



# Global Analysis of Small RNA Dynamics during Seed Development of *Picea glauca* and *Arabidopsis thaliana* Populations Reveals Insights on their Evolutionary Trajectories

Yang Liu\* and Yousry A. El-Kassaby\*

Department of Forest and Conservation Sciences, University of British Columbia, Vancouver, BC, Canada

## OPEN ACCESS

### Edited by:

Mathew G. Lewsey,  
La Trobe University, Australia

### Reviewed by:

German Martinez,  
Swedish University of Agricultural  
Sciences, Sweden  
Reena Narsai,  
La Trobe University, Australia

### \*Correspondence:

Yang Liu  
yliu2011@interchange.ubc.ca  
Yousry A. El-Kassaby  
y.el-kassaby@ubc.ca

### Specialty section:

This article was submitted to  
Plant Genetics and Genomics,  
a section of the journal  
Frontiers in Plant Science

**Received:** 22 June 2017

**Accepted:** 20 September 2017

**Published:** 04 October 2017

### Citation:

Liu Y and El-Kassaby YA (2017)  
Global Analysis of Small RNA  
Dynamics during Seed Development  
of *Picea glauca* and *Arabidopsis*  
*thaliana* Populations Reveals Insights  
on their Evolutionary Trajectories.  
Front. Plant Sci. 8:1719.  
doi: 10.3389/fpls.2017.01719

While DNA methylation carries genetic signals and is instrumental in the evolution of organismal complexity, small RNAs (sRNAs), ~18–24 ribonucleotide (nt) sequences, are crucial mediators of methylation as well as gene silencing. However, scant study deals with sRNA evolution via featuring their expression dynamics coupled with species of different evolutionary time. Here we report an atlas of sRNAs and microRNAs (miRNAs, single-stranded sRNAs) produced over time at seed-set of two major spermatophytes represented by populations of *Picea glauca* and *Arabidopsis thaliana* with different seed-set duration. We applied diverse profiling methods to examine sRNA and miRNA features, including size distribution, sequence conservation and reproduction-specific regulation, as well as to predict their putative targets. The top 27 most abundant miRNAs were highly overlapped between the two species (e.g., miR166, –319 and –396), but in *P. glauca*, they were less abundant and significantly less correlated with seed-set phases. The most abundant sRNAs in libraries were deeply conserved miRNAs in the plant kingdom for *Arabidopsis* but long sRNAs (24-nt) for *P. glauca*. We also found significant difference in normalized expression between populations for population-specific sRNAs but not for lineage-specific ones. Moreover, lineage-specific sRNAs were enriched in the 21-nt size class. This pattern is consistent in both species and alludes to a specific type of sRNAs (e.g., miRNA, tasiRNA) being selected for. In addition, we deemed 24 and 9 sRNAs in *P. glauca* and *Arabidopsis*, respectively, as sRNA candidates targeting known adaptive genes. Temperature had significant influence on selected gene and miRNA expression at seed development in both species. This study increases our integrated understanding of sRNA evolution and its potential link to genomic architecture (e.g., sRNA derivation from genome and sRNA-mediated genomic events) and organismal complexity (e.g., association between different sRNA expression and their functionality).

**Keywords:** adaptive strategy, *Arabidopsis thaliana*, microRNA, organismal complexity, phenotypic variation, *Picea glauca*, seed ontogeny, small RNA evolution

## INTRODUCTION

Epigenetic mechanisms exert functional roles in the upper dimension of gene regulatory networks (GRNs) and are more important to the transcription machinery than have hitherto been thought. Because changes alone in epigenetics can affect complex traits across generations in the absence of genetic variation (Cubas et al., 1999; Johannes et al., 2009; Richards et al., 2012), where salient examples of the direct impact of epimutations (e.g., epQTLs, modified polymorphisms in nucleotides) on phenotypic variation are mainly instantiated in the profiling of epialleles (i.e., single-locus DNA methylation variants). While non-coding small RNA (sRNA) molecules are paramount to genomic DNA methylation via sequence-specific recognition to recruit regulatory proteins in the RNA-directed DNA methylation (RdDM) pathway (Lewsey et al., 2016), there are, supposedly, parent-offspring transmission patterns of methylation marks and sRNAs [e.g., small interfering RNAs (siRNAs)], as put forth by Diez et al. (2014). Moreover, the expression level of protein-encoding genes (PEGs) correlates with the density of their nearby methylated transposable elements (TEs) in accessible euchromatic regions, reviewed by Diez et al. (2014) and Sigman and Slotkin (2016). sRNAs particularly in the 24 nucleotide (nt) size class are crucial regulators of TEs (Almeida and Allshire, 2005; Nosaka et al., 2012) and have natural origins from TE segments (Piriyaopongsa and Jordan, 2008; Sun et al., 2012). It is therefore highly conceivable that there are tight associations between sRNAs (type and abundance) and gene expression regulation in shaping regulatory diversity and robustness.

Plant sRNAs are generated from stem-loop regions of longer primary transcripts (or fold-back structures from single- or double-stranded RNA precursors) by Dicer-like enzymes (DCLs) and chiefly comprise microRNAs [miRNAs; prevalence of 21- or 22-nt long in suppressing target mRNAs], heterochromatic siRNAs [hc-siRNAs; 24-nt mediators in silencing DNA methylation and histone modifications], and *trans*-acting siRNAs [tasiRNAs or phasiRNAs; 22 (or 21)-nt with a phased configuration, playing similar roles as miRNAs or other uncharacterized functions], reviewed by Wendel et al. (2016). As sRNAs form duplexes when derived or pairing with another sRNA or binding to mRNA to direct cleavage and degradation, the capability to pair with other genomic sequence(s) is their fundamentally common feature. At the levels of transcription (indirect and low) and post-transcription (major), miRNAs are well-established players in various developmental programs (Dugas and Bartel, 2004; Sparks et al., 2013) and plasticity (Rubio-Somoza and Weigel, 2011). In response to environmental cues, GRNs can be overridden via instigating the biogenesis of different miRNAs. In major land plant lineages from the backbone phylogenetic tree, acquisition of different miRNA families has been summarized using parsimony approaches and only a handful of distinct families of deeply conserved miRNAs are evolutionarily ancient and stable (Zhang et al., 2006; Axtell et al., 2007; Axtell and Bowman, 2008). Large diverse sets of lineage-specific miRNAs often exceed the conserved ones (Lindow and Krogh, 2005; Cuperus et al., 2011). Presumably, myriad *MIRNAs* (miRNA genes) are expanded but unrevealed

(Fattash et al., 2007; Nozawa et al., 2012), or young *MIRNAs* are spawned, weakly expressed and eventually frequently lost (Fahlgren et al., 2007). This indicates that lineage-specific miRNAs have undergone rapid turnover in evolution (e.g., 1.2–3.3 genes per Myr in *Arabidopsis*, Fahlgren et al., 2010) and *MIRNAs* evolve adaptively, possibly driven by positive selection. From the phylogenetic perspective, distribution of miRNA families seems to be approximately proportional to the antiquity of the evolutionary lineages (Taylor et al., 2014). Novel miRNAs may adjust the variance of expression levels of their target genes to maintain the stability of transcription networks and after diversifying and purifying selection, they reset mean gene expression to improve fitness of specific phenotypes (Wu et al., 2009). Co-option of ancient and young miRNA families to conduct new functions is important in the evolution of new phenotypes (Taylor et al., 2014) or phenotypic differentiation. In addition, duplication of *MIRNAs*, especially those from whole-genome duplication (WGD), may be also conducive to miRNA diversity and their regulatory complexity (Maher et al., 2006; Vanneste et al., 2014), indicative of coevolution of *MIRNAs*, miRNAs and their targets in the context of WGD events (or genome evolution).

In this study, we chose populations (ecotypes) within two spermatophytes with contrasting seed-set time span (i.e., *Picea glauca* and *Arabidopsis thaliana*), which implies different seed developmental modes possibly contributing to acclimation and adaptive differentiation. Conifers (Pinophyta or Coniferales) are masters of adaptation due to their long endurance over periods of climatic sways (Jaramillo-Correa et al., 2004; Anderson et al., 2006; Tollefsrud et al., 2008) and wide geographic distribution (Wang and Ran, 2014). Their long generation time confines the ability of populations to respond to environmental stresses through genetic mechanisms (Bräutigam et al., 2013), while epigenetics is more labile, malleable and potentially reversible, and thus more closely aligned with the environmental exposure to create new combinations of variants (Lira-Medeiros et al., 2010; Schulz et al., 2014). Extant conifers (Pennsylvanian, 318–299 Myr ago), bearing enormous genome size (e.g., 200 × *Arabidopsis*), are evolutionarily twice as old as angiosperms (early Cretaceous, 146 Myr ago) (Schneider et al., 2004) but with, on average, seven times lower in rates of molecular evolution (De La Torre et al., 2017), and have undergone few WGDs or polyploidization (Leitch and Leitch, 2012). The contrasting differences in genome architecture suggest disparate evolutionary mechanisms of genome and epigenetic control between conifers and *Arabidopsis* (angiosperms). Conifers have high levels of methylation in PEGs, at ~40% in vegetative tissues and over 75% in megagametophytes in *Pinus taeda* (Takuno et al., 2016) and are found to yield 24-nt sRNAs only at reproductive tissues (Nystedt et al., 2013), supportive to high levels of epigenetic modifications at conifer seed set. In *Arabidopsis thaliana*, specific heritable methylation patterns account for 60–90% of the heritability of two studied complex traits, flowering time and primary root length (Cortijo et al., 2014). Moreover, lines of evidence in *Picea abies* and *P. taeda* have shown that environmental conditions at seed set can substantially affect progeny performance (Johnsen et al., 2005; Kvaalen and Johnsen, 2008) and this process

is mediated by miRNAs (Oh et al., 2008; Yakovlev et al., 2010). Recently, it has been reported that environmental cues (e.g., temperature, CO<sub>2</sub>) alter the expression of miRNAs (e.g., miR156, -157, -160, -164, and -172) to regulate *Arabidopsis* development and growth (May et al., 2013). Together, sRNA dynamics encrypted at seed set may shed light on the plasticity potential in adaptation to habitat heterogeneity and plant life histories.

Local adaptation enables plants to obtain a high fitness and developmental transitions should be properly timed to coincide phases in plant growth with their favorable seasons. The seed is therefore a key evolutionary adaptation of seed plants that facilitates dispersal and reinitiates development only at suitable environmental conditions. Seeds have evolved the ability to start their life cycle at the right time via timing their germination. This timing in plant life history is controlled by seed dormancy (i.e., innate constraint on seed germination under conditions that would otherwise promote germination in non-dormant seeds), which is the target phenotype in this study. Plants use dormancy in seeds to move through time and space, which finally maximizes their fitness. Seed dormancy is under strong evolutionary selection, because improper timing of germination may lead to outright extinction (Huang et al., 2010). As post-zygotic quiescence may be the starting point of the evolution of seed dormancy (Mapes et al., 1989), dormancy modulation is assumed to be the consequence of finely tuned programs at seed set. We therefore hypothesize that selected populations can represent different modes of reproductive development that are regulated by sRNAs and exert cascading effects on ensuing phenotypes, including seed dormancy. We focus on miRNA-mRNA nodes responsible for seed dormancy formation (Liu and El-Kassaby, 2017a) and three key conserved genes [i.e., *ABA INSENSITIVE 3 (ABI3)*, *AUXIN RESPONSE FACTOR 10/16 (ARF10/16)*, and *DELAY OF GERMINATION 1 (DOG1)*]. *ABI3* is a conserved gene at embryogenesis (Fischerova et al., 2008) and, in association with *ARF10/16*, regulates seed dormancy (Liu et al., 2013), while *DOG1* is involved in dormancy cycling as a response to seasonal environmental signals (Vidigal et al., 2016) and subject to non-coding RNA-mediated mechanisms (Fedak et al., 2016).

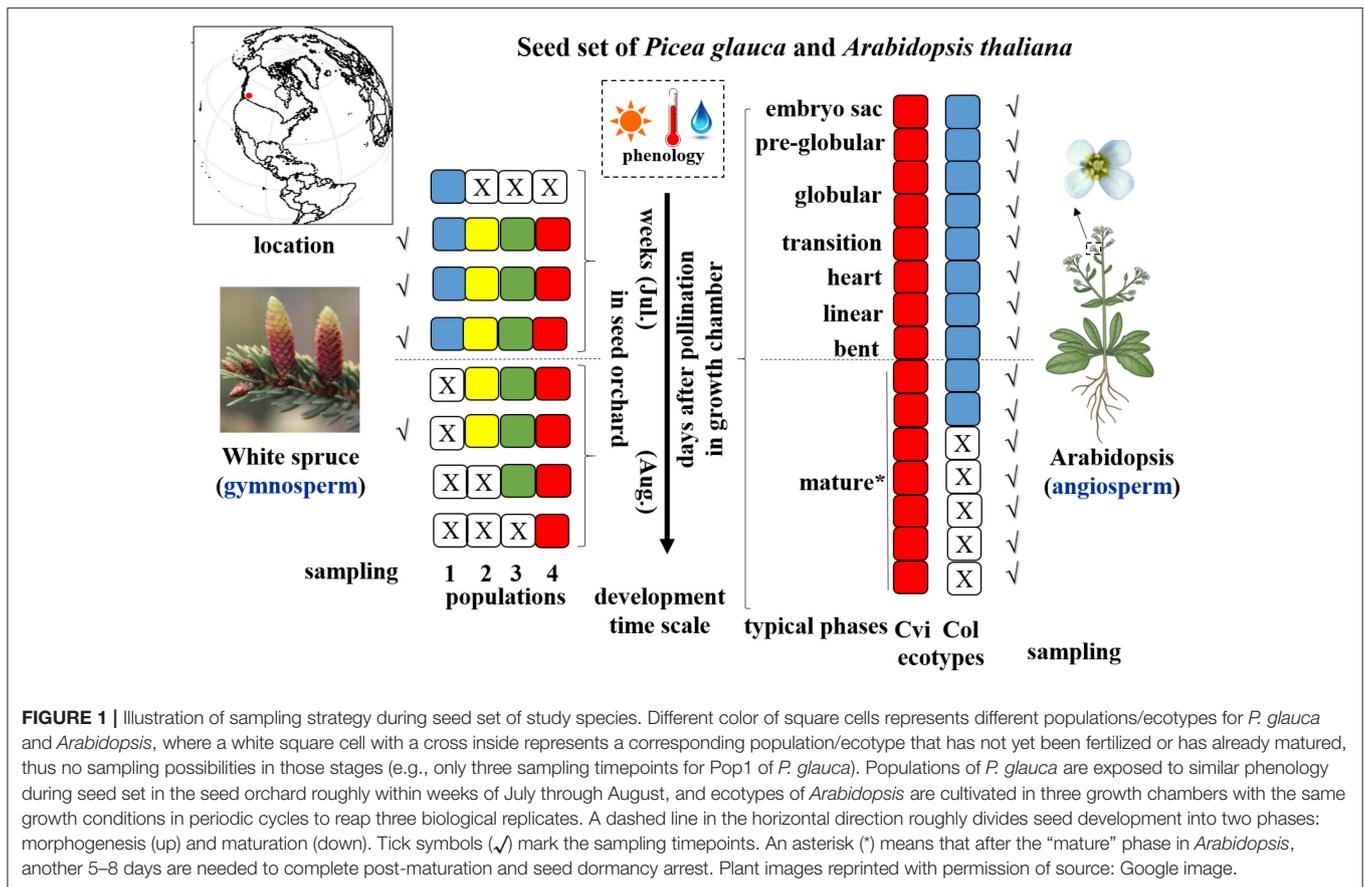
A growing body of knowledge reinforces the notion that sRNA dynamics at seed set help shape the capacity of phenotypic variation and local adaptation. Here, we employ an Illumina sequencing approach to identify and comparatively examine enriched sRNAs and their possible targets throughout seed-set phases among populations within and between *P. glauca* and *Arabidopsis*. Through this study, we intend to address three sets of compelling questions: (i) Are top highly expressed miRNAs overlapped in *P. glauca* and *Arabidopsis*? And are they also the deeply conserved ones throughout the plant kingdom? (ii) What are the expression dynamics for lineage- (i.e., sRNAs detected over time across chosen populations) and population-specific sRNAs (i.e., sRNAs detected over time but not across populations)? And are these patterns consistent in the two species? (iii) Are there putative sRNAs targeting known adaptive genes generated at seed development? And how is the role of miRNA-gene interactions at play for our study phenotype (i.e.,

