Tree Genetics & Genomes (2015) 11:806
DOI 10.1007/s11295-014-0806-1

ORIGINAL PAPER

Computational predictions and expression patterns of conserved microRNAs in loblolly pine (*Pinus taeda*)

Christina R. Quinn · Rie Iriyama · Danilo D. Fernando

Received: 12 April 2013 / Revised: 12 September 2014 / Accepted: 24 October 2014 / Published online: 12 November 2014

Abstract MiRNAs are mostly known for their expressions in sporophytic (diploid) tissues from various angiosperms, and only limited information is available on those expressed in gametophytic (haploid) tissues. Even more limited information is available for gymnosperms, particularly those that are expressed in gametophytic tissues. Homology search of the loblolly pine expressed sequence tag (EST) database using known miRNAs from *Arabidopsis thaliana* and *Picea abies* revealed 12 miRNAs previously unreported and/or uncharacterized in loblolly pine (*Pinus taeda*). Their precursor and mature sequences and secondary structures were obtained using computational approaches. PCR was used to confirm their expressions using samples that represent the sporophyte (i.e., needles) and gametophyte (i.e., mature and germinated pollen) phases of the loblolly pine life cycle. Our results showed that all 12 miRNAs were expressed in the needles, three of which were detected only in the needles, two were expressed only in the needles and germinated pollen, and seven were expressed in all of the three tissues examined. None was expressed only in the mature and germinated pollen, but nine had expressions in both or either of these samples. Target predictions suggest that these newly identified and characterized loblolly pine miRNAs regulate messenger RNAs (mRNAs) that are involved in gene regulation, metabolism, and signal transduction. This study demonstrates the usefulness of computational approach in identifying conserved miRNAs in species with unsequenced genomes like loblolly pine and reveals that conserved miRNAs can be used to differentiate sporophytic and gametophytic tissues and between stages of the male gametophyte.

Keywords Germinated pollen · Male gametophyte · Needles · Mature pollen

Communicated by J. L. Wegrzyn

Electronic supplementary material The online version of this article (doi:10.1007/s11295-014-0806-1) contains supplementary material, which is available to authorized users.

C. R. Quinn · R. Iriyama · D. D. Fernando (✉)
Department of Environmental and Forest Biology, State University of New York College of Environmental Science and Forestry, One Forestry Drive, Syracuse, NY 13210, USA
e-mail: dfernando@esf.edu

Introduction

Gymnosperms and angiosperms exhibit substantial differences in their evolution, morphology, and development (Singh 1978; Gifford and Foster 1989; Fernando et al. 2005, 2010; Williams 2008, 2012). Their difference has also been found to be reflected in the nature of their small RNAs (Dolgosheina et al. 2008; Morin et al. 2008). Gymnosperms have been shown to mostly express a very diverse population of 21 nt small RNAs, whereas angiosperms produce significant amounts of 24 nt small RNAs (Dolgosheina et al. 2008; Morin et al. 2008). Included in the diverse 21 nt small RNA class in gymnosperms are numerous deeply conserved miRNAs. MiRNAs are endogenous and non-coding RNAs that are involved in post-transcriptional gene regulation related to development and physiology (Hamilton et al. 2002; Xie et al. 2004; Henderson et al. 2006; Axtell and Bartel 2005). Many conserved miRNAs target the same messenger RNA (mRNA) families as their homologs in other plant lineages including bryophytes, lycophytes, monilophytes, and angiosperms (Dolgosheina et al. 2008; Axtell and Bartel 2005). Some of these homologs have been identified through the use of computational methods tailored to predict miRNAs in species with unsequenced genomes. By systematically mining expressed sequence tag (EST) databases, conserved miRNAs have been identified in several gymnosperms (Axtell and Bartel 2005; Zhang et al. 2005, 2006a). Such analyses provide

further evidence that many miRNA families are highly conserved across the major land plant lineages. However, much remains to be discovered in gymnosperms, particularly in species with immense ecological and economic values such as loblolly pine (*Pinus taeda*).

Loblolly pine, a key representative of the conifers, is an economically and ecologically important native forest tree species in the southern USA. It is relatively fast growing and amenable to intensive tree farming and has become the cornerstone of the country's forest products industry accounting for 58 % of the harvested industrial wood and 16 % of the world's timber (Connor and Hartsell 2002; Prestemon and Abt 2002; Wear and Greis 2002). With markets including lumber, pulp, and paper, loblolly pine has also recently become a major bioenergy feedstock used for lignocellulosic ethanol production (Frederick et al. 2008). As a frequent or predominant species on about 25 million hectares of non-planted forest land, loblolly pine also provides food and cover for numerous animal species (Connor and Hartsell 2002). Likewise, loblolly pine is expected to play a major role in efforts to curb greenhouse gas levels via carbon sequestration (Perlack et al. 2005; Gough and Seiler 2004). What is learned in loblolly pine will likely be applicable to the Pinaceae and across much of the Coniferophyta, whose members exhibit high levels of genetic similarity (Brown et al. 2001; Krutovsky et al. 2004). It is also likely to impact the study of gymnosperms as a whole, as the genus *Pinus* alone represents about 20 % of all known gymnosperms (Richardson and Rundel 1998). Despite the importance of loblolly pine, support for genomics research in conifers has lagged significantly behind most major agricultural crops and model species (Neale and Wheeler 2004). The complete genome sequencing of loblolly pine has only recently been funded by the National Institute of Food and Agriculture, and thus, the availability of genetic sequence information is still limited. Nevertheless, many large public EST discovery projects have provided a resource of sequence information on the loblolly pine genome that should be utilized to their fullest potential.

In angiosperms, numerous miRNAs have been identified and characterized from various sporophytic tissues and many of these have also been shown to be expressed in the gametophytic tissues (Le Trionnaire et al. 2011; Wei et al. 2011). In gymnosperms and conifers, in particular, most of the information on types and expression patterns of miRNAs are also based on the examination of sporophytic tissues (e.g., needles, young stems, wood, roots, etc.). There are only two reports on miRNA gametophytic expressions in gymnosperms and both are from loblolly pine, one on the female gametophyte (Oh et al. 2008) and the other is our work on the male gametophyte (Quinn et al. 2014). Whereas our previous report is based on microRNA array analysis of gametophyte development, the present study utilizes a computational approach to identify conserved miRNAs in loblolly pine and compares their expression patterns in sporophytic and gametophytic tissues. Unlike the reports on miRNA expression in angiosperm male gametophytes where only the pre-pollination stages (developing to mature or ungerminated pollen) were examined, this study compares miRNA expressions between two stages of the male gametophyte that represent pre- and post-pollination stages, particularly the mature and germinated pollen, respectively. As gymnosperms remain understudied in many areas of plant biology, this study contributes to our knowledge of their miRNA profiles, particularly for loblolly pine. This study shows that most of the conserved miRNAs identified and characterized through computational predictions from loblolly pine are expressed in both sporophytic and gametophytic tissues. Interestingly, some of these miRNAs show differential expressions between mature and germinated pollen.

Materials and methods

MiRNA reference set used for EST analysis

Conserved miRNAs from *Arabidopsis thaliana* and *Picea abies* were retrieved from MiRBase (Release 16). To avoid redundancy, only one sequence of identical miRNAs was used, and those miRNAs reported in MiRBase for loblolly pine (*P. taeda*) were removed. A total of 208 sequences were used as the reference set to search for conserved miRNAs in the loblolly pine NCBI EST database.

