://github.com/winni2k/abeona.

**Genotyping**

Genomic DNA was extracted from tissue samples using a CTAB protocol (Kim et al., 1997). PCR-fragments covering *acrocona* specific SNPs were PCR-amplified using 100 ng of genomic DNA as template, Phusion DNA Polymerase (F-530; Thermo-Fisher Scientific) and primers listed in Table S2a. The resulting PCR-products were purified using the QIAquick PCR Purification Kit according to manufacturer’s instructions (28104; Qiagen) and sent to Eurofins Genomics for Sanger sequencing. The presence of a polymorphism was detected as double peaks in the resulting ab1-files (Fig. S1).

**Allele specific expression analysis**

The *PaSPL1* alleles were used to generate an index for kallisto, which we used to run each mRNA sample on. In the output, we could detect the allele frequencies from the TPM values, presented as average allele frequency across three biological samples.

**PaSPL1 transcript degradation estimated by 5’ RLM RACE**

RNA Ligase-Mediated 5’ Rapid Amplification of cDNA Ends (RLM RACE) was performed essentially as described by Llave et al. (2011). Samples used in the RLM RACE experiments are listed in Table S3. Briefly, RNA oligonucleotide adaptors were ligated to the 5’ terminus of cleaved transcripts using T4 RNA ligase (EL0021; ThermoFisher Scientific). The ligated RNA samples were subsequently reverse transcribed into first-strand cDNA using Superscript IV Reverse Transcriptase (18090010; Invitrogen). To amplify *PaSPL1* degradation products, we performed a primary touch-down PCR, followed by a nested secondary PCR using Phusion DNA Polymerase (F-530, ThermoFisher Scientific) and primers listed in Table S2c. The resulting PCR products were size separated on an agarose gel, cloned into Zero Blunt TOPO cloning vector (K2280010; Invitrogen) and transformed into chemically competent OneShot TOP10 *Escherichia coli* cells (C404010; Invitrogen). Transformed cells were pre-screened for the presence of the *PaSPL1* sequence using colony PCR, and selected clones were sent to Eurofins Genomics for Sanger sequencing (Sanget et al., 1977).

To quantify the *PaSPL1* degradation products, we performed qPCR experiments using Maxima SYBR Green qPCR Master Mix (KO221; ThermoFisher Scientific) and *PaSPL1* specific primer pairs (Table S2c). PCR fragments were quantified on a CFX Connect Real-Time PCR Detection System (Bio-Rad). The relative abundance of *PaSPL1* degradation products was normalized against the Cv value of the 5’-fragment in each sample, as outlined by Muller et al. (2002).

**Results**

**Apical buds on leading branches formed *acrocona* transition shoots**

In order to identify genes important to the *acrocona* transition shoot phenotype we sequenced the mRNA fraction of samples collected from leading shoots and lateral shoots (Fig. 1a) on branches situated in the cone-setting region of two *P. abies* genotypes: an *acrocona* mutant and a wild-type comparator. Lateral shoot meristems produced female cones in both wild-type *P. abies* and in the *acrocona* mutant (Fig. 1b,d) whereas leading shoot meristems produced vegetative shoots in the wild-type and transition shoots in the *acrocona* mutant (Fig. 1c,d).

We collected samples at the initiation of bud development and when the buds had started to differentiate. The buds harboured enlarged shoot apical meristems with only a few or no lateral organs at the early time-point. The later samples bore bud primordia with differentiating lateral organs while only a small meristem remained. Wild-type vegetative shoots initiated needles, whereas female shoots from both genotypes produced bracts and ovuliferous scales. In the *acrocona* transition shoots, needles had been formed in the basal part of the shoot, and bracts and ovuliferous scale-like structures had been formed in the apical part of the shoot (Fig. 1d).

Fourteen genes were commonly upregulated in *acrocona* transition shoots and female shoots.

A PCA was carried out to analyse the relationships between samples and mRNA transcription profiles (Fig. 2a). The first principal component (PC1) explained 48% of the total variation, and samples grouped according to collection dates along this axis (Fig. 2a). The second principal component (PC2) explained 15% of the variation in the samples. Notably, all *acrocona* samples grouped close to each other along this axis and were distinct from wild-type bud samples (Fig. 2a). Thus, the samples formed distinct groups and the grouping could be attributed to collection date (i.e. growth phase) and genotype.