seed dormancy)? As of December 2016, 494 unique miRNA entities in *Picea* were documented, of which 155 are known on miRBase and 21 are deeply conserved miRNA families in plants (Källman et al., 2013; Xia et al., 2015). Most embryogenesis-related genes in *Arabidopsis* have homologs, with high congruity, in conifers (Cairney and Pullman, 2007), such as in *P. taeda* (83%) (Cairney et al., 2006) and *Larix kaempferi* (78%) (Zhang et al., 2012). The transcriptomic profiling of the zygotic embryo in *P. pinaster* is highly correlated with that in *Arabidopsis* (Xiang et al., 2011; de Vega-Bartol et al., 2013). Their high similarities in gene homologs make possible and meaningful this comparative study between a conifer and *Arabidopsis*. This study provides us with important clues to further investigating how sRNAs and GRNs coevolve to influence adaptive capacity and organismal diversity.

## MATERIALS AND METHODS

### Plant Material, Growing Conditions, and Sample Collection

White spruce (*P. glauca*), a keystone species of boreal forests in the North American taiga, is anemophilous (outcrossing) and its seed and pollen cones develop on the same tree and are diclinous (i.e., unisexual). According to seed orchards station records, four populations of *P. glauca* (Pop 1~4), characterized by different pollination timing and seed developmental duration, were chosen and 20 developing cones for each population were collected at early, middle, and late developmental stages for a total of four timepoints at the Kalamalka Research Station seed orchards (50°–51°37'N, 119°16'–120°29'W), British Columbia, Canada (Figure 1). Note that population 1 only had three sampling timepoints due to its short seed developmental span. Accordingly, the climatic data were retrieved from the Station. Likewise, as per contrasting seed-set durations, we selected two wild strains of a model annual selfing plant *A. thaliana*, Cvi-0 and Col-0, originating from the Cape Verde Islands and Columbia (Missouri, USA), respectively. The two *Arabidopsis* ecotypes exhibit contrasting dormancy intensities (Koornneef et al., 2000; Ali-Rachedi et al., 2004). They were cultivated in 2-in planting trays containing soil mixed with slow-release fertilizer 14-14-14 in a growth chamber with 16/8 h day/night photoperiod, PPFD (photosynthetic photon flux density) of 250 μmol·m<sup>-2</sup>·s<sup>-1</sup>, and constant temperature of 22°C. Individual flowers were tagged on the day of flowering and developing seeds were sampled manually every day for 9 consecutive days after pollination (DAP) (Figure 1). Five more sampling time points (10–14 DAP) were performed for Cvi-0 due to its slow seed development (Figure 1). Developing seeds at each timepoint were sampled three times (i.e., three biological replicates). After extraction from white spruce cones, direct collection of siliques at *Arabidopsis* embryogenesis (0–6 DAP) and quick hand-dissection in water from ~30 inflorescences at maturation stages (7 DAP onward), entire developing seeds or siliques were immediately frozen in liquid nitrogen and stored at –80°C until further use. Note that the seed samples from the same developing timepoint were pooled in the same vial for the subsequent analyses.



## RNA Isolation, Library Construction, and sRNA Sequencing

Total RNAs were extracted and divided into two aliquots (~15 µg each) from developing seed samples of *P. glauca* and *Arabidopsis* using PureLink Plant RNA Reagent (Ambion) according to the manufacturer's instructions. The integrity and quantity of the RNA samples were assessed on a BioAnalyser 2100 (Agilent Technologies) and a Nanodrop ND-1000 spectrophotometer (Thermo Fisher Scientific). The sRNA-seq libraries were constructed using a strand-specific and plate-based protocol. To enrich sRNAs, total RNA samples underwent polyA selection using Miltenyi MultiMACS mRNA isolation kit (cat. 130-092-519) following the manufacturer's protocol and the flowthrough (i.e., containing sRNA species without mRNAs) was used for plate-based sRNA construction. A 3' adapter that is an adenylated single-stranded DNA was selectively ligated to the sRNA template using a truncated T4 RNA ligase 2 (NEB Canada, cat. M0242L). A 5' adapter was then added using a T4 RNA ligase (Ambion USA, cat. AM2141) and ATPs. After ligation, first strand cDNA was synthesized using a Superscript II Reverse Transcriptase (Invitrogen, cat. 18064 014) and one RT primer. This was the template for the final library PCR, into which 6-nt mers index was introduced to identify libraries (i.e., demultiplexed) from a sequenced pool. Constructed libraries were pooled by phylum; that is, 15 *P. glauca* and 25 *Arabidopsis* samples were pooled

separately, and both lanes were 31 base SET lanes. Sequencing (Illumina HiSeq™ 2500) was implemented using one short SET indexed lane per pool (BC Cancer Agency, Genome Sciences Centre, Vancouver, Canada).

## Small RNA Dataset Analysis

The sequence data were partitioned into individual libraries based on the index read sequences, and the reads underwent an initial QC assessment. After being preprocessed to clean reads by trimming adapters and barcode sequences using an internal matching algorithm (BC Cancer Agency), the raw sequencing data (in bam format) were converted into sam, fastq, fasta, and txt formats under Linux in a command-line environment for subsequent use. The sRNA toolbox was used to profile sRNAs and size distribution, perform conserved miRNA analysis using miRNAs for *Arabidopsis* or a high confidence set of miRNAs on miRBase for *P. glauca*, and their consensus miRNA families and unique sequences (Rueda et al., 2015). sRNAs in sequencing libraries were computationally predicted against the *P. glauca* genome assemblies (highly fragmented, containing 3.1 million scaffolds longer than 500 bases and a N50 of 43.5 kbp; PG29 v3, 20Gb divided into 30 Mb per file) (Warren et al., 2015) and the *Arabidopsis* genome (TAIR10), respectively, using miRPlant (An et al., 2014) at default settings with a sequence length cutoff of 18~24 nt. As the size of our *P. glauca* sequence libraries exceeded

the maximal single load on miRPlant, we divided each library into several sub-ones with a maximum size of 120 Mb under Linux and after combining the output files in the same library (e.g., summing up raw abundance for the same unique reads from different sub-files), duplicated reads were removed and we only retained one copy of unique reads with the highest prediction score for each library using R 3.2.2 (The R Project for Statistical Computing), followed by a visual inspection to ensure the R coding has attained our objectives.

The libraries subtracted r/t/sn/snoRNAs (i.e., ribosomal RNA, transfer RNA, small nuclear ribonucleic RNA, and small nucleolar RNA, originally sourced from Sanger RNA family database 12.0, Nawrocki et al., 2015) and non-sRNAs from the total library size, and the resulting number was used for abundance normalization. Abundances were expressed throughout this study in reads per million (RPM) unless otherwise indicated. Due to the unavailability of complete r/t/sn/snoRNA sequences in *P. glauca*, we utilized *Arabidopsis* r/t/sn/snoRNA annotations to classify non-sRNAs in *P. glauca*, which include reads of non-predictable secondary RNA structure and non-mapped to the genome. After genome-wide identification of sRNAs, the unique sRNA sequences with raw abundance above 10 at least in one library, along with raw and normalized counts and precursors, were archived and compared with previously identified spruce sRNAs (Källman et al., 2013; Xia et al., 2015). To isolate conserved miRNAs by using a homolog search, sRNAs in this study were aligned versus a Viridiplantae-specific miRBase reference file containing 100,014 miRNA sequences from 1,397 miRNA families (Chávez Montes et al., 2014), curated in Table S1. Note that miRBase v21 (<http://www.mirbase.org/>) has archived c. 8,000 miRNAs from 73 plant species.

Heat maps graphically depicted the most conserved and differentially expressed miRNAs at seed set across populations/ecotypes, whereby cluster analyses were performed for seed-set phases and key miRNAs within phylum using an euclidean method. Note that raw processed data (i.e., log<sub>2</sub> of count per million) underwent a log<sub>10</sub>-transformation before being fitted on the map.

Top-one targets of the archived unique sRNAs were predicted using transcripts without miRNA genes on psRNATarget (Dai and Zhao, 2011). We focused on the sRNAs of high strength of prediction (score  $\geq 0$ ) in *Arabidopsis*, while in *P. glauca*, on sRNAs that were detected in at least 14 of 15 libraries across four populations, termed “most conserved” sRNAs hereafter (N.B. total reads in one library are much less than the others). To annotate target mRNA functions, the top predicted target gene for each sRNA of interest was aligned against the Gene Ontology (GO) protein database for GO term classification and KEGG pathway enrichment (Ashburner et al., 2000; Kanehisa and Goto, 2000).

## Identification of Adaptation-Associated sRNAs

In a parallel analysis to search for potential sRNAs targeting genes involved in adaptation to climate, we manually refined sRNAs through identifying their target genes that function in stress responses for *Arabidopsis*. Reportedly, there were 73 adaptive

genes in *P. glauca* (17 overlapped between two articles, Hornoy et al., 2015; Yeaman et al., 2016). We adopted the pipeline of user-submitted small RNAs and transcripts on psRNATarget (Dai and Zhao, 2011) to detect the sRNAs that may target the 73 genes. As most conifer genes were unannotated, we employed a reciprocal BLAST to identify homologs. Specifically, miRNA-targeted genes in *P. glauca* were retrieved via a BLASTN search against the *Arabidopsis* genome on EnsemblPlants (<http://plants.ensembl.org>) and then putative proteins in *Arabidopsis* were searched via a tBLASTN (six frames) against the *P. glauca* PlantGDB Putative Unique Transcripts (PUTs) database on ConGenIE (<http://congenie.org/>). Each PUT underwent another BLASTN search against the *Arabidopsis* genome. Homologs were identified only if the same pair of sequences could be found among the top three candidates in the two BLASTNs. Adaptive genes functions were classified based on the records in TAIR10 (<https://www.arabidopsis.org/>).

## Environment Association Analysis

To extract expression structures of genotypes (represented by miRNA and mRNA relative expression) that can be explained by environments (temperature at seed set and phenology represented by developmental phase and pattern), redundancy analysis (RDA) was conducted (Oksanen et al., 2015) and visualized by a constrained ordination triplot with both response and explanatory variables in the same coordinate using a scaling 2 method. RDA combines multivariate regression with PCA of multiple dependent variables and we used it to test and quantify the overall contribution of a primary climate variable to the expression pattern of different types of miRNA and mRNAs. This analysis was carried out under R 3.2.2.