Prediction of conserved miRNAs and their precursor sequences

A summary of the bioinformatic pipeline used to search for conserved loblolly pine miRNA homologs of known miRNAs from *A. thaliana* and *P. abies* is presented in Fig. 1. The mature miRNA sequences from the reference set were used as query sequences against the loblolly pine EST database using BLASTn (Altschul et al. 1990). The parameters used were as follows: EST database was selected, maximum target sequences were set to 1000, expected threshold was raised to 1000, and word size was set to 7; the remaining parameters were left at default settings. If matching sequences were less than the length of the mature miRNA sequence, the non-aligned segments were manually inspected to determine the total number of matching nucleotides. Sequences with four or less mismatches with the known mature miRNA sequences were used to search against the loblolly pine EST database using BLASTn to identify repeated sequences. The non-redundant sequences were then searched against the protein database to remove protein-coding ESTs. The remaining sequences were considered as candidate miRNAs. The secondary structures of these candidates were predicted using RNAfold (Hofacker et al. 1994). The following criteria were

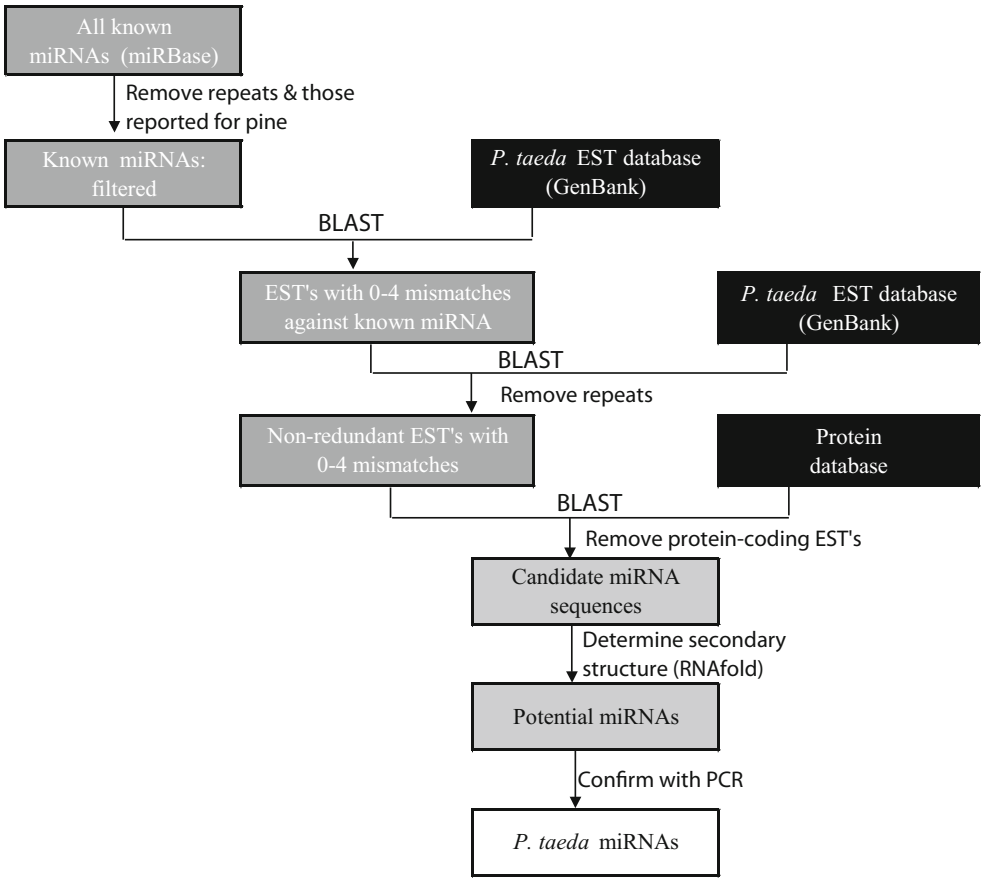


Fig. 1 Schematic of computational analysis used to predict conserved miRNAs in loblolly pine (*Pinus taeda*)

used for selecting potential miRNAs: (1) the sequence could fold into an appropriate stem-loop secondary structure at least 60 nucleotides in length, (2) the mature miRNA sequence was located within one arm of the hairpin, (3) there were less than six mismatches between the mature miRNA and the miRNA* sequence in the opposite arm, (4) no loops or breaks were present in the miRNA or miRNA* sequences, and (5) the minimal free energy (MFE) was less than −15 kcal/mol. This screening was used to ensure that the predicted miRNAs fit the criteria proposed by Ambros et al. (2003). Sequences that formed stable stem-loop structures were considered as potential loblolly pine miRNAs.

Validation of predicted conserved miRNAs

PCR was used to determine if the predicted miRNAs were expressed in loblolly pine. Total RNA was isolated from mature (0 h in culture) and germinated (48 h in culture) loblolly pine pollen using TRIzol Reagent (Invitrogen, Inc.). RNA isolation was done three times from three different biological samples. A pollen grain is considered germinated if the pollen tube is at least as long as the diameter of the grain and 48 h represents the time when at least 95 % has germinated. Pollen sources and germination protocol were based on the study of Fernando et al. (1997). Pollen cultures for germination and 3-year-old loblolly pines were grown under the same temperature (25 °C). Samples used for ungerminated pollen came straight from the freezer, which were collected and stored at the time that they were dispersed and thus represent gene expression pattern at pollination (i.e., 0 h in culture). MasterPure Plant RNA Purification Kit (Epicentre, Madison, WI) was used to isolate total RNA from young needles, also from three different biological samples. The RNA was resuspended in 25 μl RNase-free water, and quantity and quality were measured using a ND-1000 spectrophotometer (NanoDrop Technologies, Wilmington, DE) and electrophoresis, respectively. Complementary DNA was synthesized from total RNA isolated from ungerminated pollen, germinated pollen, and needles using the ProtoScript AMV First Strand complementary DNA (cDNA) Synthesis Kit (New England Biolabs, Ipswich, MA). Primers were designed using the web-based program Primer3 (http://frodo.wi.mit.edu/ primer3/) to be 18–25 base pairs in length, a GC content between 40 and 60 %, and a melting temperature between 50 and 65 °C. The primers used were specific to areas of the stem-loop region that flanked the 5′ and 3′ end of the mature miRNA sequence (Supplement 1). The PCR reaction was set up in 20 μl composed of 10 μl Taq Hot Start PCR mix (Qiagen, Valencia, CA), 0.4 μl of each primer (10 μM), 2 μl cDNA template, and the rest with DEPC-treated water.

An 18S ribosomal RNA (rRNA) fragment was also amplified as a positive control. The specificity of all the primers was verified through PCR. All PCR reactions were done in triplicate using a MiniOpticon thermal cycler (Bio-Rad, Hercules, CA) with the following cycling conditions: 95 °C for 5 min, followed by 35 cycles of 95 °C for 30 s, 60 °C for 30 s, and 72 °C for 1 min, and a final elongation at 72 °C for 10 min. Ten microliters of each reaction was mixed with 2 μl loading dye, loaded into a 2 % agarose gel, and run at 110 V until the dye front had migrated two thirds of the way down the gel. The gel was visualized under UV light and photographed.