We identified mRNA transcripts with a significant difference in expression levels in at least one of three sample types (*acrocona* transition shoots, *acrocona* and wild-type female shoots) as
compared to the base-line sample (wild-type vegetative leading shoots). We did this separately for both meristem (early) and primordia (late) samples. In total 8407 genes were significantly differentially expressed between buds collected in the early developmental stages (Fig. 2b). Among those, 515 genes were either upregulated or downregulated in all three comparisons (Figs 2b, S2; Dataset S1). Similarly, a total of 10 542 genes were significantly differentially expressed between bud samples collected in the late developmental stages (Fig. 2c). In those samples, 390 genes were differentially expressed in all three comparisons (Figs 2c, S3; Dataset S1).

In the differential expression analysis, we used the *P. abies* v.1.0 cDNAs as a reference (Sundell et al., 2015). In this assembly, known transcripts frequently map to several scaffolds due to assembly fragmentation. For example, the transcript of the MADS-box gene *DAL10* (Carlsbecker et al., 2003) (GenBank accession no. AF064080) maps in 5' to 3' direction to four ConGenIE scaffolds: MA_15122, MA_18073, MA_121040, and MA_86475g0010 (Sundell et al., 2015). To connect different ConGenIE scaffolds we performed a de novo transcriptome assembly using a novel assembly tool, CLUSTRAST (Westrin et al., 2022). We used the assembly to connect scaffolds that mapped to the same transcript. This reduced the list of differentially expressed transcripts to 461 in the meristem samples and to 352 in the primordia samples (Dataset S2a,b). Fifty-five genes were differentially expressed in both meristem and primordia samples. Out of these, 14 genes were upregulated in *acrocona* transition shoots, female *acrocona*, and female wild-type shoots, as compared wild-type vegetative leading shoots (Table 1). We verified the upregulation of these 14 genes in female cones from four additional wild-type genotypes (Fig. S4).

Three transcription factors were commonly upregulated in *acrocona* transition shoots and female shoots

We were primarily interested in transcription factors that may influence the shift from vegetative to reproductive shoot identity. The most significantly differentially expressed candidate among the 14 upregulated genes, MA_15381g0010, encoded a...
transcription factor belonging to the SPL-gene family (Table 1). Notably, in our de novo transcriptome-wide assembly we were able to connect MA_15381g0010 (on ConGenIE scaffold MA_15381) and the second most significant gene, MA_22749g0010 (on ConGenIE scaffold MA_22749), into a single transcript (ClusTrast ID 9986_s_0_0) suggesting that...
they are in fact part of the same gene. We named this gene *P. abies* SQUAMOSA BINDING PROTEIN-LIKE1 (*PaSPL1*). MA_15381g0010 harbours the signature domain of the SPL-gene family (Pfam domain PF03110) and MA_22749g0010 harbours a conserved *miRNA156* binding site commonly found in SPL-genes, as well as a binding site for *miRNA529* (Fig. 3a). To find independent proof of the connection between MA_15381g0010 and MA_22749 g0010 we PCR-amplified and Sanger-sequenced the corresponding full-length cDNA clone (Fig. S5). Next, we confirmed the upregulation of *PaSPL1* in female shoots and acrocona transition shoots as compared to wild-type vegetative shoots using independent RT-qPCR experiments (Fig. 3b). Among the upregulated genes we detected two additional transcription factors that both belong to gene families important for flowering and floral meristem identity in angiosperms: (1) The MADS-box gene *DAL10* (MA_86473g0010) (Fig. 3c; Table 1), suggested to be a marker for reproductive shoot identity in *P. abies* (Carlsbecker et al., 2003), and (2) a previously uncharacterized FLOWERING LOCUS T-LIKE gene (MA_194736g0010) belonging to the PEBP-family (Karlsgren et al., 2011; Klintenas et al., 2012; Liu et al., 2016) of transcription factors (Table 1).