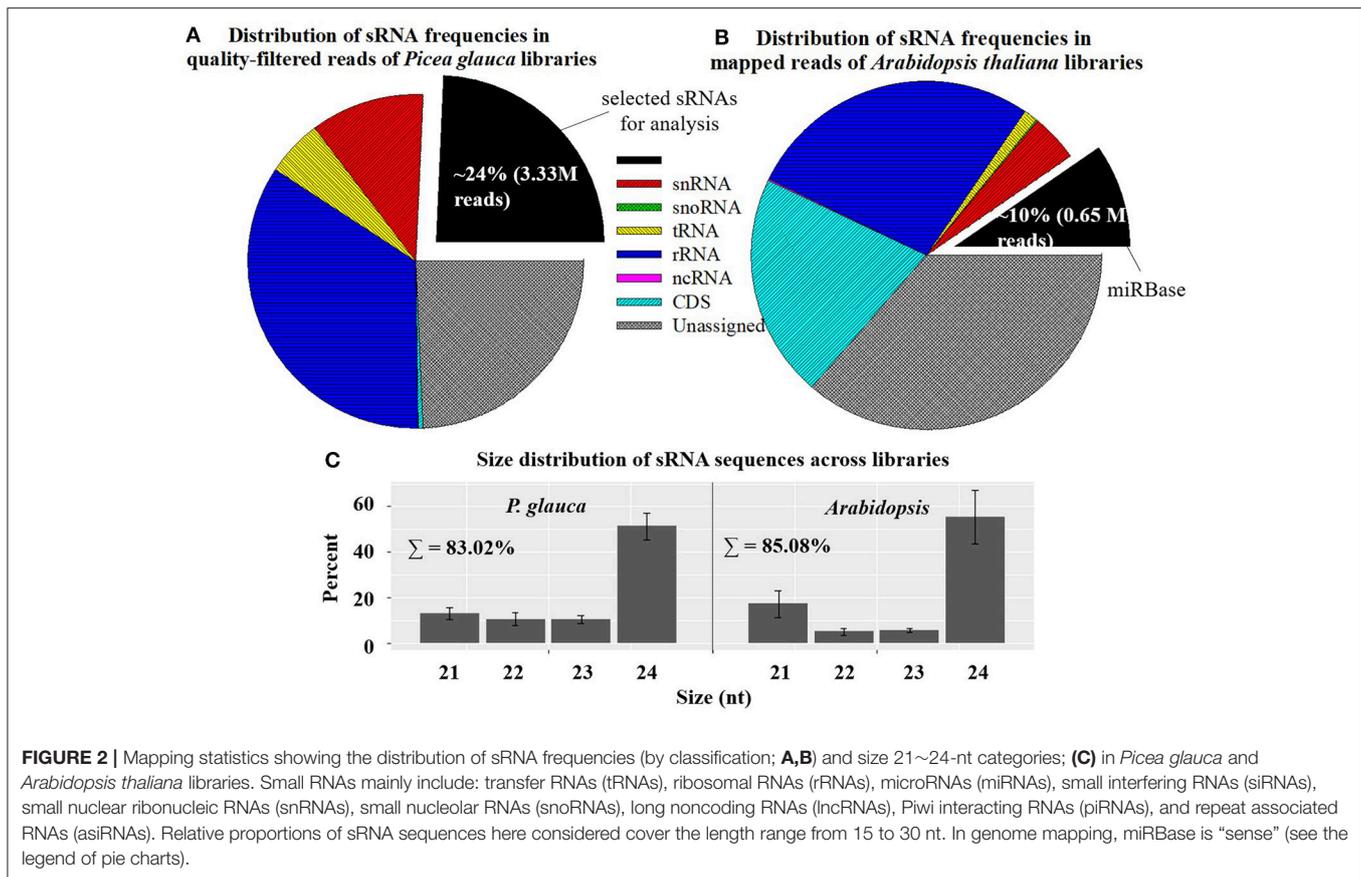
## Gene Expression Analysis

With high confidence, genes targeted by conserved miRNAs of interest were experimentally validated using a quantitative RT-PCR (qRT-PCR) assay as follows. Two microgram of the other aliquot of total RNAs was reverse-transcribed into cDNAs using the EasyScript Plus™ kit (abmGood) with oligo-dT primers following the manufacturer's instructions and first-strand cDNA synthesis products were diluted fivefold as qRT-PCR templates. qRT-PCR was run in 15  $\mu$ l reaction volumes on an ABI StepOnePlus™ machine (Life Technologies) using the PerfeCTa® SYBR® Green SuperMix with ROX (Quanta Biosciences). The reaction components and procedure were carried out as previously described (Liu et al., 2015). We adopted three technical replicates for each pooled sample of three biological replicates. Reference genes were used as previously described (Czechowski et al., 2005; Liu et al., 2015). Gene homolog identifications and primer pairs for the qPCR amplification were listed in Tables S2, S3.

## RESULTS

### Small RNA Transcriptomic Profiling throughout Seed Set

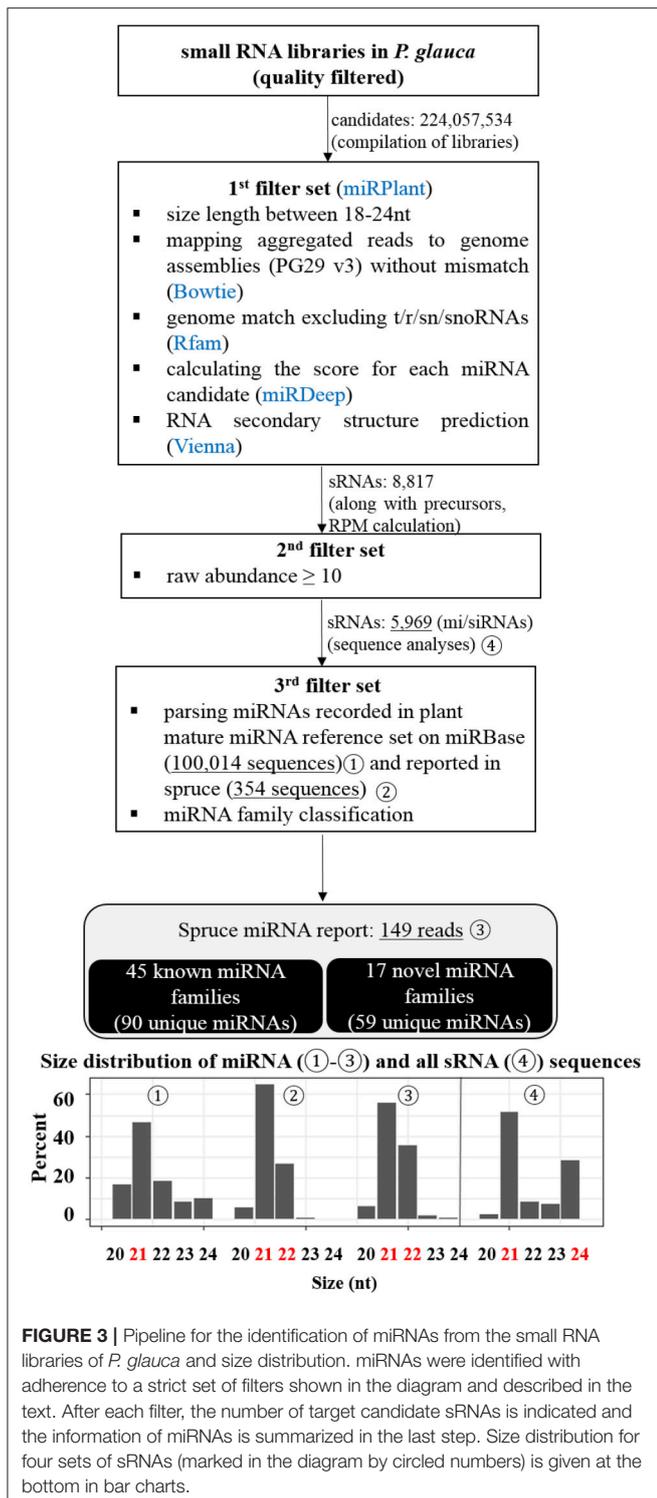
Small RNA libraries were prepared from four seed developmental stages of four populations in *P. glauca* (N.B. only three timepoints



for population 1 due to its short seed developmental duration) and 10 consecutive days after pollination (0~9 DAP) of Cvi-0 and Col-0 in *Arabidopsis* plus another 5 days (10~14 DAP) for Cvi-0 due to its slow seed development (Figure 1). After a quality filtering, the Illumina deep sequencing yielded 15.1M reads, on average, in 15 *P. glauca* libraries (Table S4). By contrast, there were on average 10.3M reads in 25 *A. thaliana* libraries (Table S4). Mapping the quality-filtered reads to their corresponding genome with predictable hairpin RNA secondary structures generated an average of 3.33 M ( $24 \pm 6\%$ ) and 7.33 M ( $84.3 \pm 5\%$ ) reads in *P. glauca* and *Arabidopsis* libraries, respectively (Figure 2A and Table S4). Low proportion of sRNAs for *P. glauca* may be due to the state of scaffold-level genome assemblies in *P. glauca* compared with the complete assembled genome by chromosomes in the model organism, *Arabidopsis*. Using the archived miRNA reads under *Arabidopsis* species on miRBase, the known miRNAs among quality-filtered sRNA reads occupied on average 0.3% (0.05 M) and 10% (0.65 M) in *P. glauca* populations and *Arabidopsis* ecotypes over time, respectively (Table S4 and Figure 2B). This indicates that an appreciable amount of other types of sRNAs (e.g., siRNAs) is largely produced at seed set of *P. glauca*. The libraries of *Arabidopsis* and *P. glauca* showed, in consistency, that 24-nt sRNAs were dominantly produced at seed set (Figure 2C). In addition, sRNA reads in *P. glauca* classified into

r/t/sn/snoRNAs took up 34.4, 5.5, 10.9, and 0.08% of the total sequences, respectively (Figure 2A); analogously, these small RNA categories were 27.3, 1.3, 4.4, and 0.1%, respectively, in *Arabidopsis* (Figure 2B).

After a suite of filters for *P. glauca* libraries, we identified 5,969 sRNA sequences, including 149 miRNAs from 62 miRNA families (Figure 3), curated in Tables S5, S6, respectively. Compared with previous reports (Källman et al., 2013; Xia et al., 2015), 90 were known miRNAs, whilst 59 were novel (Table S6). Size distribution showed that miRNAs are mainly 21-nt long throughout the plant kingdom (Figure 3-①) and identified miRNAs in spruce consistently exhibited that 21- and 22-nt were enriched in previous and this study (Figures 3-②,③). As the 21-/22-nt size class mainly consists of miRNA and phasiRNAs (or tasiRNAs), this result suggests that tasiRNAs or phasiRNAs (triggered by miRNA pairing) may play an important role in conifers, as already examined in *P. abies* (Xia et al., 2015). Additionally, the size distribution of all sRNAs showed that both 21- and 24-nt were abundant (Figure 3-④), indicating that hc-siRNAs for the reinforcement of silencing chromatin marks were highly generated typically at seed set in spruce. This observation from developing seeds (~30%) is significantly different from that in buds (~1%) (Källman et al., 2013), which confirms that the 24-nt long sRNA class is specific to reproductive tissues in conifers (Nystedt et al., 2013).



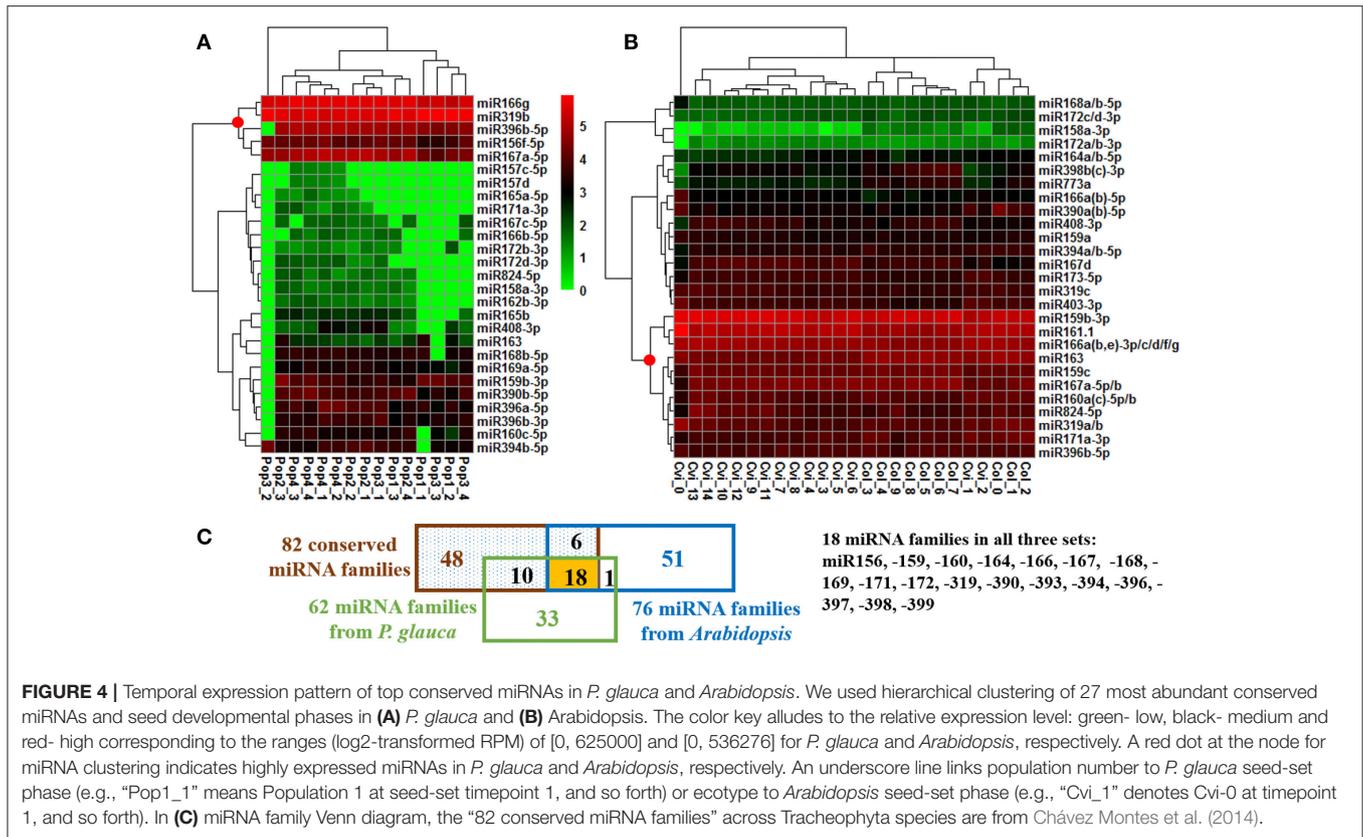
## Association between sRNA Expression Pattern and Seed-Set Phases

The top 40 conserved and differentially expressed miRNA sequences between different timepoints at seed set were respectively selected from *P. glauca* and *Arabidopsis* libraries and

their expression was showcased in bivariate plots (Figure S1). In small panels of Figure S1, histograms along the diagonal were distributed in a right-skewed manner and scatter plots for *P. glauca* and *Arabidopsis* generally showed a monotonic and linear relationship, respectively. This indicates that the expression of conserved miRNAs is intrinsically correlated among different phases of seed set and populations/ecotypes.

The expression of the top 27 conserved miRNA families in *P. glauca* and *Arabidopsis* seed set was respectively employed to create a heat map and perform cluster analyses for seed-set phases and normalized miRNA expressions over time (Figures 4A,B). In *P. glauca*, the seed-set phases in the same population were not completely classified in one of the four major “clades” (Figure 4A). However, in *Arabidopsis*, the phases at 3DAP onward were neatly clustered into two primary groups by ecotype (Figure 4B). The expression pattern for Cvi-0 at flowering was different from other phases and it constituted a single “clade” (Figure 4B), while early seed-set for Cvi-0 (i.e., Cvi\_1~2) and Col-0 (i.e., Col\_0~2) were within the same group (Figure 4B). This indicates that the pattern of relative expression of conserved miRNAs was highly correlated with seed-set phases in *Arabidopsis* but not in *P. glauca*, and miRNAs significantly regulate seed development at least as early as 3DAP in globular embryos of *Arabidopsis* ecotypes. Regarding the most highly expressed and conserved miRNAs, miR166g and -319b were most abundant, followed by miR396b-5p, -156f-5p, and -167a-5p, in *P. glauca* (Figure 4A), while in *Arabidopsis*, most enriched ones were miR159b-3p, -161.1, -166a(b,e)-3p/c/d/f/g, -163, and -159c, 167a-5p/b (Figure 4B). Taken together, miR166, -319, and -396 families were highly expressed in both *P. glauca* and *Arabidopsis* and reportedly, they play regulative roles in seed/cell development (see a summary for selected miRNAs and their functions in Table S8). In general, the most abundant miRNA families (under the “clade” marked by a red dot in Figure 4) were overlapped in *P. glauca* and *Arabidopsis* but abundant miRNA families were less in number in *P. glauca* (Figure 4). Moreover, 18 conserved miRNA families across vascular plants were identified in both *P. glauca* and *Arabidopsis* (Figure 4C), in which some were engaged in auxin and GA signaling pathways, including miR159, -160, and -167 (Figure 4 and Table 1). In addition, there were conserved miRNAs uniquely detected at seed set in the Col-0 ecotype or in the population of late maturation (Pop 4) in *P. glauca* (Table 2).