Target predictions for conserved miRNAs

The mRNA targets of the loblolly pine miRNAs were predicted using the psRNA-Target program (Dai and Zhao 2011) using the *Pinus* (Pine) DFCI Gene Index Release 9 as the sequence library for target searches and default parameters. All predicted target sequences were evaluated based on a mismatch scoring scheme of Jones-Rhoades and Bartel (2004). Using a 20 base sequence window around the miRNA:mRNA binding site, 0.5 points were assigned to each G:U wobble, 1 point to each non-G:U wobble mismatch, and 2 points to each indel/bulged nucleotide. In addition, as strong Watson-Crick base pairing seems to be necessary in the 5′ seed region (positions 2–9) of the miRNA and at positions 10 and 11 as they are assumed cleavage sites in the miRNA sequence, no more than one mismatch was allowed in the complementary region of the miRNA at positions 1–9, and no mismatches were accepted at positions 10 and 11 (Yin et al. 2008; Unver and Budak 2009; Yu et al. 2011). Up to three additional mismatches were permitted between positions 12 and 21, but no more than two continuous mismatches within this region. Predicted targets with a score of 3.0 or less (Jones-Rhoades and Bartel 2004) and the base-pairing rules stated above were accepted as authentic targets with a high confidence.

Validation of potential target genes

Potential targets of pta-miR161, pta-miR172, and pta-miR399 in loblolly pine were chosen because sufficient sequence information was available for PCR primer design and experimentation. The Plant Gene Index accession numbers of the potential targets that were useful in this study were TC133159, CX712622, and TC115098 for pta-miR161, pta-miR172, and pta-miR399, respectively. Forward and reverse primers for TC133159 were 5′-GCGAAATGGTGGAGGA GTAA-3′ and 5′-GATTCAGCGTTGTGACAGGA-3′, respectively. Primers for CX712622 were 5′-GGGGCCATAA CCAAAGAGTA-3′ and 5′-TCACTTTTGGCTTTGCAGTG -3′. Primers for TC115098 were 5′-GGGGCCATAACCAA AGAGTA-3′ and 5′- CAATTACCCCTTCCCCATCT-3′. The 18 s rRNA was used as the reference gene (using 5′-CATGGC CGTT CTTAGTTGGT-3′ and 5′-GAGTTGATGACACGCG CT TA-3′ as forward and reverse primers, respectively). qRT-PCR was done in triplicate for each combination of primers and cDNA templates including the reference gene using QuantiFast SYBR Green PCR Kit (Qiagen, Valencia, CA) following the protocol as described in our previous work (Quinn et al. 2014).

Results

Characteristics of predicted conserved miRNAs

The unique mature miRNA sequences from *A. thaliana* and *P. abies* were used as reference set to query the loblolly pine EST database. After filtering out repeated ESTs and protein-coding sequences, the sequences that remained had four or fewer mismatches with the miRNAs. These were then submitted to RNAfold to determine their secondary structures. Twelve miRNAs were identified to be conserved in loblolly pine following NCBI EST database search. Ten of the miRNAs were homologous to *A. thaliana* and two miRNAs were homologous to *P. abies* (Table 1). These 12 miRNAs represented nine miRNA families in loblolly pine. Six families (miR161, miR172, miR399, miR776, miR3701, and miR3704) had only one member each, while three families (miR157, miR164, and miR166), had two members each. This indicated that each of the latter three miRNAs is encoded from at least two different genes. The sequences upstream and downstream of each mature miRNA were submitted to RNAfold and all were determined to form stable stem-loop miRNA precursor structures (Fig. 2) and that met the criteria for miRNA identification following (Ambros et al. 2003). The length of the miRNA precursors varied from 66 to 187 nucleotides. For each pre-miRNA, the mature miRNA sequence was found in the 5′ or 3′ arm. Four of the 12 miRNAs were located at the 5′ end, while eight were located at the 3′ end. In addition, all precursors were greater than 60 nt in length with no loops or breaks in the miRNA or miRNA* region, have four or less mismatches between the miRNA and miRNA* sequences, and a minimum free energy of less than −18.20 kcal/mol (Table 1).

Expression of potential conserved miRNAs

The predicted miRNAs identified from the EST loblolly pine database suggest that they were expressed in the loblolly pine genome. To validate their expressions, PCR analyses were performed for the 12 miRNAs using total RNA from mature and germinated pollen. PCR analysis was also carried out using total RNA from young needles to compare the expression of the miRNAs between gametophyte and sporophyte

Table 1 Features of conserved loblolly pine (*Pinus taeda*) miRNAs

miRNA	Mature sequence	Mature miRNA length (nt)	Pre-miRNA length (nt)	EST acc. no.	Arm location	MFE (kcal/mol)
Loblolly pine miRNAs homologous to *Arabidopsis thaliana*						
pta-miR157a	-UGACAGAAGAUAGAGAGCAC	21	83	DR088695	5′	−36.20
pta-miR157b	-UGACAGAAGAUAGAGAGCAC	21	83	CX651705	5′	−40.70
pta-miR161.2	UUAAUGCAUUAAAAGUGAAUG	21	80	CF396469	5′	−25.00
pta-miR164a	UGGAGAAGCAUGGCACUGGAG	21	148	DR744906	3′	−38.40
pta-miR164b	UGGAGAAGCAUGGCACUGGAG	21	176	DR686108	3′	−54.40
pta-miR166a	UCGGACCAGGCUUCAUUCCCC	21	88	DR178396	3′	−50.80
pta-miR166b	UCGGACCAGGCUUCAUUCCCC	21	102	DR012818	3′	−52.10
pta-miR172	AACAACUUGAUGAUGAUGCAU	21	66	DR165657	3′	−18.20
pta-miR399	UGG-AAAGGAGUGUUGCCCUG	21	174	DT631247	3′	−47.20
pta-miR776	UCUAAGUGUUCUAUUGAUGGU	21	88	AW62658	3′	−33.90
Loblolly pine miRNAs homologous to *Picea abies*						
pta-miR3701	UGAACAAUGCCCACCCUUCAUC	22	187	CO163803	3′	−72.92
pta-miR3704	GGUCUCGGUGGAGUUGGAAAAA	22	147	CO201834	5′	−56.30

Underscores indicate nucleotide substitutions or indels in comparison with the homologs from *A. thaliana* or *P. abies*

tissues. The primers used were based on pre-miRNA sequences. The results showed that all 12 predicted miRNAs exhibited expression in loblolly pine but in different tissues and at varying levels (Fig. 3), and therefore, they can now be appropriately referred to as valid miRNAs. Pta-miR161 and pta-miR164b were expressed in the needles and germinated pollen, but no expression was detected in mature pollen (Fig. 3). Interestingly, no expression was observed for pta-miR164a in both mature and germinated pollen. For pta-miR166a, pta-miR166b, pta-miR172, pta-miR399, and pta-miR776, expressions were detected in the needles and both of the male gametophyte stages. Expressions of pta-miR157a and pta-miR157b were strongly detected in the needles and mature pollen, but they have faint expressions in germinated pollen. Interestingly, the expressions of these two miRNAs in mature pollen were the most intense among all the miRNAs detected in this study. Finally, pta-miR3701 and pta-miR3704 (and pta-miR164a as indicated above) were only detected in the needles (Fig. 3).