*PaSPL1* is homologous to angiosperm SPL-genes involved in reproductive phase change

Several publications have reported phylogenetic reconstructions of the MADS-box gene family (e.g. Carlsbecker et al., 2003; Gramzow et al., 2014; Akhter et al., 2018). In those analyses, the *DAL10* gene commonly grouped into a gymnosperm specific sub-clade, which appears to be lost in the angiosperm lineage. In order to analyse the evolutionary relationship between conifer and angiosperm SPL-genes, we used the conserved SBP-domain of *PaSPL1* as bait to search for additional members of this gene family in the *P. abies* genome v.1.0 (Sundell et al., 2015). In total we retrieved 10 additional members of the SPL-gene family from *P. abies*, here named *PaSPL2-11*. Among those, *PaSPL1*, *PaSPL2*, *PaSPL10* and *PaSPL11* harbour the conserved *miR156/529* binding site (Table S4). The evolutionary relationship between the *P. abies* SPL-genes and genes from the model species *Arabidopsis thaliana* and *Populus trichocarpa* was analysed using Bayesian phylogenetics (Fig. S6). In this phylogeny, *PaSPL1*, *PaSPL10* and *PaSPL11* formed a clade that grouped basal to the *Arabidopsis thaliana* genes *AtSPL2*, *AtSPL6*, *AtSPL9*, *AtSPL10*, *AtSPL11* and *AtSPL15*, which all contain the *miR156* binding site, and

**Fig. 3** Verification of expression of *PaSPL1* and *DAL10* using reverse transcription quantitative polymerase chain reaction (RT-qPCR). (a) Graphical representation of the *PaSPL1* transcript assembled by ConGenIE with the coverage of the ConGenIE scaffolds MA_15381 and MA_22749 indicated in blue and green colour, respectively. Indicated are also the positions of the two introns present in the *PaSPL1* open reading frame, the signature SBP-domain and the overlapping binding sites of *miR156* and *miR529*. The boxplots in (b) and (c) show the normalized expression of *PaSPL1* and *DAL10* assayed by RT-qPCR. Veg, vegetative; F, female; TS, transition shoot; wt, wild-type; ac, acrocona. Box-plot elements: Line, median; box limits, upper and lower quartiles; whiskers, points.
have implicated roles in reproductive phase change (Preston & Hileman, 2013). The other *P. abies* SPL-genes included in the analysis grouped with other *A. thaliana* and *Populus trichocarpa* genes, e.g. PaSPL3, PaSPL4 and PaSPL5 grouped together with AtSPL8 and PtSPL8 (Fig. S6).

**MicroRNA156 and miR529 were upregulated in vegetative shoots and acrocona transition shoots**

Next, we analysed the expression of miRNAs in the meristematic samples (Table S1). Illumina sequencing reads of the small RNA fraction were mapped against previously known miRNAs present in miRBase (Kozomara et al., 2019). The read length of the RNA-Seq reactions allowed us to identify both precursor and mature miRNAs. We identified miRNAs with a significant difference in expression levels in at least one sample type (acrocona transition shoots, acrocona and wild-type female shoots) as compared to the same base-line sample as used in mRNA analysis (wild-type vegetative). Next, we performed hierarchical clustering of 1231 precursor and 966 mature miRNAs based on their estimated abundance levels (Figs S7, S8). Among the differentially expressed miRNAs, we identified **miR156t and miR529c**, which both have the capacity to bind the PaSPL1 mRNA in a partly overlapping manner (Fig. 3a). Both **miR156t and miR529c** were upregulated in wild-type vegetative shoots and acrocona transition shoots, compared to female shoots from both genotypes (Fig. S9a,b; *P < 0.001*; Dataset S3). The estimated expression levels of **miR156t** and **miR529c** were negatively correlated to the expression level of PaSPL1 in wild-type *P. abies*, whereas PaSPL1 and the miRNAs **miR156t** and **miR529c** were all upregulated in acrocona transition shoot meristems. A similar expression pattern of **miR156t** and **miR529c** was also detected in the late primordia samples (Fig. S10).