Because the majority of miRNA families has a very limited taxonomic distribution (Cuperus et al., 2011), it is meaningful to carry out a lineage- and population-specific sRNA analysis. Here, lineage-specific sRNAs were represented by sRNAs identified throughout the studied populations in respective species. In *P. glauca*, there was no significant difference in expression abundance and pattern for lineage-specific sRNAs ( $p > 0.05$ ; a total of 187 sequences used) (Figure 5IA). However, the expression pattern was significantly different between populations, when population-specific sRNAs (435, 1181, 433, and 1065 sequences for Pop 1~4, respectively) were used for comparison (all  $p$ -values  $< 0.05$ ) (Figure 5IB). The sequence length peaked only in the 21-nt size class (no longer at 24-nt size; compare Figures 3-④, 5IC), supportive to miRNAs



and/or tasiRNAs that were abundantly generated and may be selected for.

Analogously, there were 91 conserved miRNAs identified from sequence libraries (Table S9), in which 85 were expressed in both early (Cvi\_0~7 and Col\_0~4) and late (Cvi\_8~14 and Col\_5~9) seed-set phases (Figure S3). Conserved ath-miRNAs had a dominant length of 21 nt (Figure 5IIC) and there was no significant difference in the expression pattern for known ath-miRNAs identified across the seed-set period in both ecotypes ( $p = 0.786$ ; totalizing 23 sequences used) (Figure 5IIA), while significantly different when known ath-miRNAs in either ecotype were compared ( $p = 0.029$ ; 23 and 47 sequences for Cvi-0 and Col-0, respectively) (Figure 5IIB). Interestingly, there was a significant difference for sRNAs in both ecotypes ( $p = 0.027$ ; totalizing 69 sequences; known ath-miRNAs excluded) (Figure 5IID) and no significant difference was found for the ones in either ecotype ( $p = 0.21$ ; 85 and 99 sRNAs for Cvi-0 and Col-0, respectively) (Figure 5IIE). The size of these sRNAs peaked at the 21- and 24-nt size classes (Figure 5IIF). This indicates that hc-siRNAs, as well as tasiRNAs and novel miRNAs may have undergone selection in *Arabidopsis*.

## Small RNA Frequent Emergence and Demise throughout Time

In *Picea glauca*, considerable sRNAs were generated across seed-set phases in populations (Figure S2). The number of sRNAs detected in all four populations (1,318 reads) was as many as

that of unique sRNAs in different populations (1,200 reads on average) (Figure S2A). The unique sRNAs that were expressed across developing phases in all populations were less in number than those that were only detected in one single population (Figure S2B). Burgeoning sets of sRNAs are analogous to the frequent emergence and decay previously reported in novel miRNAs in *Arabidopsis* (Fahlgren et al., 2007). In *Arabidopsis*, the number of sRNA reads without conserved miRNAs was almost 10 times as many as conserved miRNA sequences (Figure S3A). Different from conserved miRNAs that were expressed in both early and late seed-set phases, these unknown sRNAs were differently produced in phases as well as between ecotypes (Figures S3A,B). A total of 705 unknown sRNA reads out of 5,749 reads were detected in both early and late seed-set phases in the two ecotypes (Cvi\_0-7 vs. Cvi\_8-14 and Col\_0-4 vs. Col\_5-9; Figure S3B), in which 44 sequences were expressed across all seed-set phases in two ecotypes.

## Targets of Abundant sRNAs and their Involvement in Adaptation

In light of the tight correlation between sRNA conservation and expression abundance (Chávez Montes et al., 2014), GO classification for target genes of the 107 “most conserved” sRNAs in *P. glauca* (Table S7) showed that sRNAs were mostly involved in the regulation of pathways related with metabolic and cellular processes (Figure 6A), which was in line with the observation in conserved miRNAs (Table S9) in *Arabidopsis* (Figure 6B).



TABLE 1 | Continued

Name <sup>a</sup>	Predicted target <sup>b</sup>		Alignment		Target gene function <sup>c</sup>	E <sup>d</sup>	UPE <sup>e</sup>
miR390b-5p	AT5G03650.1	miRNA	21 CCGCGAUAGGGAGGACUCGAA	1	Starch branching enzyme	1.5	21
		Target	410 GGCGUUAUCCUUCUUGAGCUU	430			
	EX354481	miRNA	21 CCGCGAUAGGGAGGACUCGAA	1	–	1.5	17
		Target	137 GGCGUUAUCCUUCUUGAGCUU	157			
miR394b-5p	AT1G27350	miRNA	20 CCUCCACCUGUCUUACGGUU	1	Ribosome associated membrane protein	1	15
		Target	2592 GGAGGUUGACAGAAUGCCAA	2611			
	BT112917	miRNA	20 CCUCCACCUGUCUUACGGUU	1	–	1	20
		Target	830 GGAGGUAGACAGAAUGCCAA	849			
miR396-5p	AT1G53910	miRNA	20 UCAAGUUCUUUCGACACCUU	1	Ethylene response factor	0	14
		Target	526 AGUUCAAGAAAGCUGUGGAA	545			
	BT102125	miRNA	21 UUCAAGUUCUUUCGACACCUU	1	–	1.5	34
		Target	164 AAGUUCAAGAAAGCCGUGGAA	184			
miR408-3p	AT2G02860	miRNA	20 GGUCCCUUCUCCGUCACGUA	1	SUCROSE TRANSPORTER 3	1	23
		Target	687 CAAGGGAAGAGGCAGUGCAU	706			
	BT103532	miRNA	20 GGUCCCUUCUCCGUCACGUA	1	–	2	25
		Target	102 CCAGGGAGGAGGCAGUGCGA	121			
miR824-5p	AT3G57230	miRNA	21 AGGGAAGAGUGUUUACCAGAU	1	MADS-box transcription factor	0.5	15
		Target	2516 UCUCUUCUCACAAAUGGUCUA	2536			
	BT112142	miRNA	21 AGGGAAGAGUGUUUACCAGAU	1	–	3	13
		Target	266 UCUCUCUCAUAAAUGGUCUC	286			

<sup>a</sup>Nomenclature of miR: (organism)miR<sub>n</sub>x - precursor arm and/ or y, where n, a sequential number representing family of miR; x, lettered suffixes representing family member (i.e., closely related mature sequences); –5p or –3p denote 5' or 3' arm of the precursor; y, integer denoting occurrence of more than one mature sequence from the same precursor.

<sup>b</sup>For each miRNA, shown is the most confidently predicted in *Arabidopsis* (above) and *P. glauca* (below).

<sup>c</sup>Refer to GO enrichment analysis, TAIR10 and NCBI.

<sup>d</sup>Expectation (E), stringent threshold [0–0.2] gives lower false positive prediction.

<sup>e</sup>Maximum energy to unpair the target site (UPE), small value (range € [0, 100]) is better.

The experimentally validated miRNA-gene pairs were listed in **Figure 6C** and their target genes are mainly involved in seed development (**Figure 6C**).

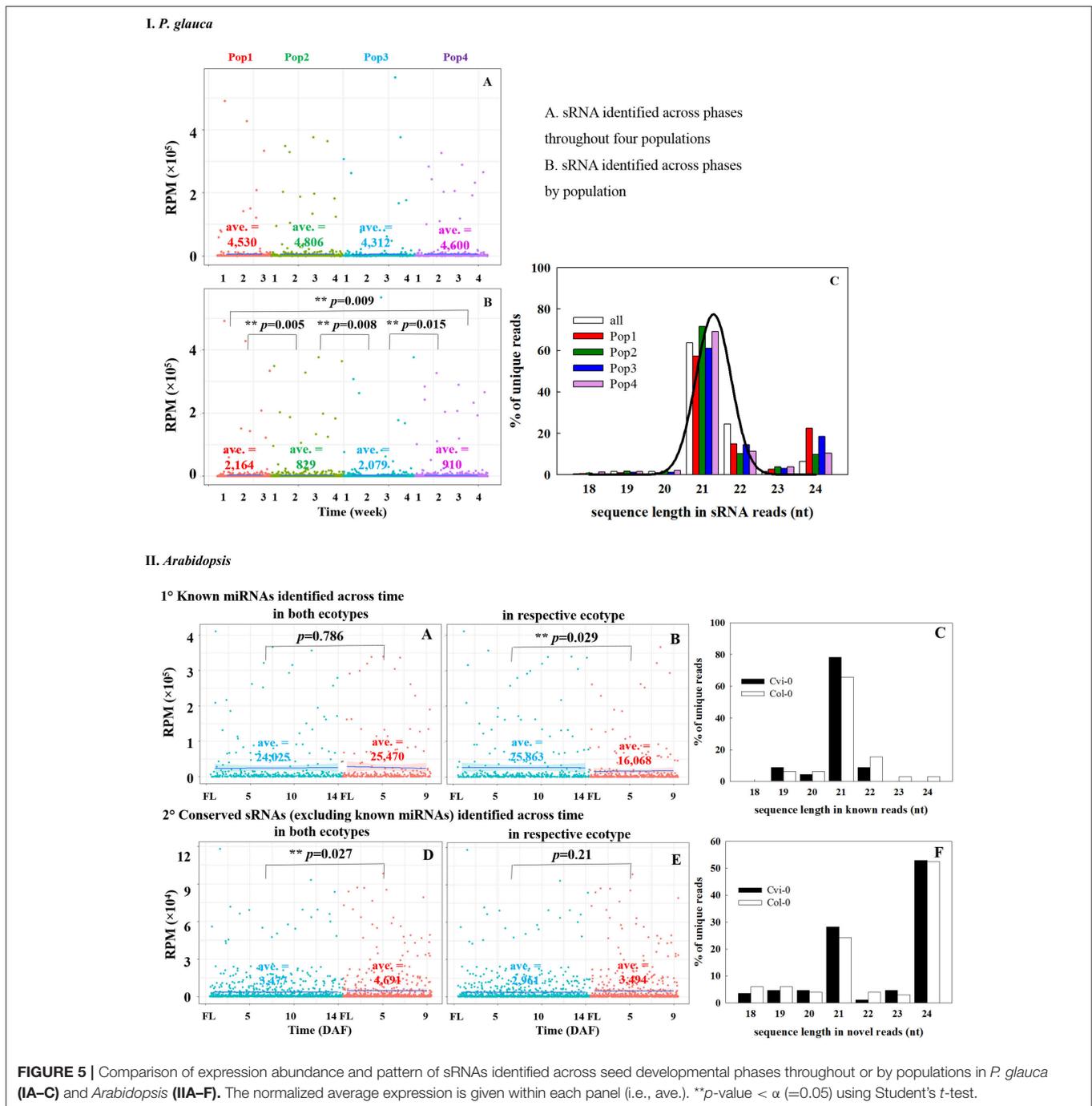
We adopted an Illumina sequencing approach to get a deep coverage of mature sRNAs and predicted sRNA candidates that may target the reported genes involved in adaptation to climate in *P. glauca*. We found 24 sRNA candidates (**Table 3**) targeting genes putatively functioning mainly in abiotic and biotic stresses (**Figure 7A**). In model organisms, a group of key and conserved miRNA sequences involved in plant seed development and phase transitions was summarized in Table S8, some of which are involved in adaptation to stresses (see description in the “function” column). They were identified by fold changes in miRNAs between stress-treated and control samples. With reference to previous reports, we identified 9 miRNAs

involving in adaptation during *Arabidopsis* seed development and they were miR159, –167, –169, –393, –395, –397, –398, –399, and –408 (**Figure 7B**). These miRNA families also have functionalities in seed development via regulating transcriptional factors in hormone-based GRNs, such as miR159, –167, –393 targeting *GAMYB*, *AUXIN RESPONSE FACTORS* (*ARFs*), auxin-receptors and cyclin-like F-box, respectively (see a summary in Table S8).