Predicted targets of potential conserved miRNAs

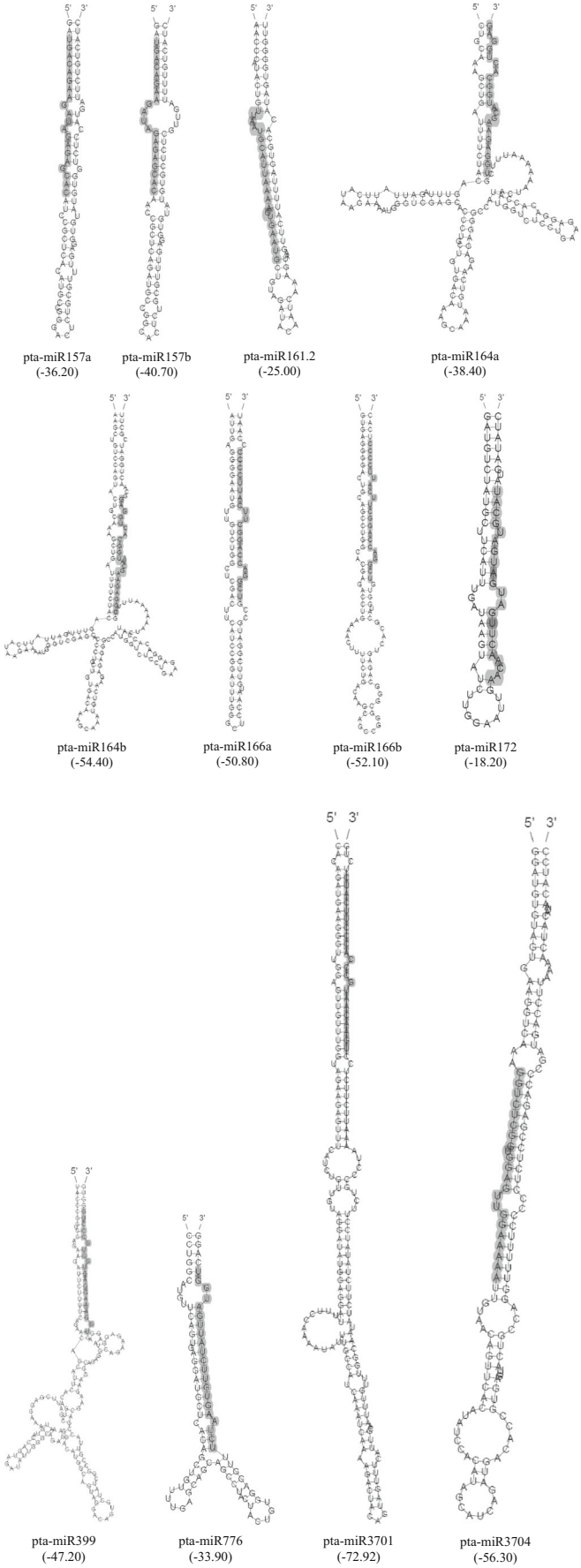
The web-based program psRNA Target was used to predict the complementary mRNA sequences, target sites, and functions of the potential loblolly pine miRNAs using the *Pinus* DFCI Gene Index as source of sequence information to search for targets. MiRNAs and their predicted mRNA targets had four or fewer mismatches with no gaps or bulges (Supplement 2). The G:U non-canonical pair was treated as a mismatch, as per the screening criteria used. For those miRNAs that commonly silence mRNA targets through translational inhibition (Yin et al. 2008; Unver and Budak 2009; Yu et al. 2011; Aukerman and Sakai 2003; Brodersen et al. 2008), mismatches were allowed at nucleotide positions 10 and 11. A total of 16 potential mRNA targets were identified for seven of the nine miRNA families, most of which had homologs that have known functions in other species (Table 2). Members of the pta-miR157 family were predicted to target mRNAs coding for a nucleic acid-binding protein and alcohol dehydrogenase, whereas pta-miR166 targets HD-Zip III proteins and pta-miR172 targets UBX domain-containing and zinc binding proteins (Table 2). Sequences of the targets of two (miR3701 and miR3704) of the nine miRNA families were identified, but no function can be inferred (Table 2).

Validation of potential target genes

The potential targets of computationally predicted loblolly pine miRNAs were confirmed using three representative target genes (Fig. 4). These genes were expressed in both of the male gametophyte stages (mature and germinated pollen) and in the needles (Fig. 4). Expression levels were higher in the needles than in any of the gametophytic stages examined. There is also generally a higher expression in mature pollen compared to germinated pollen.

Discussion

Many miRNAs are deeply conserved in land plants, some of which are perfectly conserved across the major land plant



◀ **Fig. 2** Mature and precursor sequences and the stem-loop structures of miRNAs predicted in loblolly pine (*Pinus taeda*) by EST analysis. Mature miRNA sequences are highlighted in *gray* and minimum-free energies are indicated in *parentheses*

lineages (Axtell and Bartel 2005; Floyd and Bowman 2004). Therefore, miRNA genes in one taxa may be found as homologs in another taxa. This suggests a powerful strategy to predict new miRNAs by computational methods (Zhang et al. 2006a, c, d, 2009; Mishra and Lobiyal 2011; Frazier and Zhang 2011; Chambers and Shuai 2009; Grant-Downton et al. 2009a; Le Trionnaire and Twell 2010; Willmann and Poethig 2007; Axtell and Bowman 2008; Rhoades et al. 2002; Elhiti and Stasolla 2009), which is particularly relevant considering the relatively large amounts of sequence information available from various groups of plants, mostly from angiosperms. In fact, our computational prediction has identified five miRNAs (miR157a, miR157b, miR161, miR399, and miR776) which are known in angiosperms, but not in gymnosperms, until this report. The mature miR156 sequence is identical to the mature sequence of miR157 (except for one

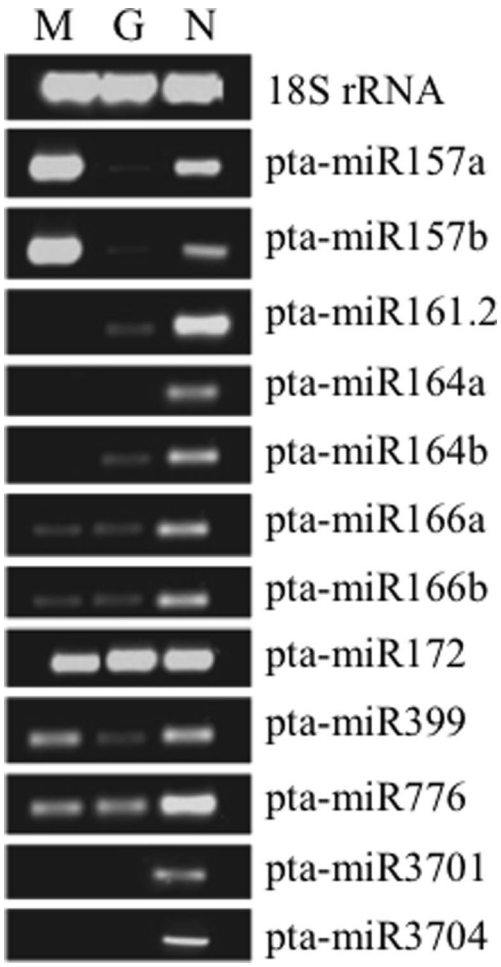


Fig. 3 Expression of conserved miRNAs in loblolly pine revealed through PCR of the following samples: mature pollen (*M*), germinated pollen (*G*), and young needles (*N*)