Apart from **miR156t** we also detected expression of other miRNAs that have implicated roles in the regulation of flowering or floral patterning in angiosperms (Spanudakis & Jackson, 2014), e.g. **miR159, miR172, miR167, miR319** and **miR390** (Figs S7, S8). Among those, **miR172** stood out as it, like **miR156t**, had a clear differential expression pattern between samples of different bud-types. In contrast to **miR156t,** **miR172i** was upregulated in female shoot meristems instead of wild-type vegetative meristems (Fig. S9c). In acrocona leading shoot meristems, **miR172i** had a more variable expression.

**The acrocona mutant harbours a SNP in the miR156/529 binding-site of PaSPL1**

A point mutation or a SNP in the **miR156/529 binding-site** could explain the simultaneous expression of PaSPL1, and the miRNAs **miR156 and miR529**, in acrocona transition shoots. To address this notion, we performed a separate **de novo** assembly of PaSPL1 transcripts from either wild-type samples, or acrocona samples using Abeona assembly. This method can be used to identify alleles and individual SNPs in short read transcriptome datasets (Akhter et al., 2018). As a reference, we performed similar allele-specific assemblies of the remaining candidate genes listed in Table 1 and four additional genes that also harbour the **miR156/529** binding-site. Thirteen of the assembled genes had SNPs in the **acrocona** mutant background (Table S5; Dataset S4). Next, we compared SNPs present in the assembled acrocona-transcripts with SNPs present in the Swedish breeding population of *P. abies* (Wang et al., 2020). In this comparison, only four genes had **acrocona** specific SNPs, and among those, only PaSPL1 also had a **miR156/529** binding site (Table S5). PaSPL1 had two **acrocona** specific SNPs, one located 256 nucleotides from the assumed start codon, and a second in the **miR156/529** binding-site at nt 1421 (Fig. 4a). Sanger sequencing of full-length cDNA
clones derived from the acrocona and the wild-type comparator confirmed the presence of this acro-SNP at nucleotide 1421.

Expression of miR156 and miR529 affected the allelic proportion of expressed PaSPL1
To test if the expression of miR156 or miR529 affects PaSPL1 expression levels in an allele-specific manner, we estimated the proportion mRNA expressed from each allele in the transcriptome datasets (Fig. 4b; Dataset S2c). We also assessed miR156 and miR529 levels in the same samples (Fig. 4c; Dataset S2d). Wild-type female meristems and primordia expressed low levels of miR156 and miR529, and the two PaSPL1 alleles were expressed in equal proportions (Fig. 4b,c, left column). In acrocona female shoots, the PaSPL1 allele proportion (acrocona vs wild-type) was 70 : 30 in the early meristematic samples and 50 : 50 in the later primordia samples (Fig. 4b, middle column). In the meristem samples, we observed elevated expression of an additional miR529 variant, miR529e, whereas primordia samples expressed low levels of all variants of miR156 and miR529 (Fig. 4c, middle column). Similar to the acrocona females, the allele proportion in meristematic samples from transition shoots was about 70 : 30. In the later transition shoot samples, the acro-allele accounted for almost 95% of all PaSPL1 transcripts (Fig. 4b, right column). As noted in our differential expression analysis of miRNAs (Fig. S9), acrocona transition shoots expressed elevated levels of both miR156 and miR529; and in the early meristem samples, miR529e (Fig. 4c, right column). In short, in samples that expressed elevated levels of miR156 or miR529, the allele proportion of PaSPL1 was affected in favour of the allele which harboured the acro-SNP. This pattern was evident in the early meristematic samples and became even more pronounced in the later primordia samples.