### Expression Pattern of Selected miRNA and Genes Explained by the Environment

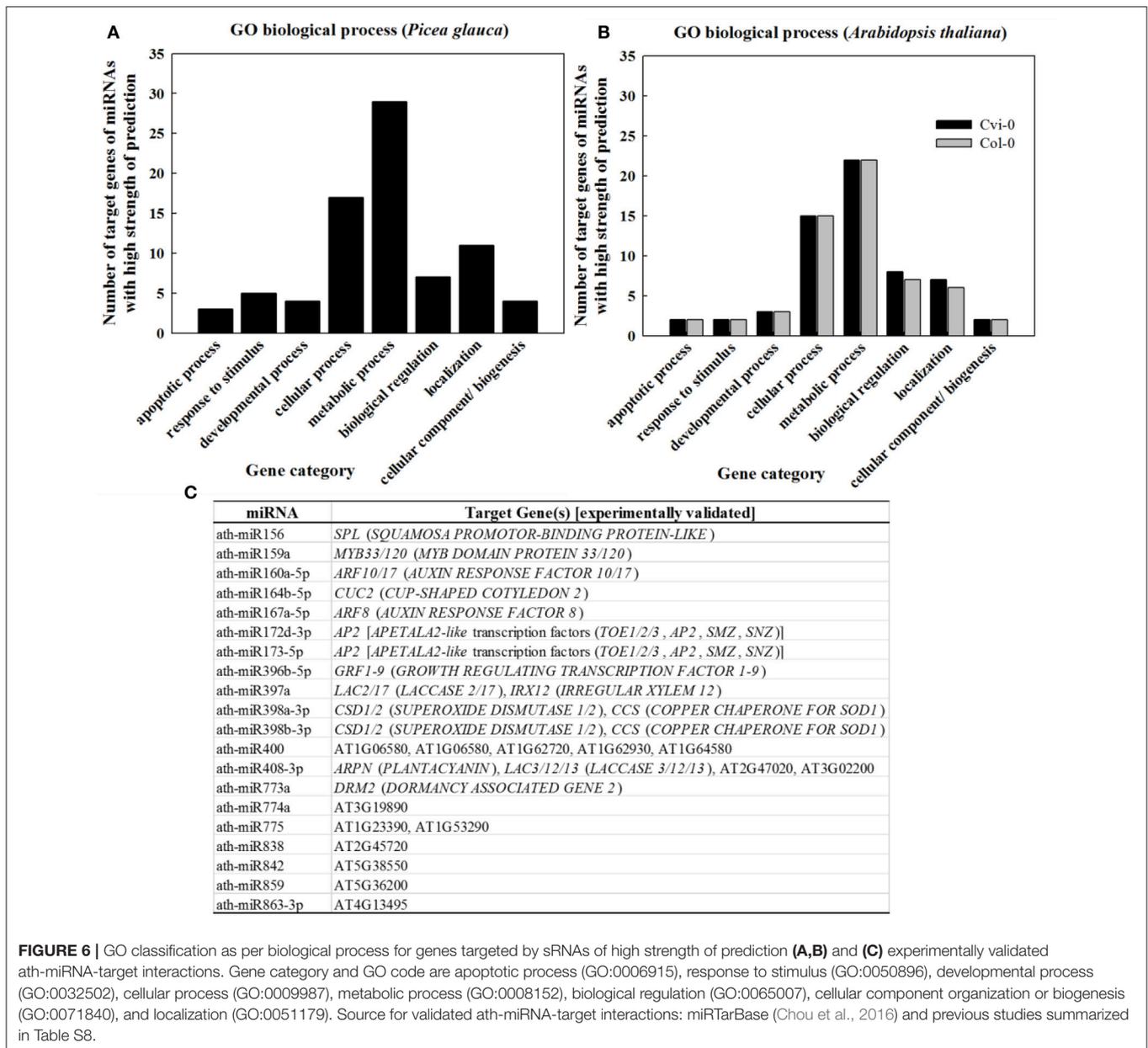
We chose the seed dormancy phenotype to investigate how environmental signals impinge on the expression of conserved miRNAs targeting genes related to seed dormancy and also that of key conservative genes conducive to the phenotype, thus





collectively manipulating phenotypic variation. Gene phylogeny for *ARF10/16* showed that gymnosperm and model angiosperm species were separated into different clades except for *Picea*, which has a higher sequence similarity with angiosperms than species in the same taxonomic category (Figure S4). This indicates that *ARF10/16* is ancient and evolutionarily conserved within the plant kingdom. Conserved domain analysis showed that putative *ARF10* in *P. glauca* harbors an Aux\_IAA super family domain (Figure S5). The pivotal activation function of

*ARF* proteins is conferred by their four-domain architecture, including DNA binding region (a B3 and an ARF domain) and protein dimerization motifs (Ulmasov et al., 1999; Tiwari et al., 2003). Loss of the canonical four-domain structure has promoted functional shifts within the ARF family by disrupting either dimerization or DNA-binding capacities (Finet et al., 2013). As such, the putative *ARF10* in *P. glauca* may not correspond to its counterpart in *Arabidopsis* or the essential Aux\_IAA domain is sufficient to have the function of *ARF10* in *P. glauca*. No homolog



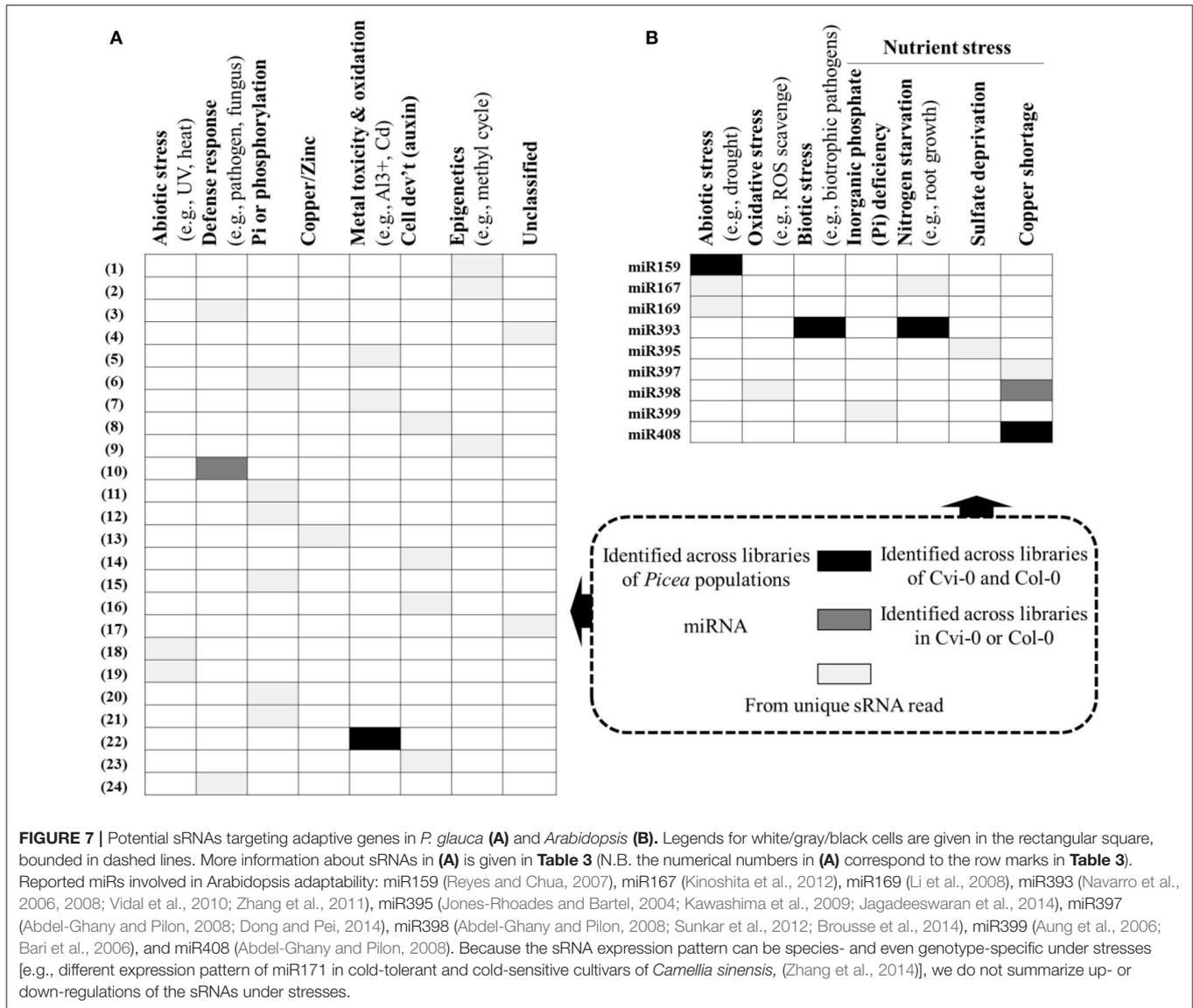
hits for *ABI3* and *DOG1* were obtained in some species (Table S2) and the phylogeny showed that they might have undergone substantial selection (Figure S4). The miR160 targets *ARF10* and its hairpin structure was computationally predicted (Figure S6).

The relative expression of aforementioned genes was displayed in Figure S7. As predicted, *AtDOG1*, *AtABI3*, *AtARF10*, and *AtARF16* were highly expressed in Cvi-0 than Col-0 (Figure S7, left panels). However, such pattern was not observed for gene counterparts in populations of *P. glauca* (Figure S7, first three panels on the right side), indicating that these genes in *P. glauca* may have different functions or that the studied phenotype in *P. glauca* is regulated by other unknown mechanisms. In the RDA triplot, the percentage of accumulated constrained eigenvalues showed that the first axis explained 43.8% variance (Figure S7, last panel), indicating that the major trends have

been modeled by RDA. In addition to species and dev\_phase, developmental temperature played an important role in the dispersion of developmental phases along the first axis and had high correlation with miR160 and ARF10/16 (Figure S7, last panel). As transcripts of *ARF10/16* are targeted by miR160 (Liu et al., 2013), the expression patterns of *ARF10/16* and miR160 were highly positively correlated with each other but negatively correlated with that of *DOG1* (Figure S7). In addition, the projection of the same developmental phases on the first axis was overlapped across populations in *P. glauca* (Figure S7). This suggests that the same phase between populations has more similarities in gene expression pattern than different phases within populations, and in turn, prompts the conservation of gene regulation at a temporal scale across populations.







**FIGURE 7 |** Potential sRNAs targeting adaptive genes in *P. glauca* (A) and *Arabidopsis* (B). Legends for white/gray/black cells are given in the rectangular square, bounded in dashed lines. More information about sRNAs in (A) is given in Table 3 (N.B. the numerical numbers in (A) correspond to the row marks in Table 3). Reported miRNAs involved in *Arabidopsis* adaptability: miR159 (Reyes and Chua, 2007), miR167 (Kinoshita et al., 2012), miR169 (Li et al., 2008), miR393 (Navarro et al., 2006, 2008; Vidal et al., 2010; Zhang et al., 2011), miR395 (Jones-Rhoades and Bartel, 2004; Kawashima et al., 2009; Jagadeeswaran et al., 2014), miR397 (Abdel-Ghany and Pilon, 2008; Dong and Pei, 2014), miR398 (Abdel-Ghany and Pilon, 2008; Sunkar et al., 2012; Brousse et al., 2014), miR399 (Aung et al., 2006; Bari et al., 2006), and miR408 (Abdel-Ghany and Pilon, 2008). Because the sRNA expression pattern can be species- and even genotype-specific under stresses [e.g., different expression pattern of miR171 in cold-tolerant and cold-sensitive cultivars of *Camellia sinensis*, (Zhang et al., 2014)], we do not summarize up- or down-regulations of the sRNAs under stresses.

## DISCUSSION

To date, evolutionary analysis on miRNAs is almost comprehensive and profound throughout species in the tree of life (Nozawa et al., 2012), but there is a dearth of representative species in a subgroup of gymnosperms—conifers, from which flowering plants bifurcate, thus occupying an important taxonomic position. Although recently there have been sporadic reports on conifer small RNA sequences (e.g., Källman et al., 2013; Xia et al., 2015), no small RNA study is tailed to examine the reproductive period, during which small RNAs of the 24-nt size class are specifically and uniquely yielded (Nystedt et al., 2013) and the whole genome is highly methylated (Takuno et al., 2016). This implies a different landscape of small RNAs during conifer seed set. Furthermore, in multicellular organisms gene expression fine-tuned by miRNAs can decrease phenotypic variation (i.e., canalization) among

individuals and even among cells, thus reducing conflicts among cells of different genetic background (Michod and Roze, 2001; Hornstein and Shomron, 2006; Flynt and Lai, 2008). Together, a supplemental characterization of sRNAs in conifer seed development therefore motivates and necessitates this study. In quest of molecular underpinnings of today's diversity of life and epigenetic phenomena involved in shaping responses to environmental signals, we rely on comparative studies of deeply conserved and known miRNA homologs as well as lineage- and population-specific sRNAs in two spermatophytes, *Picea glauca* and *A. thaliana*, to unravel, from evolutionary perspectives, the pattern of expression dynamics of sRNAs and miRNAs (also via comparing with deeply conserved ones in the plant kingdom), sRNA candidates targeting adaptive genes, and miRNA-gene regulation in one studied phenotype (i.e., seed dormancy). Understanding the underlying mechanism of local adaptation helps reveal evolutionary signatures that may comprise selective

forces for adaptation and speciation. This study attests to the correlation between the sRNA (and miRNA) repertoire and the organismal complexity and provides a potential rationale to explore the evolutionary mechanism at the organism level via *de novo* methylation and sRNA interactions.

## Insights into miRNA Evolution

Of 37 miRNA families that are deeply conserved in plant development throughout the plant kingdom (Willmann and Poethig, 2007), we identified 12 miRNA families at seed set between phyla (**Table 1**). Some are conserved across spermatophytes (i.e., miR163, -394, and -396), tracheophytes (i.e., miR159), and embryophytes (i.e., miR160, -166, -167, -171, -319, -390, and -408). The ubiquitously conserved miRNAs had significantly differential expression abundance between *P. glauca* and *Arabidopsis* as well as between populations within species (e.g., **Figure 5IA** vs. **Figure 5IIA**), and this observation is correlated to genome evolution (Hodgins et al., 2016). These mature miRNA sequences have a well-conserved consensus with minor variation at the 5' or 3' end between *P. glauca* and *Arabidopsis* (**Table 1**), and their targets may not always be conserved partially due to no homologs found in *P. glauca* (**Table 1**). This indicates that the cognate miRNA-target pairs may be acquired prior to the split of gymnosperms and angiosperms and have undergone natural selection. Conserved miRNAs in *Arabidopsis* (**Table S8**) were enriched and expressed across seed set (**Figure 5IIA**), while abundant sRNAs expressed throughout seed set in *P. glauca* did not belong to miRNA families documented on miRBase (**Table S7**). Nonetheless, their target genes were classified into the same GO categories (**Figure 6**). This indicates that spruce and *Arabidopsis* may use different sRNAs to attain similar regulations at molecular levels. Note that the known miRNA-gene pairs in *Arabidopsis* are associated with seed development and nodes in hormone-based GRNs (**Figure 6C**).