Table 2 Potential targets of conserved loblolly pine (*Pinus taeda*) miRNAs

miRNA	Target EST	Score	Annotation of targeted protein	*E* value
pta-miR157	TC118278	2.5	NP_191557, nucleic acid-binding protein (*Arabidopsis thaliana*)	$2e^{-22}$
	TC154061	3.0	AAS18891, alcohol dehydrogenase (*Zea mays*)	$1e^{-24}$
pta-miR161	TC140959	2.5	ABK1348, unknown (*Picea sitchensis*)	$2e^{-62}$
	TC151202	2.5	Unknown	
	TC133159[a]	2.5	ABR17115, unknown (*Picea sitchensis*)	0.0
pta-miR164	Not found in loblolly pine		Unknown	
pta-miR166	TC145397	2.0	ABG73245, class III HD-Zip protein (*Pinus taeda*)	$1e^{-78}$
	TC139716	2.0	ABG73246, class III HD-Zip protein (*Pinus taeda*)	0.0
	CX713974	2.0	ABG73247, class III HD-Zip protein (*Pinus taeda*)	$6e^{-108}$
	BX253120	2.0	ADV04325, class III homeodomain leucine zipper protein (*Picea glauca*)	$4e^{-118}$
pta-miR172	TC115543[b]	2.0	ABR18217, unknown (*Picea sitchensis*)	$1e^{-173}$
	TC140475[b]	2.0	NP_567262, UBX domain-containing protein (*Arabidopsis thaliana*)	$4e^{-31}$
	CX712622[a]	3.0	ACG24900, zinc-binding protein (*Zea mays*)	$9e^{-62}$
pta-miR399	TC129934	3.0	ABK25258, unknown (*Picea sitchensis*)	$2e^{-26}$
	TC115098[a]	3.0	NP_001183236, hypothetical protein (*Zea mays*)	$6e^{-50}$
pta-miR776	Not found in loblolly pine		Unknown	
pta-miR3701	TC154085	0.5	Unknown	
pta-miR3704	TC125400	1.0	Unknown	

[a] Source of target gene analyzed for qRT-PCR

[b] Predicted to be silenced through translational inhibition

nucleotide difference) and sometimes complementary to the same sites within the target mRNAs (Rhoades et al. 2002). Nevertheless, miR157 has not yet been reported in loblolly pine or any gymnosperm. We also report two other miRNAs (miR3701 and miR3704) from loblolly pine that still has no homolog in angiosperms or any other plant lineages. However, these two miRNAs are known from *P. abies* (Yakovlev et al. 2010) and *Pinus densata* (Wan et al. 2012). It remains to be seen if these two miRNAs are conifer-specific. Although miRNA analyses have been done in gymnosperms, they are relatively limited and so this study expands our knowledge of the identities and characterizations of miRNAs from this less derived and lesser explored group of seed plants.

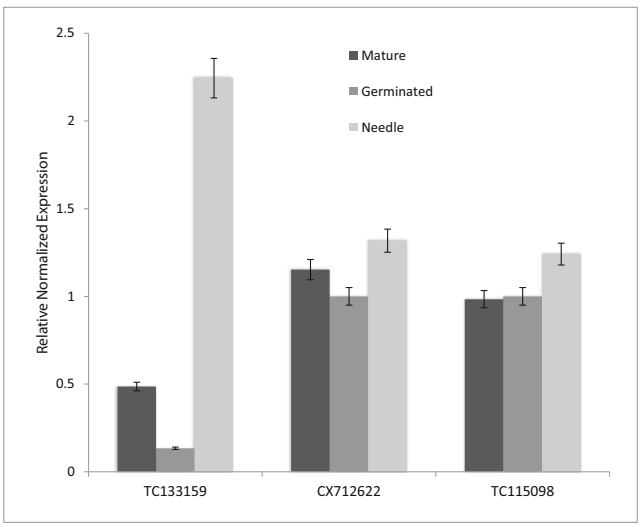


Fig. 4 QRT-PCR analyses of miRNA target genes in loblolly pine. TC133159, CX712622, and TC115098 are accession numbers based on the Pine DFCI Gene Index and represent the target genes of pta-miR161, pta-miR172, and pta-miR399, respectively

In seed plants, the expression of miRNAs has been predominantly examined from various sporophytic (diploid) tissues and only limited studies have reported expressions in gametophytic (haploid) tissues. There is one report each on miRNA expression in the female (Oh et al. 2008) and male (Quinn et al. 2014) gametophytes of loblolly pine. There are a few reports on male gametophytes in angiosperms, but unlike in loblolly pine (Quinn et al. 2014 and this paper), these are only from pre-pollination stages, particularly developing to mature (ungerminated) pollen (Wei et al. 2011; Chambers and Shuai 2009; Grant-Downton et al. 2009a; Le Trionnaire and Twell 2010). Therefore, much remains to be discovered regarding the presence and function of miRNAs from the male gametophytes of seed plants. The present study reports nine conserved miRNAs (miR157a, miR157b, miR161, miR164b, miR166a, miR166b, miR172, miR399, and miR776) with expressions in mature and/or germinated loblolly pine pollen, with at least four of these exhibiting differential expressions between these two developmental stages. This is the first report of the expression of these nine miRNAs in mature and germinated pollen from any seed plant. Pta-miR166a, pta-miR166b and pta-miR166c were identified by Lu et al. (2007) through computational prediction but no validation

was done. Our computational prediction identified only two of these (pta-miR166a and pta-miR166b) and from other loblolly pine ESTs. Therefore, we included these two miRNAs in our report to validate their occurrence in loblolly pine through PCR and describe their expression in the needles and mature and germinated pollen. The expression of pta-miR172 was shown using loblolly pine female gametophytes (Oh et al. 2008), but this loblolly pine miRNA is not found in MiRBase. Therefore, our study provides its mature and precursor sequences, stem-loop structure, and expression in three different tissues.

The mature sequences of the 12 conserved miRNAs in loblolly pine range in size from 19 to 22 nt, but most of these are 21 nt long, and thus, our results provide support to the predominance of this RNA class size in gymnosperms (Dolgosheina et al. 2008; Morin et al. 2008; Quinn et al. 2014). These 12 miRNAs represent nine families. A total of 16 potential target sequences were identified for seven of the families where three of these have known functions including nucleic acid-binding protein and alcohol dehydrogenase for miR157, class III HD-Zip protein for miR166, and UBX domain-containing and zinc-binding proteins for miR172. For four of the miRNAs (miR161, miR399, miR3701, and miR3704), target sequences are available but no function can be inferred. The functions of several of the predicted target genes were unknown in loblolly pine, and so the reported functions of targets for homologous plant miRNAs were considered. This is possible since the targets have been shown to be similar or functionally related in many conserved miRNAs across different plant species (Axtell and Bartel 2005; Zhang et al. 2006a; Willmann and Poethig 2007; Axtell and Bowman 2008). Therefore, based on information from *Arabidopsis* and maize, four of the families have been shown to target transcription factors, including squamosa promoter-binding proteins, NAC (NAM, ATAF1/2, and CUC2) domain-containing proteins, HD-Zip III proteins, and APETALA2 for miRNA families miR157, miR164, miR166, and miR172, respectively (Richardson and Rundel 1998; Willmann and Poethig 2007; Axtell and Bowman 2008). Other targets include pentatricopeptide repeat proteins, phosphatase transporter, and serine/threonine kinase for families miR161, miR399, and miR776, respectively (Le Trionnaire et al. 2011; Rhoades et al. 2002; Zhang et al. 2006c).