MicroR156/529 cleave PaSPL1 in an allele specific manner
The allele-specific reduction of PaSPL1 transcripts in acrocona transition shoots could be explained by miR156/529 mediated transcript cleavage. To examine this possibility, we performed 5′ RLM RACE experiments. Putative PaSPL1 cleavage- and degradation products of the expected size (c. 450 bp) could be readily observed in wild-type vegetative samples and, to some extent, in samples from acrocona transition shoots (Fig. S11a). Quantification of short degradation products that end at, or downstream of, the miR156/529 binding-site compared to longer general degradation products that span the entire miR156/529 binding-site reflects this pattern (Fig. S11b). We identified two putative PaSPL1 cleavage products from the wild-type vegetative samples by cloning and Sanger sequencing gel-purified DNA-fragments of the expected size. Seventeen out of 24 cloned fragments ended at nt 1427, i.e. within the miR156/529 binding site, and four fragments ended at nt 1431 (Fig. S11c). By cloning and Sanger sequencing of gel-purified fragments from the acrocona transition shoot, we detected two longer degradation products, three additional putative cleavage products that ended at nt 1427, and several shorter PaSPL1 fragments (Table S6). Furthermore, SNPs present in the cloned DNA fragments showed that the long uncleaved degradation products were expressed from the acrocona allele. In contrast, the putative cleavage products that ended at nt 1427, and the shorter fragments, were all from the wild-type allele. This indicates that expression of miR156 and miR529 preferentially mediate a cleavage of the wild-type allele of PaSPL1, and that this cleavage could explain the differences in the expression levels of the two alleles in acrocona transition shoots.

The acro-SNP co-segregated with the early cone-setting acrocona phenotype
We have previously performed inbred crosses of adult ramets of the acrocona mutant (Uddenberg et al., 2013). One-fourth of the resulting siblings displayed an enhanced early cone-setting phenotype and produced cones during the third growth cycle. As these trees have grown older, they now form rounded bushes with no clear apical dominance and regularly produce transition shoots on almost every shoot (Fig. S12). We expect that the causal mutation for the acrocona phenotype should be homozygous for the acro-allele in the early cone-setting siblings. To test if any of our candidate genes met this criterion, we analysed the segregation pattern of the acro-specific SNPs identified in the genes PaSPL1, MA_381942g0010, MA_10436587g0010 and MA_21055g010 in a sub-set of trees from the sibling population (Table S7). In this analysis only PaSPL1 was homozygous for its acro-specific SNP in early cone-setting trees.

To provide further support for this segregation pattern, we genotyped the entire segregating inbred sibling population generated in Uddenberg et al. (2013) (Table 2; Dataset S2e). In this analysis, 32% of the segregating sibling trees were homozygous for the acro-SNP present in PaSPL1, 57% were heterozygous, and 11% homozygous wild-type. Among the sibling trees that were homozygous for the acro-SNP, 92% displayed either an early cone-setting (21/24) or an intermediate acrocona (1/24) phenotype. Only two homozygous trees produced vegetative topshoots, and both of those trees had stunted growth. All heterozygous trees (42/42) produced vegetative shoots only. Similarly, none of the trees that were homozygous wild-type had any cones. Hence, we detected a highly significant (P<0.00001, Fisher’s exact test) correlation between trees homozygous for the acro-SNP and the early cone-setting phenotype.

We also genotyped wild stands of the acrocona mutant which all displayed a semidominant phenotype (Fig. S13). All trees were heterozygous with respect to the acro-SNP (Table S8; Dataset S2f), whereas two of the trees were also homozygous for the upstream acrocona specific SNP at nucleotide 256 (Dataset S2f). This

<table>
<thead>
<tr>
<th>Genotype/phenotype</th>
<th>WT (G/G)</th>
<th>Het (G/A)</th>
<th>acro (AA)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Apical cone</td>
<td>0</td>
<td>0</td>
<td>21 (88%)</td>
</tr>
<tr>
<td>Intermediate</td>
<td>0</td>
<td>0</td>
<td>1 (4%)</td>
</tr>
<tr>
<td>Vegetative</td>
<td>8 (100%)</td>
<td>43 (100%)</td>
<td>2 (8%)</td>
</tr>
<tr>
<td>Total</td>
<td>8</td>
<td>43</td>
<td>24</td>
</tr>
</tbody>
</table>
suggests that the upstream SNP is not necessary for the enhanced phenotype displayed by homozygous acrocona plants.