In general, ancient miRNAs are more highly and broadly expressed than younger ones (Fahlgren et al., 2007); while newly emerged miRNAs may be used as substrates for natural selection and form specific miRNA circuitry interplaying with GRNs, whereby adaptation and speciation occur (Chen and Rajewsky, 2007). During evolution, deleterious miRNA-target pairs are selected against and miRNAs by gradually decreasing the number of target genes are retained into GRNs (Chen and Rajewsky, 2007). Through constraining variance or mean gene expression, miRNAs render phenotypic traits evolvable as well as heritable (Wagner and Altenberg, 1996) and after selection, these miRNAs may improve fitness of phenotypes (Wu et al., 2009). However, most of young miRNAs appear to be selectively neutral [ $S = 0$ ] (Fahlgren et al., 2010), because the majority of them does not have target genes or does not regulate targets in such a way that plant fitness can be enhanced, and they were quickly lost during evolution. In regulation of target genes with similar functions, different species may rely on variants of the same *MIRNA* families or produce distinct miRNAs through changes in miRNA precursor sequences and/or binding sites. From this study, the conserved miRNAs throughout the plant kingdom and lineage-specific sRNAs (**Figures 4, 5**) represent different levels

of sRNA conservation serving for basic needs in plants and particular requirements in a given lineage, respectively. On the other hand, although members of plant *MIRNA* families are often highly similar, the same gene family varies in size and genomic structure between species, indicating dosage effects and spatiotemporal differences in *MIRNA* regulation (Li and Mao, 2007). There are miRNA orthologs that function quite differently and several miRNA isoforms have specific tissue expression patterns (Lelandais-Brière et al., 2009), indicative of functional divergence. We detected some miRNAs solely expressed in one population/ecotype, which were conserved in miRNA families on miRBase (**Table 2**). As per target gene function, target genes do not seem to be key components in GRNs and they primarily target repeats, intergenic sequences, and genes in cell-cell communication and signaling (**Table 2**). This suggests that altering mean miRNA expression in different plant populations is subject to fine-tuning to regulate the expression of target genes, especially for non-key ones in GRNs. The dosage balance hypothesis, a mechanism in maintaining *MIRNA* duplicates after WGDs (possibly also in small-scale tandem or segmental duplication events due to little impact on maintaining optimal stoichiometry among gene products; e.g., Birchler and Veitia, 2012), supports our inference, as key genes in GRNs are dosage sensitive, such that deletion of one copy may disrupt the whole network.

As requirements for a functional *MIRNA* are less demanding than a protein-encoding gene (PEG), *MIRNA* can easily evolve from various sources of unstructured transcripts, such as gene duplication, intergenic regions, transposable elements (Nozawa et al., 2012). As well, like PEGs, *MIRNAs* are subject to the same evolutionary processes, such as, substitution, insertion, recombination, and natural selection. There are three proposed models that explain the origination of *MIRNAs*, that is, transposable elements, inverted duplication of target genes, and random hairpin sequences with subsequent mutations, reviewed by Cui et al. (2016). A study in spruce supports the second model, evidenced by the precursor sequence of a copy of miR482 highly similar to nucleotide-binding site and leucine-rich repeat domains (*NBS-LRR*) genes (Xia et al., 2015). The third model emphasizes that a serendipitous DNA sequence transcribed and Dicer-processed yields a sRNA that interacts with a homolog target. Coevolution of miRNA sequence and target genes refines such interaction (Felippes et al., 2008). The miRNA-target complementarity in plants is extensive and stringent (Rhoades et al., 2002), suggestive of coevolution of miRNAs and their targets. Taken together, miRNA-target systems have undergone evolution in the context of genome evolution; despite divergences in miRNA sequences, the function of their target genes is largely of commonality (e.g., indirect evidence from **Figure 6**) and this supports similar molecular mechanisms and signaling cascades at seed set across phyla.

## Roles of sRNAs in Local Adaptation

To cope with the vagaries of environmental perturbations, plants have evolved mechanisms of stress avoidance (acclimation) and stress tolerance (adaptation) via well-honing gene expression. Reprogramming of gene expression via sRNAs is a major defense

mechanism in plants as response to stresses (see a list of references in the **Figure 7** legend). The sRNA pathways through (P)TGS intersect with mechanisms regulating different steps in the life of an mRNA, starting from transcription and ending at mRNA decay, in both nuclear and cytoplasmic compartments. Stresses usually trigger destabilization of chromatin states, manifested as changes in DNA methylation and reductions in nucleosome occupancy, and some of these processes involve miRNAs via the RdDM pathway and TE-associated 21-nt siRNAs (Dowen et al., 2012). Recently, a type of stress-induced, UTR-derived siRNAs in 24-nt long has been identified in *Brachypodium* and proposed to mask *cis*-elements (i.e., pre-mRNA) and direct the splicing machinery to use the correct splice site (Wang et al., 2015). Under stress, alternative splicing results in production of different splice variants and thus proteins with different functions required for survival, reviewed by Mastrangelo et al. (2012).

Studies in genomic basis of adaptation in long-lived organisms have largely focused on identifying intraspecific DNA variation and comparative gene expression that are associated with adaptation, reviewed by Prunier et al. (2016). In boreal and temperate regions, conifers with complex life cycles have been found to rely on the regulation of certain gene expression to defense against climatic stresses (Hornoy et al., 2015; Yeaman et al., 2016). These adaptive genes mainly participate in the process against abiotic stresses (e.g., cold injury, resistance to aridity, high salinity, UV-B, etc.) and fungal pathogens, growth cessation and bud set to synchronize with seasonal climatic changes. As such, divergence in gene expression possibly through stabilizing or directional selection (because studies in animals indicate that most gene expression variability may be deleterious and selected against [ $S = -\infty$ ]) is important to species adaptability and evolution. However, the long generation time (i.e., old age of sexual maturity) imposes limitations on natural selection and adaptive evolution for genes to cope with rapidly changing climate. Consequently, these organisms confront a real need to use epigenetics, owing to its merits of high flexibility, to deal with a wide range of environmental conditions throughout their life time. Our reported sRNAs (**Figure 7A** and **Table 3**) have opened a venue into better understanding evolutionary trajectories among small RNAs targeting adaptive genes (albeit not experimentally confirmed), as natural selection of sRNAs in different lineages and among populations within lineage may exhibit different evolutionary patterns and delving into these patterns for different organismal complexity helps uncover sRNA evolution in line with the complexity of organisms.

## Effects of the Environment on Phenotypic Variation

The phenology or temporal control of the life cycle provides adaptive strategies to avoid adverse consequences in harsh environments at seed set and seedling establishment (Krämer, 2015). In the life cycle, temperature is of utmost importance in seed set, germination, and seedling stage due to epigenetic imprinting and their fragile state (Liu et al., 2016). At seed set, temperature signals are a critical selective pressure and

have a strong influence on life history traits, such as timing of seed set and seed dormancy depth (Springthorpe and Penfield, 2015; Vidigal et al., 2016). Specifically, the temperature-mediated control of flowering has evolved to constrain the maternal environment for setting seeds to a specific temperature window, thus yielding seeds with dormancy variation. Maternal environments (e.g., temperature cues) have a persistent and transgenerational effect on the expression pattern of regulatory molecules in GRNs (e.g., *DOG1*, Chiang et al., 2013) and on the biogenesis of versatile miRNAs in life histories. The latter was confirmed by miRNAs (i.e., miR160) targeting key genes (i.e., *ABI3* and *ARF10/16*) that modulate timing of seed set and the dormancy state (Huo et al., 2016). In this study, the chosen populations of *P. glauca* and ecotypes of *Arabidopsis* are characterized by different fertilization timepoints, contrasting seed-set durations and ensuing phenotypic variation (i.e., dormancy intensities; **Figure 1**). We tested how the role of miRNA-gene interactions is at play for one study phenotype, seed dormancy, in which we chose a miRNA-gene pair (i.e., miR160-*ARF10/16*) and key genes (i.e., *ABI3* and *DOG1*) (**Figure S7**). Although the chosen gene may have different functions in *P. glauca*, our results reinforce the notion that environment factors, such as temperature have great impact on the expression of genotypes (miRNAs and genes).

## CONCLUSIONS AND PROSPECTS

This study accentuated that roles of sRNAs in phenotypic variation are canalized at the reproductive period and coupled with species of different evolutionary time, which, metaphorically like writing on a palimpsest for the next generation, sets molecular imprinting on phenotypes that will be expressed in the adult stage for local adaptation. Through global analysis of small RNA dynamics across and among populations as well as within and between *P. glauca* and *Arabidopsis*, we demonstrated small RNA evolution and shed light upon its link to organismal complexity and genome evolution, as evidenced by (i) the expression pattern of deeply conserved miRNAs reflected different seed-set programs in *Arabidopsis* ecotypes (**Figure 4**); (ii) expression patterns of lineage-specific sRNAs enriched at the 21(or 22)-nt size class had no significant difference between populations but population-specific sRNAs were differently expressed in both *Arabidopsis* and *P. glauca* (**Figure 5**); (iii) sRNAs targeting reported adaptive genes were computationally predicted and both of them were not overlapped between the two species (**Figure 7** and **Table 3**); and (iv) environmental factors (e.g., temperature) had significant influence on the expression of key genes and miRNAs at seed development in both species (**Figure S7**). Notwithstanding deeply conserved miRNA families play an important role across a plurality of plants, few studies isolate conserved miRNAs and sRNAs (including siRNAs and novel miRNAs) to comparatively uncover the frequent presence and absence of sRNAs in large quantities and its evolutionary significance shaped by natural selection.

In the future, we wish to advance the non-coding sRNA study in conifers with emphasis on the following two facets.

One fascinating aspect is to examine the pattern of changes in DNA methylation in developing seeds and vegetative tissues among conifer populations and test whether the pattern in (de-)methylation is congruent with that observed in selection of sRNA sequences especially in the short size class (i.e., 21~22-nt). The motivation for this aspect comes from intriguing reports showing conspicuous 24-nt small RNAs (Nystedt et al., 2013) and a very high level of DNA methylation (Takuno et al., 2016) uniquely yielded at the reproductive period in conifers. Further on, comparative profiling of epigenetic changes in different populations has the potential to manifest how epigenetic mechanisms contribute to the gene expression regulation. The other aspect aims to investigate whether *MIRNAs* originate from TEs and undergo evolution coupled in time with genome evolution between species within spermatophyte and how the environment contributes to adaptive variation at molecular levels (e.g., epigenetic imprinting, genetic variants, etc.). This conception is emanated from our results concerning *MIRNAs* for abundant sRNAs in *Arabidopsis* (i.e., conserved miRNAs) containing more DNA repeat modules than those for enriched ones in *P. glauca* (Figure S8; also see Liu and El-Kassaby, 2017b). This may be linked to their giga-genome evolution, in the sense that genetic divergence is suppressed in conifers, thus leading to few WGDs. As our increased knowledge of macro-structural features of giga-genomes, this study will open new perspectives for understanding the evolutionary mechanism of sRNAs in association with *MIRNAs*, sRNA targets and genome evolution (e.g., avoiding pseudogenization via sub- or neofunctionalization, relative dosage or increased gene dosage) in forest trees.

## DATA AVAILABILITY

The sRNA sequencing data has been deposited at Sequence Read Archive (SRA) in the National Center for Biotechnology

Information (NCBI) under the accession numbers, SRP096198, SRP096194 and SRP072220.

## AUTHOR CONTRIBUTIONS

YL conceived this study, performed data analyses, and wrote the manuscript; YE coordinated the project.

## FUNDING

This project was funded by the Johnson's Family Forest Biotechnology Endowment and the National Science and Engineering Research Council of Canada Discovery and Industrial Research Chair to YE.

## ACKNOWLEDGMENTS

We would like to extend our sincere gratitude to B. Jaquish (Ministry of Forestry) for sampling developing cones of white spruce, to J. Denny (CMMT) for measuring RNA quality of the developing seed samples using Bioanalyzer, to D. Miller (BCCA) for sequencing sRNAs from our samples, to R. Hamelin (UBC) for providing ABI StepOnePlus™ machine for qRT-PCR assays, to S. Aitken (UBC) for providing fine equipment for total RNA extraction, and to R. Baranowski (UBC) for setting up our data analysis pipeline on WestGrid (Compute Canada). Lastly, we are thankful to the reviewers for their constructive comments.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <http://journal.frontiersin.org/article/10.3389/fpls.2017.01719/full#supplementary-material>

## REFERENCES

- Abdel-Ghany, S. E., and Pilon, M. (2008). MicroRNA-mediated systemic down-regulation of copper protein expression in response to low copper availability in *Arabidopsis*. *Journal of Biological Chemistry* 283, 15932–15945. doi: 10.1074/jbc.M801406200
- Ali-Rachedi, S., Bouinot, D., Wagner, M. H., Bonnet, M., Sotta, B., Grappin, P., et al. (2004). Changes in endogenous abscisic acid levels during dormancy release and maintenance of mature seeds: studies with the cape verde islands ecotype, the dormant model of *Arabidopsis thaliana*. *Planta* 219, 479–488. doi: 10.1007/s00425-004-1251-4
- Almeida, R., and Allshire, R. C. (2005). RNA silencing and genome regulation. *Trends Cell Biol.* 15, 251–258. doi: 10.1016/j.tcb.2005.03.006
- An, J., Lai, J., Sajjanhar, A., Lehman, M. L., and Nelson, C. C. (2014). miRPlant: an integrated tool for identification of plant miRNA from RNA sequencing data. *BMC Bioinform.* 15:275. doi: 10.1186/1471-2105-15-275
- Anderson, L. L., Hu, F. S., Nelson, D. M., Petit, R. J., and Paige, K. N. (2006). Ice-age endurance: DNA evidence of a white spruce refugium in Alaska. *Proc. Natl. Acad. Sci. U.S.A.* 103, 12447–12450. doi: 10.1073/pnas.0605310103
- Ashburner, M., Ball, C. A., Blake, J. A., Botstein, D., Butler, H., Cherry, J. M., et al. (2000). Gene ontology: tool for the unification of biology. the gene ontology consortium. *Nat. Genet.* 25, 25–29. doi: 10.1038/75556
- Aung, K., Lin, S. I., Wu, C. C., Huang, Y. T., Su, C. L., and Chiou, T. J. (2006). *pho2*, a phosphate overaccumulator, is caused by a nonsense mutation in a MicroRNA399 target gene. *Plant Physiology* 141, 1000–1011. doi: 10.1104/pp.106.078063
- Axtell, M. J., and Bowman, J. L. (2008). Evolution of plant microRNAs and their targets. *Trends Plant Sci.* 13, 343–349. doi: 10.1016/j.tplants.2008.03.009
- Axtell, M. J., Snyder, J. A., and Bartel, D. P. (2007). Common functions for diverse small RNAs of land plants. *Plant Cell* 19, 1750–1769. doi: 10.1105/tpc.107.051706
- Bari, R., Datt Pant, B., Stitt, M., and Scheible, W. R. (2006). *PHO2*, microRNA399, and *PHR1* define a phosphate-signaling pathway in plants. *Plant Physiology* 141, 988–999. doi: 10.1104/pp.106.079707
- Birchler, J. A., and Veitia, R. A. (2012). Gene balance hypothesis: connecting issues of dosage sensitivity across biological disciplines. *Proc. Natl. Acad. Sci. U.S.A.* 109, 14746–14753. doi: 10.1073/pnas.1207726109
- Bräutigam, K., Vining, K. J., Lafon-Placette, C., Fossdal, C. G., Mirouze, M., Marcos, J. G., et al. (2013). Epigenetic regulation of adaptive responses of forest tree species to the environment. *Ecol. Evol.* 3, 399–415. doi: 10.1002/ece3.461
- Brousse, C., Liu, Q. K., Beauclair, L., Deremetz, A., Axtell, M. J., and Bouché, N. (2014). A non-canonical plant microRNA target site. *Nucleic Acids Res.* 42, 5270–5279. doi: 10.1093/nar/gku157
- Cairney, J., and Pullman, G. S. (2007). The cellular and molecular biology of conifer embryogenesis. *New Phytol.* 176, 511–536. doi: 10.1111/j.1469-8137.2007.02239.x
- Cairney, J., Zheng, L., Cowels, A., Hsiao, J., Zismann, V., Liu, J., et al. (2006). Expressed sequence tags from loblolly pine embryos reveal