All 12 miRNAs identified in the present study were expressed in the needles, and nine of these were expressed in either or both mature or germinated loblolly pine pollen. Three of the 12 miRNAs (miR164a, miR3701, and miR3704) appear to have specific expression in the needles, but none of the miRNAs was found to be specific to the either mature or germinated pollen or both. This is not surprising considering that the available loblolly pine ESTs were mostly derived from cDNAs isolated from vegetative or sporophytic tissues, particularly stem, wood, needle, and root samples. Although pollen cones have also been used, these samples likely represent limited sequence information related to male gametophyte development. Five of the conserved miRNAs (miR166a, miR166b, miR172, miR399, and miR776) were found to be expressed in all the three tissues examined and at generally the same level. This indicates that these miRNAs regulate genes that are involved in housekeeping functions that are common to sporophyte and gametophyte development. MiR166 is also one of the miRNA families that were expressed in both sporophytes and gametophytes in rice, which were also interpreted to have a housekeeping role (Wei et al. 2011). MiR166 is known to target transcription factors that are unique to plants including class III HD-Zip transcription factors (Elhiti and Stasolla 2009; Zhou et al. 2007; Carlsbecker et al. 2010; Ariel et al. 2007), whereas miR172 targets UBX domain-containing and zinc-binding proteins and APETALA2 (Rancour et al. 2004; Alberts et al. 2009). These two genes have been validated as targets for miR172 in *Pinus resinosa*, *P. taeda, Taxus chinensis*, and *Larix leptolepis* (Axtell and Bartel 2005; Chambers and Shuai 2009; Qiu et al. 2009; Zhang et al. 2010). In addition, miR172 has been predicted through EST analysis to be present in *Picea engelmannii*×*Picea sitchensis*, *Picea glauca*, and *P. sitchensis* (Zhang et al. 2005, 2006c). MiR172 has been included in this report to reiterate its occurrence in the gametophytes since it is not listed as one of the loblolly pine miRNAs in miRBase, which may explain why previous reports did not identify it for loblolly pine (Zhang et al. 2006a). Although it is interesting that miR172 is expressed in both the female (Oh et al. 2008) and male (this report) gametophytes of loblolly pine, it is also expressed in the needles and thus further supports a housekeeping function. The targets of miR399 and miR776 are unknown in loblolly pine, but a homology search suggests that the former targets mRNAs coding for phosphatase transporters which are involved in the regulation of phosphate metabolism, as reported for maize (Zhang et al. 2005; Floyd and Bowman 2004), whereas the latter targets serine/threonine kinase, as reported for Arabidopsis (Fahlgren et al. 2007).

Three miRNAs (miR164a, miR3701, and miR3704) are expressed only in the needles, but their specific functions in loblolly pine are unknown. Homology search indicates that miR164a targets NAC (NAM, ATAF1/2, and CUC2) domain proteins that constitute one of the largest families of plant-specific transcription factors that occur in a wide range of plants (Olsen et al. 2005). Homology search for targets of miR3701 and miR3704 failed to reveal any possible function. Nevertheless, the needle-specific expression of these two miRNAs suggests that their targets have regulatory roles in vegetative or sporophyte development, but not in male gametophyte development. These two miRNAs have also been identified from the needles of *P. abies* (Yakovlev et al. 2010) and *P. densata* (Wan et al. 2012), with no reported function.

Therefore, our results validate and extend the expression of these two miRNAs in loblolly pine needles. It remains to be seen whether these two miRNAs are conifer-specific.

Pta-miR157a and pta-miR157b were highly expressed in the needles and mature pollen, but very low expression can be detected in germinated pollen. Nevertheless, these miRNAs are the most highly expressed of all the miRNAs detected in the present study. Interestingly, these miRNAs are also two of the most highly expressed miRNAs in the mature pollen of *Arabidopsis* and rice (Le Trionnaire et al. 2011; Chambers and Shuai 2009; Grant-Downton et al. 2009a). However, there is no information available on their expression in germinated pollen from any angiosperm species. In loblolly pine, pta-miR157a and pta-miR157b show different expression levels between mature and germinated pollen. These miRNAs are predicted to target a nucleic acid-binding protein and alcohol dehydrogenase (Grant-Downton et al. 2009a, b; Dolferus et al. 1994; Tadege and Kuhlemeier 1997) and squamosa promoter-binding protein (Willmann and Poethig 2007; Tadege and Kuhlemeier 1997; Grant-Downton et al. 2009b). In *Arabidopsis*, miR157 has been found to be enriched in the mature pollen as compared to sporophyte tissues (Grant-Downton et al. 2009b), which is similar to the trend obtained in the present study. It is interesting to note that there are miRNAs that are highly expressed in the mature pollen and then have very low expression in germinated pollen, two developmental stages that are only separated by about 48 h in loblolly pine. This trend is reversed for pta-miR161.2 and pta-miR164b. These four miRNAs provide evidence that the mature and germinated pollen are regulated differently in loblolly pine, which supports our previous report using microRNA array (Quinn et al. 2014).

The functions of the targets of pta-miR161.2 and pta-miR164 in loblolly pine are also unknown, but both miRNAs have known functions in other species. Their targets include pentatricopeptide repeat proteins for miR161 (Willmann and Poethig 2007) and NAC domain-containing proteins for miR164 (Chen 2009; Schwarz et al. 2008; Sieber et al. 2007). In contrast to our results for loblolly pine, miR161 is relatively abundant in *Arabidopsis* mature pollen and expression is slightly higher in pollen than in the sporophyte (Chambers and Shuai 2009; Grant-Downton et al. 2009a, b). While miR164 is not expressed in mature loblolly pine pollen, this miRNA is expressed in *Arabidopsis* pollen although at very low levels (Chambers and Shuai 2009). This is interesting considering the different developmental patterns involved in gametophyte formation in gymnosperms and angiosperms (Singh 1978; Gifford and Foster 1989; Fernando et al. 2005, 2010; Williams 2008, 2012).

The potential targets of pta-miR161.2, pta-miR172, and pta-miR399 in loblolly pine that were examined include TC133159, CX712622, and TC115098, respectively (Table 2, Fig. 4). CX712622 encodes a zinc-binding protein, whereas the other two have no known function. Our results show that these three target genes were expressed in all the tissues examined, with the expression in the needle tissues higher than any of the gametophytic tissues. This also shows a positive correlation between computational prediction of miRNAs and occurrence of their target genes in loblolly pine. Interestingly, the expression level of pta-miR161.2 negatively correlates with the expression level of its target gene. Anti-correlated miRNA and target gene expression are a typical trend in miRNAs. However, in a few cases, miRNAs and target gene expression levels are generally similar, as in pta-miR172 and pta-miR399. According to Alonso-Peral et al. (2012), miRNA-mediated silencing is not solely determined by miRNA abundance and target transcript levels but also through other mechanisms, but yet unknown.