Discussion

In most of Sweden’s planting zones, there is a shortage of domestically produced P. abies seeds (Rosvall, 2011). This shortage has two primary causes: irregular cone-setting of P. abies and damages to cones and seeds caused by insects and fungi (Almqvist et al., 2010). Conifer breeding and research also face a significant obstacle because of the very long generation times (Flachowsky, 2010). Conifer breeding and research also face a significant obstacle because of the very long generation times (Flachowsky, 2010). Therefore, there is a strong desire to learn more about the molecular mechanism that regulates cone-setting in conifers (Uddenberg et al., 2015). Understanding the genetic mechanism that regulates reproductive phase change in conifers could also increase our understanding of the evolutionary relationship between extant seed plants, i.e. angiosperms and gymnosperms.

In this study, we utilized the unique features of an adult, naturally occurring and presumably heterozygous, acrocona mutant. In this acrocona mutant, apical shoots on leading branches commonly develop into transition shoots. Like vegetative shoots, the first lateral organs that initiate in acrocona transition shoots are needles. Later in the growing season, the acrocona transition shoots produce ovuliferous scale-like structures. Hence, we collected samples that allowed us to identify genes expressed in the acrocona transition shoots before any morphological signs of the reproductive shift were apparent, and we compared their transcriptome profiles to profiles generated from corresponding wild-type vegetative shoot meristems and female meristems. This selection of shoots allowed us to address the hypothesis that the acrocona transition shoots express genes related to reproductive shoot development before the morphological shift. It was also likely that the identified candidate genes would be active in the meristem rather than acting in the down-stream morphological differentiation. Genes upregulated in both transition shoots and female meristems of acrocona (relative to wild-type vegetative meristems) would therefore be candidates for genes important for reproductive meristem identity.

Apart from the meristematic samples collected in early August, we also collected primordia samples during the autumn, when lateral organ differentiation occurs. In this growth phase, vascular strands connect to the lateral organs and cellular differentiation occurs within the ovuliferous scales and sterile bracts in female cones, and within the needles in vegetative shoots. Similar cellular differentiation also occurs in the acrocona transition shoots.

By combining the results from comparisons of meristem and primordia samples, we identified 14 genes that were upregulated in acrocona transition shoots and female cones as compared to wild-type vegetative leading shoots. In line with the hypothesis that these 14 genes represent genes important for reproductive development, we identified DAL10, a marker for reproductive shoot identity (Carlsbecker et al., 2003). Among the top candidate genes, we also identified a member of the SPL gene family, here named PaSPL1. In our phylogenetic analysis of the SPL gene family, PaSPL1 groups together with angiosperm SPL-genes that have been shown to regulate flowering (Preston & Hileman, 2013). This is interesting because a certain position in a phylogenetic tree may be indicative not only of shared ancestry, but also of conserved function between closely located genes (Theissen et al., 1996; Tandré et al., 1998). Although sub-functionalization and neo-functionalization frequently occur (Irish & Litt, 2005).

In angiosperms, members of the SPL-gene family are key regulators of the age-dependent flowering pathway (Wang et al., 2009; Preston & Hileman, 2013). This pathway also includes miR156, which acts as a negative regulator of the SPL-genes during juvenile stages and in vegetative meristems. Expression of miR156-resistant variants of AaSPL15 in the perennial herb A. alpina is known to induce early flowering and flowering in positions which in wild-type would continue as vegetative shoots (Hyun et al., 2019). We note that this resembles the acrocona mutant phenotype. Analysis of the PaSPL1 sequence revealed that PaSPL1 harbours a conserved miR156 binding site located 1421 nucleotides downstream from the start codon. Partially overlapping was also the binding site of miR529, a miRNA that has been lost in the core eudicots but that is still present in, e.g. bryo-phytes and monocots such as Oryza sativa (Morea et al., 2016). The occurrence of an overlapping binding site indicates that both miRNAs may negatively regulate PaSPL1. In line with this hypothesis, we detected an elevated expression of both miR156 and miR529 in wild-type vegetative leading meristem compared with female meristems of both assayed genotypes. This supports the hypothesis that the SPL/miR156 module of the age-dependent flowering pathway regulates reproductive phase change in conifers – possibly with the additional involvement of miR529.