- similarities with angiosperm embryogenesis. *Plant Mol. Biol.* 62, 485–501. doi: 10.1007/s11103-006-9035-9
- Chávez Montes, R. A., De Fátima Rosas-Cárdenas, F., De Paoli, E., Accerbi, M., Rymarquis, L. A., Mahalingam, G., et al. (2014). Sample sequencing of vascular plants demonstrates conservation and divergence of microRNAs. *Nat. Commun.* 5:3722. doi: 10.1038/ncomms4722
- Chen, K., and Rajewsky, N. (2007). The evolution of gene regulation by transcription factors and microRNAs. *Nat. Rev. Genet.* 8, 93–103. doi: 10.1038/nrg1990
- Chiang, G. C., Barua, D., Dittmar, E., Kramer, E. M., De Casas, R. R., and Donohue, K. (2013). Pleiotropy in the wild: the dormancy gene *DOG1* exerts cascading control on life cycles. *Evolution* 67, 883–893. doi: 10.1111/j.1558-5646.2012.01828.x
- Chou, C. H., Chang, N. W., Shrestha, S., Hsu, S. D., Lin, Y. L., Lee, W. H., et al. (2016). miRTarBase 2016: updates to the experimentally validated miRNA-target interactions database. *Nucleic Acids Res.* 44, D239–D247. doi: 10.1093/nar/gkv1258
- Cortijo, S., Wardenaar, R., Colome-Tatché, M., Gilly, A., Etcheverry, M., Labadie, K., et al. (2014). Mapping the epigenetic basis of complex traits. *Science* 343, 1145–1148. doi: 10.1126/science.1248127
- Cubas, P., Vincent, C., and Coen, E. (1999). An epigenetic mutation responsible for natural variation in floral symmetry. *Nature* 401, 157–161. doi: 10.1038/43657
- Cui, J., You, C., and Chen, X. (2016). The evolution of microRNAs in plants. *Curr. Opin. Plant Biol.* 35, 61–67. doi: 10.1016/j.pbi.2016.11.006
- Cuperus, J. T., Fahlgren, N., and Carrington, J. C. (2011). Evolution and functional diversification of MIRNA genes. *Plant Cell* 23, 431–442. doi: 10.1105/tpc.110.082784
- Czechowski, T., Stitt, M., Altmann, T., Udvardi, M. K., and Scheible, W. R. (2005). Genome-wide identification and testing of superior reference genes for transcript normalization in Arabidopsis. *Plant Physiol.* 139, 5–17. doi: 10.1104/pp.105.063743
- Dai, X., and Zhao, P. X. (2011). psRNATarget: a plant small RNA target analysis server. *Nucleic Acids Res.* 39, W155–W159. doi: 10.1093/nar/gkr319
- De La Torre, A. R., Li, Z., Van De Peer, Y., and Ingvarsson, P. K. (2017). Contrasting rates of molecular evolution and patterns of selection among gymnosperms and flowering plants. *Mol. Biol. Evol.* 34, 1363–1377. doi: 10.1093/molbev/msx069
- de Vega-Bartol, J. J., Simões, M., Lorenz, W. W., Rodrigues, A. S., Alba, R., Dean, J. F., et al. (2013). Transcriptomic analysis highlights epigenetic and transcriptional regulation during zygotic embryo development of *Pinus pinaster*. *BMC Plant Biol.* 13:123. doi: 10.1186/1471-2229-13-123
- Diez, C. M., Roessler, K., and Gaut, B. S. (2014). Epigenetics and plant genome evolution. *Curr. Opin. Plant Biol.* 18, 1–8. doi: 10.1016/j.pbi.2013.11.017
- Dong, C.-H., and Pei, H. X. (2014). Over-expression of *mir397* improves plant tolerance to cold stress in *Arabidopsis thaliana*. *J. Plant Biol.* 57, 209–217. doi: 10.1007/s12374-013-0490-y
- Down, R. H., Pelizzola, M., Schmitz, R. J., Lister, R., Down, J. M., Nery, J. R., et al. (2012). Widespread dynamic DNA methylation in response to biotic stress. *Proc. Natl. Acad. Sci. U.S.A.* 109, E2183–E2191. doi: 10.1073/pnas.1209329109
- Dugas, D. V., and Bartel, B. (2004). MicroRNA regulation of gene expression in plants. *Curr. Opin. Plant Biol.* 7, 512–520. doi: 10.1016/j.pbi.2004.07.011
- Fahlgren, N., Howell, M. D., Kasschau, K. D., Chapman, E. J., Sullivan, C. M., Cumbie, J. S., et al. (2007). High-throughput sequencing of Arabidopsis microRNAs: evidence for frequent birth and death of MIRNA genes. *PLoS ONE* 2:e219. doi: 10.1371/journal.pone.0000219
- Fahlgren, N., Jogdeo, S., Kasschau, K. D., Sullivan, C. M., Chapman, E. J., Laubinger, S., et al. (2010). MicroRNA gene evolution in *Arabidopsis lyrata* and *Arabidopsis thaliana*. *Plant Cell* 22, 1074–1089. doi: 10.1105/tpc.110.073999
- Fattash, I., Voss, B., Reski, R., Hess, W. R., and Frank, W. (2007). Evidence for the rapid expansion of microRNA-mediated regulation in early land plant evolution. *BMC Plant Biol.* 7:13. doi: 10.1186/1471-2229-7-13
- Fedak, H., Palusinska, M., Krzyczmonik, K., Brzezniak, L., Yatusевич, R., Pietras, Z., et al. (2016). Control of seed dormancy in Arabidopsis by a cis-acting noncoding antisense transcript. *Proc. Natl. Acad. Sci. U.S.A.* 113, E7846–E7855. doi: 10.1073/pnas.1608827113
- Felippes, F. F., Schneeberger, K., Dezulian, T., Huson, D. H., and Weigel, D. (2008). Evolution of *Arabidopsis thaliana* microRNAs from random sequences. *RNA* 14, 2455–2459. doi: 10.1261/rna.1149408
- Finet, C., Berne-Dedieu, A., Scutt, C. P., and Marlétaz, F. (2013). Evolution of the ARF gene family in land plants: old domains, new tricks. *Mol. Biol. Evol.* 30, 45–56. doi: 10.1093/molbev/mss220
- Fischerova, L., Fischer, L., Vondráková, Z., and Vágner, M. (2008). Expression of the gene encoding transcription factor PaVP1 differs in *Picea abies* embryogenic lines depending on their ability to develop somatic embryos. *Plant Cell Rep.* 27, 435–441. doi: 10.1007/s00299-007-0469-6
- Flynt, A. S., and Lai, E. C. (2008). Biological principles of microRNA-mediated regulation: shared themes amid diversity. *Nat. Rev. Genet.* 9, 831–842. doi: 10.1038/nrg2455
- Hodgins, K. A., Yeaman, S., Nurkowski, K. A., Rieseberg, L. H., and Aitken, S. N. (2016). Expression divergence is correlated with sequence evolution but not positive selection in conifers. *Mol. Biol. Evol.* 33, 1502–1516. doi: 10.1093/molbev/msw032
- Hornoy, B., Pavy, N., Gérardi, S., Beaulieu, J., and Bousquet, J. (2015). Genetic adaptation to climate in white spruce involves small to moderate allele frequency shifts in functionally diverse genes. *Genome Biol. Evol.* 7, 3269–3285. doi: 10.1093/gbe/evv218
- Hornstein, E., and Shomron, N. (2006). Canalization of development by microRNAs. *Nat. Genet.* 38, S20–S24. doi: 10.1038/ng1803
- Huang, X. Q., Schmitt, J., Dorn, L., Griffith, C., Effgen, S., Takao, S., et al. (2010). The earliest stages of adaptation in an experimental plant population: strong selection on QTLs for seed dormancy. *Mol. Ecol.* 19, 1335–1351. doi: 10.1111/j.1365-294X.2010.04557.x
- Huo, H., Wei, S., and Bradford, K. J. (2016). *DELAY OF GERMINATION1 (DOG1)* regulates both seed dormancy and flowering time through microRNA pathways. *Proc. Natl. Acad. Sci. U.S.A.* 113, E2199–E2206. doi: 10.1073/pnas.1600558113
- Jagadeeswaran, G., Li, Y. F., and Sunkar, R. (2014). Redox signaling mediates the expression of a sulfate-deprivation-inducible microRNA395 in Arabidopsis. *The Plant Journal* 77, 85–96. doi: 10.1111/tj.12364
- Jaramillo-Correa, J. P., Beaulieu, J., and Bousquet, J. (2004). Variation in mitochondrial DNA reveals multiple distant glacial refugia in black spruce (*Picea mariana*), a transcontinental North American conifer. *Mol. Ecol.* 13, 2735–2747. doi: 10.1111/j.1365-294X.2004.02258.x
- Johannes, F., Porcher, E., Teixeira, F. K., Saliba-Colombani, V., Simon, M., Agier, N., et al. (2009). Assessing the impact of transgenerational epigenetic variation on complex traits. *PLoS Genet.* 5:e1000530. doi: 10.1371/journal.pgen.1000530
- Johnsen, Ø., Fosdal, C. G., Nagy, N., Molmann, J., Daehlen, O. G., and Skroppa, T. (2005). Climatic adaptation in *Picea abies* progenies is affected by the temperature during zygotic embryogenesis and seed maturation. *Plant Cell Environ.* 28, 1090–1102. doi: 10.1111/j.1365-3040.2005.01356.x
- Jones-Rhoades, M. W., and Bartel, D. P. (2004). Computational identification of plant MicroRNAs and their targets, including a stress-induced miRNA. *Mol. Cell* 14, 787–799. doi: 10.1016/j.molcel.2004.05.027
- Källman, T., Chen, J., Gyllenstrand, N., and Lagercrantz, U. (2013). A significant fraction of 21-nucleotide small RNA originates from phased degradation of resistance genes in several perennial species. *Plant Physiol.* 162, 741–754. doi: 10.1104/pp.113.214643
- Kanehisa, M., and Goto, S. (2000). KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Res.* 28, 27–30. doi: 10.1093/nar/28.1.27
- Kawashima, C. G., Yoshimoto, N., Maruyama-Nakashita, A., Tsuchiya, Y. N., Saito, K., Takahashi, H., et al. (2009). Sulphur starvation induces the expression of microRNA-395 and one of its target genes but in different cell types. *The Plant Journal* 57, 313–321. doi: 10.1111/j.1365-313X.2008.03690.x
- Kinoshita, N., Wang, H., Kasahara, H., Liu, J., Macpherson, C., Machida, Y., et al. (2012). IAA-Ala Resistant3, an evolutionarily conserved target of miR167, mediates Arabidopsis root architecture changes during high osmotic stress. *Plant Cell* 24, 3590–3602. doi: 10.1105/tpc.112.097006
- Koornneef, M., Alonso-Blanco, C., Bentsink, L., Blankestijn-De Vries, H., Debeajon, I., Hanhart, C. J., et al. (2000). *The Genetics of Seed Dormancy in Arabidopsis thaliana*. Wallingford, CT: CAB International.
- Krämer, U. (2015). Planting molecular functions in an ecological context with *Arabidopsis thaliana*. *eLife* 4:e06100. doi: 10.7554/eLife.06100
- Kvaalen, H., and Johnsen, Ø. (2008). Timing of bud set in *Picea abies* is regulated by a memory of temperature during zygotic and somatic embryogenesis. *New Phytol.* 177, 49–59. doi: 10.1111/j.1469-8137.2007.02222.x