While much remains to be discovered regarding the roles of miRNAs in loblolly pine or gymnosperms in general, this study demonstrates the usefulness of computational predictions in identifying conserved miRNAs in species with unsequenced genomes. Likewise, this study has revealed that most of these conserved miRNAs identified are expressed in both the sporophytic and gametophytic tissues, although with varying patterns of expression levels. It is also interesting that computationally predicted conserved miRNAs can be used to characterize reproductive development, particularly male gametophyte development and differentiate two stages (mature and germinated pollen) that are distinctly temporally separated, i.e., pre- and post-pollination. Whereas additional studies are necessary to confirm many of the targets predicted here, our results present an array of genes whose roles may help in our understanding of male gametophyte development in seed plants, particularly the post-pollination stage where information is very limited.

Data archiving statement The precursor and mature sequences of the miRNAs identified in this study have been submitted to the MiRBase Registry.

References

Alberts SM, Sonntag C, Schafer A, Wolf DH (2009) UBX4 modulates CDC48 activity and influences degradation of misfolded proteins of the endoplasmic reticulum. J Biol Chem 284:16082–16089

Alonso-Peral MM, Sun C, Millar AA (2012) MicroRNA159 can act as a switch or tuning microRNA independently of its abundance in Arabidopsis. Plos One 7:e34751

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. J Mol Biol 215:403–410

Ambros V, Bartel B, Bartel D, Burge CB, Carrington JC, Chen XM, Dreyfuss G, Eddy SR, Griffiths-Jones S, Marshall M, Matzke M, Ruvkun G, Tuschl T (2003) A uniform system for microRNA annotation. RNA 9:277–279

Ariel FD, Manavella PA, Dezar CA, Chan RL (2007) The true story of the HD-Zip family. Trends Plant Sci 12:419–426

Aukerman MJ, Sakai H (2003) Regulation of flowering time and floral organ identity by a microRNA and its APETALA2-like target genes. Plant Cell 15:2730–2741

Axtell MJ, Bartel DP (2005) Antiquity of microRNAs and their targets in land plants. Plant Cell 17:1658–1673

Axtell MJ, Bowman JL (2008) Evolution of plant microRNAs and their targets. Trends Plant Sci 13:343–349

Brodersen P, Sakvarelidze-Achard L, Bruun-Rasmussen M, Dunoyer P, Yamamoto YY, Sieburth L, Voinnet O (2008) Widespread translational inhibition by plant miRNAs and siRNAs. Science 320:1185–1190

Brown G, Kadel EI, Bassoni D, Kiehne K, Temesgen B, van Buijtenen J, Sewell M, Marshall K, Neale DB (2001) Anchored reference loci in loblolly pine (*Pinus taeda* L.) for integrating pine genomes. Genetics 159:799–809

Carlsbecker A, Lee J-Y, Roberts CJ, Dettmer J, Lehesranta S, Zhou J, Lindgren O, Moreno-Risueno MA, Vaten A, Thitamadee S, Campilho A, Sebastian J, Bowman JL, Helariutta Y, Benfey PN (2010) Cell signaling by microRNA165/6 directs gene dose-dependent root cell fate. Nature 465:316–321

Chambers C, Shuai B (2009) Profiling microRNA expression in *Arabidopsis* pollen using microRNA array and real-time PCR. BMC Plant Biol 9:87

Chen CC (2009) Small RNAs and their roles in plant development. Annu Rev Cell Dev Biol 25:21–44

Connor R, Hartsell A (2002) Forest area and conditions. In: Wear D, Gries J (eds) Southern Forest Resource Assessment. Gen. Tech. Rep. SRS-53. USDA Forest Service, Asheville, pp 357–401

Dai X, Zhao PX (2011) psRNATarget: a plant small RNA target analysis server. Nucleic Acid Res 39:W155–W159

Dolferus R, De Bruxelles G, Dennis ES, Peacock WJ (1994) Regulation of the *Arabidopsis Adh* gene by anaerobic and other environmental stresses. Ann Bot 74:301–308

Dolgosheina EV, Morin RD, Aksay G (2008) Conifers have a unique small RNA silencing signature. RNA 14:1508–1515

Elhiti M, Stasolla C (2009) Structure and function of homeodomain-leucine zipper (HD-Zip) proteins. Plant Signal Behav 4:86–88

Fahlgren N, Howell MD, Kasschau KD, Chapman EJ, Sullivan CM, Cumbie JS, Givan S, Law TF, Grant SR, Dangl JL, Carrington JC (2007) High-throughput sequencing of *Arabidopsis* microRNAs: evidence for frequent birth and death of *MIRNA* genes. PLoS One 2:e219

Fernando DD, Owens JN, von Aderkas P, Takaso T (1997) In vitro pollen tube growth and penetration of female gametophytes in Douglas fir (*Pseudotsuga menziesii*). Sex Plant Reprod 10:209–216

Fernando DD, Lazzaro MD, Owens JN (2005) Growth and development of conifer pollen tubes. Sex Plant Reprod 18:149–162

Fernando DD, Quinn CR, Brenner ED, Owens JN (2010) Male gametophyte development and evolution in extant gymnosperms. Int J Plant Dev Biol 4:47–63

Floyd SK, Bowman JL (2004) Gene regulation: ancient microRNA target sequences in plants. Nature 428:485–486

Frazier TP, Zhang B (2011) Identification of plant microRNAs using expressed sequence tag analysis. Methods Mol Biol 678:13–25

Frederick WJ, Lien S, Couchene C, Demartini N, Ragauskas A, Iisa K (2008) Production of ethanol from carbohydrates from loblolly pine: a technical and economic assessment. Bioresour Technol 99:5051–5057

Gifford EM, Foster AS (1989) Morphology and evolution of vascular plants. W. H. Freedman and Company, New York, 626 pp

Gough C, Seiler J (2004) Below ground carbon dynamics in loblolly pine (*Pinus taeda*) immediately following diammonium phosphate fertilization. Tree Physiol 24:845–851

Grant-Downton R, Hafidh S, Twell D, Dickinson HG (2009a) Small RNA pathways are present and functional in the angiosperm male gametophyte. Mol Plant 2:500–512

Grant-Downton R, Le Trionnaire G, Schmid R, Rodriguez-Enriquez J, Hafidh S, Mehdi S, Twell D, Dickinson H (2009b) MicroRNA and tasiRNA diversity in mature pollen of *Arabidopsis thaliana*. BMC Genomics 10:643

Hamilton A, Voinnet O, Chappell L, Baulcombe D (2002) Two classes of short interfering RNA in RNA silencing. EMBO J 21:4671–4679

Henderson IR, Zhang XY, Lu C, Johnson L, Meyers BC, Green PJ, Jacobsen SE (2006) Dissecting *Arabidopsis thaliana* DICER function in small RNA processing, gene silencing and DNA methylation patterning. Nat Genet 38:721–725

Hofacker IL, Fontana W, Stadler PF, Bonhoeffec S, Tacker M, Schuster P (1994) Fast folding and comparison of RNA secondary structures. Monatsh Chem 125:167–188