Interestingly, both miR156 and miR529 were co-expressed with PaSPL1 in acrocona transition shoot meristems. The SNP present in the overlapping miR156/529 binding site of the PaSPL1 acrocona allele could explain the co-expression: we detected (1) two distinct cleavage products of PaSPL1 in samples from vegetative shoots cleaved in the putative miR156/529 target site, (2) specific cleavage of the wild-type allele of PaSPL1 in heterozygous acrocona transition shoots, and (3) a higher expression of the acrocona allele (as compared to the wild-type allele) in acrocona transition shoots. This indicates that miR156 and/or miR529 can mediate PaSPL1 cleavage, and that this cleave occurs in an allele specific manner, suggesting that the acro-SNP renders the acrocona allele miR156/529 tolerant. We note that similar dual cleavage products of SPL transcripts have been reported previously in heterologous experiments studying the ectopic expression of miR156 and miR529 from O. sativa in A. thaliana (Morea et al., 2016).

Provided that PaSPL1 regulates female reproductive identity, we would expect a co-segregation of this SNP with the acrocona phenotype. Indeed, the acro-SNP is absent in a tested set of 35 wild-type genotypes that are part of the Swedish breeding population (Wang et al., 2020). In our previous studies, we have produced an inbred population of the acrocona mutant. One quarter segregated with an early cone-setting phenotype among the sibling trees, which we then interpreted as an enhanced homozygous phenotype (Uddenberg et al., 2013). As these plants have grown...
older, they only became rounded shrubs, distinct from the heterozygous *acrocona* trees, which displayed a semi-dominant phenotype and grew taller. We have now demonstrated a co-segregation between the early cone-setting phenotype and trees that are homozygous for the *acro*-SNP by genotyping. Importantly, none of the segregating siblings that were homozygous for the wild-type allele displayed any *acrocona* phenotypes.

In conclusion, we propose that cone-setting in the conifer *P. abies* is regulated by conserved elements of the age-dependent flowering pathway. In support of this notion, we provide several independent lines of experimental evidence: (1) Using transcriptome analyses, we demonstrate an anti-correlated expression of *PaSPL1* and *miR156/529* in female and vegetative shoot meristems. (2) Using allele-specific assembly and expression analysis, we identify an *acrocona* specific SNP in the miRNA binding site of *PaSPL1*. We show that the *acrocona* allele of *PaSPL1* is upregulated in transition shoots, along with *miR156* and *miR529* in contrast to the anti-correlated expression in wild-type shoots. (3) Using RLM RACE experiments, we show that *miR156* and *miR529* preferentially mediate cleavage of the wild-type allele of *PaSPL1*. (4) Finally, we demonstrate that among our *acrocona* specific SNPs – in *PaSPL1* and other candidate genes – only the *acro*-SNP in the miRNA binding site of *PaSPL1* co-segregates with the enhanced *acrocona* phenotype. We have, however, not analysed the genomic sequence of *PaSPL1*, and it is possible that other unknown SNPs or perhaps the SNP located at nucleotide 256 in the *PaSPL1* transcript also contribute to the *acrocona* phenotype. Considering these numerous lines of evidence, together with the similarity to the situation in the angiosperm perennial herb *Arabis alpina*, the most parsimonious conclusion is that the early flowering of the *acrocona* mutant is caused by the mutation in the *miR156/miR529* binding site of the *PaSPL1* gene. However, to ultimately prove that the *acro*-SNP alone is responsible for the *acrocona* phenotype we would be required to find, or generate, independent mutations in the *PaSPL1* locus. Our results demonstrate remarkable conservation of this pathway, which is linked to perennity, between species that shared a last common ancestor 300 million years ago. Hence, the age-dependent pathway seems to be crucial to the regulation of reproductive phase change not only in conifers, but also in many other perennial seed plants.