- Leitch, A. R., and Leitch, I. J. (2012). Ecological and genetic factors linked to contrasting genome dynamics in seed plants. *New Phytol.* 194, 629–646. doi: 10.1111/j.1469-8137.2012.04105.x
- Lelandais-Brière, C., Naya, L., Sallet, E., Calenge, F., Frugier, F., Hartmann, C., et al. (2009). Genome-wide *Medicago truncatula* small RNA analysis revealed novel microRNAs and isoforms differentially regulated in roots and nodules. *Plant Cell* 21, 2780–2796. doi: 10.1105/tpc.109.068130
- Lewsey, M. G., Hardcastle, T. J., Melnyk, C. W., Molnar, A., Valli, A., Urlich, M. A., et al. (2016). Mobile small RNAs regulate genome-wide DNA methylation. *Proc. Natl. Acad. Sci. U.S.A.* 113:E801. doi: 10.1073/pnas.1515072113
- Li, A. L., and Mao, L. (2007). Evolution of plant microRNA gene families. *Cell Res.* 17, 212–218. doi: 10.1038/sj.cr.7310113
- Lindow, M., and Krogh, A. (2005). Computational evidence for hundreds of non-conserved plant microRNAs. *BMC Genomics* 6:119. doi: 10.1186/1471-2164-6-119
- Lira-Medeiros, C. F., Parisod, C., Fernandes, R. A., Mata, C. S., Cardoso, M. A., and Ferreira, P. C. G. (2010). Epigenetic variation in mangrove plants occurring in contrasting natural environment. *PLoS ONE* 5:e10326. doi: 10.1371/journal.pone.0010326
- Liu, X. D., Zhang, H., Zhao, Y., Feng, Z. Y., Li, Q., Yang, H. Q., et al. (2013). Auxin controls seed dormancy through stimulation of abscisic acid signaling by inducing ARF-mediated ABI3 activation in Arabidopsis. *Proc. Natl. Acad. Sci. U.S.A.* 110, 15485–15490. doi: 10.1073/pnas.1304651110
- Liu, Y., and El-Kassaby, Y. A. (2017a). Regulatory cross-talk between microRNAs and hormone signalling cascades controls phenotypical variations: a case study in seed dormancy modulations during seed set of *Arabidopsis thaliana*. *Plant Cell Rep.* 36, 705–717. doi: 10.1007/s00299-017-2111-6
- Liu, Y., and El-Kassaby, Y. A. (2017b). Landscape of fluid sets of hairpin-Derived 21-/24-nt-long small RNAs at seed set uncovers special epigenetic features in *Picea glauca*. *Genome Biol. Evol.* 9, 82–92. doi: 10.1093/gbe/evw283
- Liu, Y., Müller, K., El-Kassaby, Y. A., and Kermodé, A. R. (2015). Changes in hormone flux and signaling in white spruce (*Picea glauca*) seeds during the transition from dormancy to germination in response to temperature cues. *BMC Plant Biol.* 15:292. doi: 10.1186/s12870-015-0638-7
- Liu, Y., Wang, T., and El-Kassaby, Y. A. (2016). Contributions of dynamic environmental signals during life-cycle transitions to early life-history traits in lodgepole pine (*Pinus contorta* Dougl.). *Biogeosciences* 13, 2945–2958. doi: 10.5194/bg-13-2945-2016
- Li, W. X., Oono, Y., Zhu, J., He, X. J., Wu, J. M., Iida, K., et al. (2008). The Arabidopsis NFYA5 transcription factor is regulated transcriptionally and posttranscriptionally to promote drought resistance. *Plant Cell* 20, 2238–2251. doi: 10.1105/tpc.108.059444
- Maher, C., Stein, L., and Ware, D. (2006). Evolution of Arabidopsis microRNA families through duplication events. *Genome Res.* 16, 510–519. doi: 10.1101/gr.4680506
- Mapes, G., Rothwell, G. W., and Haworth, M. T. (1989). Evolution of seed dormancy. *Nature* 337, 645–646. doi: 10.1038/337645a0
- Mastrangelo, A. M., Marone, D., Laidò, G., De Leonardi, A. M., and De Vita, P. (2012). Alternative splicing: enhancing ability to cope with stress via transcriptome plasticity. *Plant Science* 185, 40–49. doi: 10.1016/j.plantsci.2011.09.006
- May, P., Liao, W., Wu, Y., Shuai, B., McCombie, W. R., Zhang, M. Q., et al. (2013). The effects of carbon dioxide and temperature on microRNA expression in Arabidopsis development. *Nat. Commun.* 4:2145. doi: 10.1038/ncomms3145
- Michod, R. E., and Roze, D. (2001). Cooperation and conflict in the evolution of multicellularity. *Heredity* 86, 1–7. doi: 10.1046/j.1365-2540.2001.00808.x
- Navarro, L., Dunoyer, P., Jay, F., Arnold, B., Dharmasiri, N., Estelle, M., et al. (2006). A plant miRNA contributes to antibacterial resistance by repressing auxin signaling. *Science* 312, 436–439. doi: 10.1126/science.1126088
- Navarro, L., Jay, F., Nomura, K., He, S. Y., and Voinnet, O. (2008). Suppression of the microRNA pathway by bacterial effector proteins. *Science* 321, 964–967. doi: 10.1126/science.1159505
- Nawrocki, E. P., Burge, S. W., Bateman, A., Daub, J., Eberhardt, R. Y., Eddy, S. R., et al. (2015). Rfam 12.0: updates to the RNA families database. *Nucleic Acids Res.* 43, D130–D137. doi: 10.1093/nar/gku1063
- Nosaka, M., Itoh, J. I., Nagato, Y., Ono, A., Ishiwata, A., and Sato, Y. (2012). Role of transposon-derived small RNAs in the interplay between genomes and parasitic DNA in rice. *PLoS Genet.* 8:e1002953. doi: 10.1371/journal.pgen.1002953
- Nozawa, M., Miura, S., and Nei, M. (2012). Origins and evolution of microRNA genes in plant species. *Genome Biol. Evol.* 4, 230–239. doi: 10.1093/gbe/evs002
- Nystedt, B., Street, N. R., Wetterbom, A., Zuccolo, A., Lin, Y. C., Scofield, D. G., et al. (2013). The Norway spruce genome sequence and conifer genome evolution. *Nature* 497, 579–584. doi: 10.1038/nature12211
- Oh, T. J., Wartell, R. M., Cairney, J., and Pullman, G. S. (2008). Evidence for stage-specific modulation of specific microRNAs (miRNAs) and miRNA processing components in zygotic embryo and female gametophyte of loblolly pine (*Pinus taeda*). *New Phytol.* 179, 67–80. doi: 10.1111/j.1469-8137.2008.02448.x
- Oksanen, J., Blanchet, F. G., Kindt, R., Legendre, P., Minchin, P. R., O'hara, R., et al. (2015). *Vegan: Community Ecology Package*. Version 2.3-2. <https://CRAN.R-project.org/package=vegan>
- Piriyaopongsa, J., and Jordan, I. K. (2008). Dual coding of siRNAs and miRNAs by plant transposable elements. *RNA* 14, 814–821. doi: 10.1261/rna.916708
- Prunier, J., Verta, J. P., and Mackay, J. J. (2016). Conifer genomics and adaptation: at the crossroads of genetic diversity and genome function. *New Phytol.* 209, 44–62. doi: 10.1111/nph.13565
- Reyes, J. L., and Chua, N. H. (2007). ABA induction of miR159 controls transcript levels of two MYB factors during Arabidopsis seed germination. *The Plant Journal* 49, 592–606. doi: 10.1111/j.1365-313X.2006.02980.x
- Rhoades, D. P., Reinhart, B. J., Lim, L. P., Burge, C. B., Bartel, B., and Bartel, M. P. (2002). Prediction of plant microRNA targets. *Cell* 110, 513–520. doi: 10.1016/S0092-8674(02)00863-2
- Richards, C. L., Schrey, A. W., and Pigliucci, M. (2012). Invasion of diverse habitats by few Japanese knotweed genotypes is correlated with epigenetic differentiation. *Ecol. Lett.* 15, 1016–1025. doi: 10.1111/j.1461-0248.2012.01824.x
- Rubio-Somoza, I., and Weigel, D. (2011). MicroRNA networks and developmental plasticity in plants. *Trends Plant Sci.* 16, 258–264. doi: 10.1016/j.tplants.2011.03.001
- Rueda, A., Barturen, G., Lebrón, R., Gómez-Martín, C., Alganza, A., Oliver, J. L., et al. (2015). sRNAtoolbox: an integrated collection of small RNA research tools. *Nucleic Acids Res.* 43, W467–W473. doi: 10.1093/nar/gkv555
- Schneider, H., Schuettelpelz, E., Pryer, K. M., Cranfill, R., Magallon, S., and Lupia, R. (2004). Ferns diversified in the shadow of angiosperms. *Nature* 428, 553–557. doi: 10.1038/nature02361
- Schulz, B., Eckstein, R. L., and Durka, W. (2014). Epigenetic variation reflects dynamic habitat conditions in a rare floodplain herb. *Mol. Ecol.* 23, 3523–3537. doi: 10.1111/mec.12835
- Sigman, M. J., and Slotkin, R. K. (2016). The first rule of plant transposable element silencing: location, location, location. *Plant Cell* 28, 304–313. doi: 10.1105/tpc.15.00869
- Sparks, E., Wachsman, G., and Benfey, P. N. (2013). Spatiotemporal signalling in plant development. *Nat. Rev. Genet.* 14, 631–644. doi: 10.1038/nrg3541
- Springthorpe, V., and Penfield, S. (2015). Flowering time and seed dormancy control use external coincidence to generate life history strategy. *Elife* 4:e05557. doi: 10.7554/eLife.05557
- Sun, J., Zhou, M., Mao, Z. T., and Li, C. X. (2012). Characterization and evolution of microRNA genes derived from repetitive elements and duplication events in plants. *PLoS ONE* 7:e34092. doi: 10.1371/journal.pone.0034092
- Sunkar, R., Li, Y. F., and Jagadeeswaran, G. (2012). Functions of microRNAs in plant stress responses. *Trends Plant Sci.* 17, 196–203. doi: 10.1016/j.tplants.2012.01.010
- Takuno, S., Ran, J. H., and Gaut, B. S. (2016). Evolutionary patterns of genetic DNA methylation vary across land plants. *Nature Plants* 2:15222. doi: 10.1038/nplants.2015.222
- Taylor, R. S., Tarver, J. E., Hiscock, S. J., and Donoghue, P. C. (2014). Evolutionary history of plant microRNAs. *Trends Plant Sci.* 19, 175–182. doi: 10.1016/j.tplants.2013.11.008
- Tiwari, S. B., Hagen, G., and Guilfoyle, T. (2003). The roles of auxin response factor domains in auxin-responsive transcription. *Plant Cell* 15, 533–543. doi: 10.1105/tpc.008417
- Tollefsrud, M. M., Kissling, R., Gugerli, F., Johnsen, Ø., Skråppa, T., Cheddadi, R., et al. (2008). Genetic consequences of glacial survival and postglacial colonization in Norway spruce: combined analysis of mitochondrial DNA and fossil pollen. *Mol. Ecol.* 17, 4134–4150. doi: 10.1111/j.1365-294X.2008.03893.x

- Ulmasov, T., Hagen, G., and Guilfoyle, T. J. (1999). Dimerization and DNA binding of auxin response factors. *Plant J.* 19, 309–319. doi: 10.1046/j.1365-313X.1999.00538.x
- Vanneste, K., Maere, S., and Van De Peer, Y. (2014). Tangled up in two: a burst of genome duplications at the end of the Cretaceous and the consequences for plant evolution. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 369:20130353. doi: 10.1098/rstb.2013.0353
- Vidigal, D. S., Marques, A. C., Willems, L. A., Buijs, G., Méndez-Vigo, B., Hilhorst, H. W., et al. (2016). Altitudinal and climatic associations of seed dormancy and flowering traits evidence adaptation of annual life cycle timing in *Arabidopsis thaliana*. *Plant Cell Environ.* 39, 1737–1748. doi: 10.1111/pce.12734
- Vidal, E. A., Araus, V., Lu, C., Parry, G., Green, P. J., Coruzzi, G. M., et al. (2010). Nitrate-responsive miR393/AFB3 regulatory module controls root system architecture in *Arabidopsis thaliana*. *Proc. Natl. Acad. Sci. USA.* 107, 4477–4482. doi: 10.1073/pnas.0909571107
- Wagner, G. P., and Altenberg, L. (1996). Perspective: complex adaptations and the evolution of evolvability. *Evolution* 50, 967–976. doi: 10.1111/j.1558-5646.1996.tb02339.x
- Wang, H. L. V., Dinwiddie, B. L., Lee, H., and Chekanova, J. A. (2015). Stress-induced endogenous siRNAs targeting regulatory intron sequences in *Brachypodium*. *RNA* 21, 145–163. doi: 10.1261/rna.047662.114
- Wang, X. Q., and Ran, J. H. (2014). Evolution and biogeography of gymnosperms. *Mol. Phylogenet. Evol.* 75, 24–40. doi: 10.1016/j.ympev.2014.02.005
- Warren, R. L., Keeling, C. I., Yuen, M. M., Raymond, A., Taylor, G. A., Vandervalk, B. P., et al. (2015). Improved white spruce (*Picea glauca*) genome assemblies and annotation of large gene families of conifer terpenoid and phenolic defense metabolism. *Plant J.* 83, 189–212. doi: 10.1111/tpj.12886
- Wendel, J. F., Jackson, S. A., Meyers, B. C., and Wing, R. A. (2016). Evolution of plant genome architecture. *Genome Biol.* 17:37. doi: 10.1186/s13059-016-0908-1
- Willmann, M. R., and Poethig, R. S. (2007). Conservation and evolution of miRNA regulatory programs in plant development. *Curr. Opin. Plant Biol.* 10, 503–511. doi: 10.1016/j.pbi.2007.07.004
- Wu, C. I., Shen, Y., and Tang, T. (2009). Evolution under canalization and the dual roles of microRNAs: a hypothesis. *Genome Res.* 19, 734–743. doi: 10.1101/gr.084640.108
- Xia, R., Xu, J., Arikait, S., and Meyers, B. C. (2015). Extensive families of miRNAs and PHAS loci in Norway spruce demonstrate the origins of complex phasiRNA networks in seed plants. *Mol. Biol. Evol.* 32, 2905–2918. doi: 10.1093/molbev/msv164
- Xiang, D., Venglat, P., Tibiche, C., Yang, H., Risseuw, E., Cao, Y., et al. (2011). Genome-wide analysis reveals gene expression and metabolic network dynamics during embryo development in *Arabidopsis*. *Plant Physiol.* 156, 346–356. doi: 10.1104/pp.110.171702
- Yakovlev, I. A., Fossdal, C. G., and Johnsen, Ø. (2010). MicroRNAs, the epigenetic memory and climatic adaptation in Norway spruce. *New Phytol.* 187, 1154–1169. doi: 10.1111/j.1469-8137.2010.03341.x
- Yeaman, S., Hodgins, K. A., Lotterhos, K. E., Suren, H., Nadeau, S., Degner, J. C., et al. (2016). Convergent local adaptation to climate in distantly related conifers. *Science* 353, 1431–1433. doi: 10.1126/science.aaf7812
- Zhang, B. H., Pan, X. P., Cannon, C. H., Cobb, G. P., and Anderson, T. A. (2006). Conservation and divergence of plant microRNA genes. *Plant J.* 46, 243–259. doi: 10.1111/j.1365-313X.2006.02697.x
- Zhang, Y., Zhang, S., Han, S., Li, X., and Qi, L. (2012). Transcriptome profiling and in silico analysis of somatic embryos in Japanese larch (*Larix leptolepis*). *Plant Cell Rep.* 31, 1637–1657. doi: 10.1007/s00299-012-1277-1
- Zhang, X., Zhao, H., Gao, S., Wang, W. C., Katiyar-Agarwal, S., Huang, H. D., et al. (2011). *Arabidopsis* Argonaute 2 regulates innate immunity via miRNA393(\*)-mediated silencing of a Golgi-localized SNARE gene, MEMB12. *Mol. Cell* 42, 356–366. doi: 10.1016/j.molcel.2011.04.010
- Zhang, Y., Zhu, X. J., Chen, X., Song, C. N. A., Zou, Z. W., Wang, Y. H., et al. (2014). Identification and characterization of cold-responsive microRNAs in tea plant (*Camellia sinensis*) and their targets using high-throughput sequencing and degradome analysis. *BMC Plant Biol.* 14, 271. doi: 10.1186/s12870-014-0271-x

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The reviewer RN and handling Editor declared their shared affiliation.

Copyright © 2017 Liu and El-Kassaby. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.