Jones-Rhoades MW, Bartel DP (2004) Computational identification of plant microRNAs and their targets, including a stress-induced miRNA. Mol Cell 14:787–799

Krutovsky K, Troggio M, Brown G, Jermstad K, Neale DB (2004) Comparative mapping in the Pinaceae. Genetics 168:447–461

Le Trionnaire G, Twell D (2010) Small RNAs in angiosperm gametophytes: from epigenetics to gamete development. Genes Dev 24:1081–1085

Le Trionnaire G, Grant-Downton R, Kourmpetli S, Dickinson H, Twell D (2011) Small RNA activity and function in angiosperm gametophytes. J Exp Bot 62:1601–1610

Lu S, Sun YH, Amerson H, Chiang VL (2007) MicroRNAs in loblolly pine (*Pinus taeda* L.) and their association with fusiform rust gall development. Plant J 51:1077–1098

Mishra AK, Lobiyal DK (2011) MiRNA prediction using computational approach. Adv Exp Med Biol 696:75–82

Morin RD, Askay G, Dolgosheina D, Ebhardt HA, Magrinio V, Mardiso ER, Cenk Sahinalp SC, Unrau PJ (2008) Comparative analysis of the small RNA transcriptomes of *Pinus contorta* and *Oryza sativa*. Genome Res 18:571–584

Neale DB, Wheeler N (2004) The Loblolly Pine Genome Project. http://dendrome.ucdavis.edu/NealeLab/lpgp/pdf/prospectus.pdf

Oh TJ, Wartell RM, Cairney J, Pullman GS (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

Olsen AN, Ernst HA, Leggio LL, Skriver K (2005) NAC transcription factors: structurally distinct, functionally diverse. Trends Plant Sci 10:79–87

Perlack R, Wright L, Turhollow A, Graham R, Stokes B, Erbach D (2005) Biomass as a feedstock for a bioenergy and bioproducts industry: the technical feasibility of a billion-ton annual supply. Biomass Program Overview. US Department of Energy/US Department of Agriculture

Prestemon J, Abt R (2002) Timber products supply and demand. In: Wear D, Gries J (eds) Southern Forest Resource Assessment. Gen. Tech. Rep. SRS-53. USDA Forest Service, Asheville, pp 299–326

Qiu DY, Pan XP, Wilson IW, Li FL, Liu M, Teng WJ, Zhang BH (2009) High throughput sequencing technology reveals that the taxoid elicitor methyl jasmonate regulates microRNA expression in Chinese yew (*Taxus chinensis*). Gene 436:37–44

Quinn CR, Iriyama I, Fernando DD (2014) Expression patterns of conserved microRNAs in the male gametophyte of loblolly pine (*Pinus taeda*). Plant Reprod 27:69–78

Rancour DM, Park S, Knight SD, Bednarek SY (2004) Plant UBX domain-containing protein 1, PUX1, regulates the oligomeric structure and activity of *Arabidopsis* CDC48. J Biol Chem 279:54264–54274

Rhoades MW, Reinhart BJ, Lim LP, Burge CB, Bartel B, Bartel D (2002) Prediction of plant microRNA targets. Cell 110:513–520

Richardson D, Rundel P (1998) Ecology and biogeography of *Pinus*: an introduction. In: Richardson D (ed) Ecology and biogeography of *Pinus*. Cambridge University Press, Cambridge

Schwarz S, Grande AV, Bujdoso N, Saedler H, Huijser P (2008) The microRNA regulated SBP-box genes SPL9 and SPL15 control shoot maturation in *Arabidopsis*. Plant Mol Biol 67:183–195

Sieber P, Wellmer F, Gheyselinck J, Riechmann JL, Meyerowitz EM (2007) Redundancy and specialization among plant microRNAs: role of the MIR164 family in developmental robustness. Development 134:1051–1060

Singh H (1978) Embryology of gymnosperms. Gebruder Borntraegar, Berlin, 302 pp

Tadege M, Kuhlemeier C (1997) Aerobic fermentation during tobacco pollen development. Mol Biol 35:343–354

Unver T, Budak U (2009) Conserved microRNAs and their targets in model grass species *Brachypodium distachyon*. Planta 230:659–669

Wan LC, Zhang H, Lu S, Zhang L, Qiu Z, Zhao Y, Zeng QY, Lin J (2012) Transcriptome-wide identification and characterization of miRNAs from *Pinus densata*. BMC Genomics 13:132

Wear D, Greis J (2002) Southern forest resource assessment: summary of findings. J For 100:6–14

Wei LQ, Yan LF, Wang T (2011) Deep sequencing on genome-wide scale reveals the unique composition and expression patterns of microRNAs in developing pollen of *Oryza sativa*. Genome Biol 12:R53

Williams JH (2008) Novelties of the flowering plant pollen tube underlie diversification of a key life history stage. PNAS 32:11259–11263

Williams JH (2012) Pollen tube growth rates and the diversification of flowering plant reproductive cycles. Int J Plant Sci 173:649–661

Willmann MR, Poethig RS (2007) Conservation and evolution of miRNA regulatory programs in plant development. Curr Opin Plant Biol 10:503–511

Xie ZX, Johansen LK, Gustafson AM, Kasschau KD, Lellis AD, Zilberman D, Jacobsen SE, Carrington JC (2004) Genetic and functional diversification of small RNA pathways in plants. PLoS Biol 2:642–652

Yakovlev IA, Fossdal CG, Johnsen O (2010) MicroRNAs, the epigenetic memory and climatic adaptation in Norway spruce. New Phytol 187:1154–1169

Yin ZJ, Li CH, Han ML, Shen FF (2008) Identification of conserved microRNAs and their target genes in tomato (*Lycopersicon esculentum*). Gene 414:60–66

Yu H, Song C, Jia Q, Wang C, Li F, Nicholas KK, Zhang X, Fang J (2011) Computational identification of microRNAs in apple expressed sequence tags and validation of their precise sequences by miR-RACE. Physiol Plant 141:56–70

Zhang BH, Pan XP, Wang QL, Cobb GP, Anderson TA (2005) Identification and characterization of new plant microRNAs using EST analysis. Cell Res 15:336–360

Zhang BH, Pan XP, Cannon CH, Cobb GP, Anderson TA (2006a) Conservation and divergence of plant microRNA genes. Plant J 46:243–259

Zhang BH, Pan XP, Cobb GP, Anderson TA (2006b) Plant microRNA: a small regulatory molecule with big impact. Dev Biol 289:3–16

Zhang B, Pan X, Anderson TA (2006c) Identification of 188 conserved maize microRNAs and their targets. FEBS Lett 580:3753–3762

Zhang WW, Luo YP, Gong X, Zeng WH, Li SG (2009) Computational identification of 48 potato microRNAs and their targets. Comput Biol Chem 33:84–93

Zhang S, Zhou J, Han S, Yang W, Li W, Wei H, Li X, Qi L (2010) Four abiotic stress-induced miRNA families differentially regulated in the embryonic and non-embryonic callus tissues of *Larix leptolepis*. Biochem Biophys Res Commun 398:355–360

Zhou G-K, Kubo M, Zhong R, Demura T, Ye Z-H (2007) Overexpression of miR165 affects apical meristem formation, organ polarity establishment and vascular development in *Arabidopsis*. Plant Cell Physiol 48:391–404