Acknowledgements

The authors acknowledge funding for salaries and consumables from the Carl Tryggers Foundation (grant number CTS 18:367), the Royal Physiographic Society (22309-000), the Swedish Foundation for Strategic Research (FID16-0030 and FFF20-0014), Formas (239-2013-650 and 2021 01026), and the Swedish Foundation for Strategic Research (FID16-0030 and 0807). The computations were enabled by resources in project SNIC 2017/7-303 provided by the Swedish National Infrastructure for Computing (SNIC) at UPPMAX and project PDC-2020-43 at Parallelldatorcentrum (PDC), partially funded by the Swedish Research Council (grant no. 2018-05973).

Author contributions

JFS and OE planned and designed the research. JFS wrote the manuscript with contributions from SA, KJW and NZ. SA, NZ and VN performed the experiments and conducted fieldwork. SA, KJW, WWK and ND analysed the data. JFS, OE, NRS and ON provided supervision, funding and materials. All authors read and edited the final version of the manuscript. SA, KJW and NZ contributed equally to this work. JFS and OE are the joint corresponding authors on the manuscript.

ORCID

Shirin Akhter https://orcid.org/0000-0003-4238-3110
Nicolas Delhomme https://orcid.org/0000-0002-3053-0796
Olof Emanuelsson https://orcid.org/0000-0002-8879-9245
Warren W. Kretzschmar https://orcid.org/0000-0002-2575-0807
Ove Nilsson https://orcid.org/0000-0002-1033-1909
Nathaniel R. Street https://orcid.org/0000-0001-6031-005X
Jens F. Sundström https://orcid.org/0000-0003-2848-5284
Karl Johan Westrin https://orcid.org/0000-0002-6937-9245

Data availability

All data generated or analysed during this study are included in this published article or its Supporting Information. The sequencing data is available at the European Nucleotide Archive (ENA, https://www.ebi.ac.uk/ena/browser/home) under the accession no. PRJEB45942. All custom-made code is available at either https://github.com/karljohanw/clustarr or https://github.com/winni2k/abeona.

References


Supporting Information

Additional Supporting Information may be found online in the Supporting Information section at the end of the article.

**Dataset S1** Differentially expressed genes.

**Dataset S2** Background data.

**Dataset S3** MicroRNA statistics.

**Dataset S4** Consensus acrocona transcripts.

**Fig. S1** Genotyping of the acro-single nucleotide polymorphism in PaSPL1.

**Fig. S2** Hierarchical clustering of differentially expressed genes in early meristematic samples.

**Fig. S3** Hierarchical clustering of differentially expressed genes in late primordia samples.

**Fig. S4** Expression of candidate genes in control samples.

**Fig. S5** Cloning and sequencing of the PaSPL1 coding sequence.

**Fig. S6** SQUAMOSA BINDING PROTEIN-LIKE (SPL) gene family phylogeny.

**Fig. S7** Hierarchical clustering of precursor microRNA expressed in meristem.

**Fig. S8** Hierarchical clustering of mature microRNAs expressed in meristem.

**Fig. S9** Expression of miR156t, miR529c, and miR172i.

**Fig. S10** Hierarchical clustering of mature microRNAs expressed in bud primordia.

**Fig. S11** PaSPL1 transcript cleavage estimated by 5' RNA Ligase-Mediated 5' Rapid Amplification of cDNA Ends (5' RLM RACE).

**Fig. S12** Phenotypes of inbred acrocona siblings after 13 growth-cycles.

**Fig. S13** Locations and phenotypes of adult stands of acrocona trees.

**Table S1** Samples subjected to RNA-sequencing.

**Table S2** Primers used in the study.

**Table S3** Samples used in 5' RNA Ligase-Mediated 5' Rapid Amplification of cDNA Ends (5' RLM RACE) experiments.

**Table S4** The PaSPL-gene family.

**Table S5** Acrocona specific single nucleotide polymorphisms.

**Table S6** Frequency of RNA Ligase-Mediated 5' Rapid Amplification of cDNA Ends (5' RLM RACE) products.

**Table S7** Genotyping of genes with acrocona specific single nucleotide polymorphisms.

**Table S8** Genotyping of adult stands of acrocona and wild-type Picea abies.

Please note: Wiley Blackwell are not responsible for the content or functionality of any Supporting Information supplied by the authors. Any queries (other than missing material) should be directed to the New Phytologist Central Office.