

Article

Further Mining and Characterization of miRNA Resource in Chinese Fir (*Cunninghamia lanceolata*)

Houyin Deng^{1,2}, Rong Huang¹, Dehuo Hu¹, Runhui Wang¹, Ruping Wei¹, Su Yan¹, Guandi Wu¹, Yuhan Sun² , Yun Li^{2,*}  and Huiquan Zheng^{1,*} 

¹ Guangdong Provincial Key Laboratory of Silviculture, Protection and Utilization, Guangdong Academy of Forestry, Guangzhou 510520, China

² National Engineering Research Center of Tree Breeding and Ecological Restoration, College of Biological Sciences and Technology, Beijing Forestry University, Beijing 100083, China

* Correspondence: yunli@bjfu.edu.cn (Y.L.); zhenghq@sinogaf.cn (H.Z.); Tel.: +86-20-8758-4306 (H.Z.)

Abstract: In this study, we aimed to expand the current miRNA data bank of Chinese fir (*Cunninghamia lanceolata* (Lamb.) Hook.) regarding its potential value for further genetic and genomic use in this species. High-throughput small RNA sequencing successfully captured 140 miRNAs from a Chinese fir selfing family harboring vigor and depressed progeny. Strikingly, 75.7% (n = 106) of these miRNAs have not been documented previously, and most (n = 105) of them belong to the novel set with 6858 putative target genes. The new datasets were then integrated with the previous information to gain insight into miRNA genetic architecture in Chinese fir. Collectively, a relatively high proportion (62%, n = 110) of novel miRNAs were found. Furthermore, we identified one MIR536 family that has not been previously documented in this species and four overlapped miRNA families (MIR159, MIR164, MIR171_1, and MIR396) from new datasets. Regarding the stability, we calculated the secondary structure free energy and found a relatively low R² value (R² < 0.22) between low minimal folding free energy (MFE) of pre-miRNAs and MFE of its corresponding mature miRNAs in most datasets. When in view of the conservation aspect, the phylogenetic trees showed that MIR536 and MIR159 sequences were highly conserved in gymnosperms.

Keywords: Chinese fir; miRNA resource; bioinformatics; free energy; evolution



Citation: Deng, H.; Huang, R.; Hu, D.; Wang, R.; Wei, R.; Yan, S.; Wu, G.; Sun, Y.; Li, Y.; Zheng, H. Further Mining and Characterization of miRNA Resource in Chinese Fir (*Cunninghamia lanceolata*). *Genes* **2022**, *13*, 2137. <https://doi.org/10.3390/genes13112137>

Academic Editor: Matthias Zytnicki

Received: 11 October 2022

Accepted: 15 November 2022

Published: 17 November 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Plant mature miRNAs are generated from stem-loop precursors by a Dicerlike (DCL) enzyme and range from 20 to 24 nucleotides (nt) in length [1]. Their precursor miRNA sequence can fold into a stable stem-loop structure with lower minimal folding free energy than other sRNAs [2]. Additionally, mature miRNAs typically combine with Argonaute proteins in order to form RNA-induced silencing complexes (RISCs) with stronger stable structure, which can silence the expression of complementary mRNAs [3]. These properties of miRNAs ensure the regular performance of the growth and development of the organism. Plant growth and development are depended on gene expression and regulation [4]. MicroRNAs (miRNAs) are versatile regulators of gene expression and play a pivotal role in the evolutionary process of plants [5,6]. Conifers dominate the world's forest ecosystems and are the most widely cultivated tree species [7]. The conifers are expected to provide important information regarding the evolution of highly conserved small regulatory RNAs [8]. An increasing number of miRNAs studies about growth and development are being carried out in coniferous trees from different genus, such as *Pinus* L. [8–11], *Picea* A. Dietrich [12], *Araucaria* Juss. [13] and *Ginkgo* L. [14]. To our knowledge, some miRNAs studies have also been conducted on several Cupressaceae trees, such as *Platycladus orientalis* (L.) Franco and Chinese fir (*C. lanceolata*) [15–17].

Chinese fir, an evergreen coniferous tree that occupies approximately 25% of plantations in southern China, has great commercial and ecological importance [18]. The

different clones within this species harbor a broad range of growth variation, which means that a plethora of valuable genetic resources remain hidden in this species [19]. With the widespread accessibility of next-generation sequencing technology, small RNA sequencing and identification were performed for Chinese fir [20]. The miRNA expression profiles of this species were primarily conducted in seedling (stem and leaf), seed dormancy, vascular cambium activity, and vascular cambium development [16,20–22]. However, there is still a considerable lack of corresponding research on selfed progeny growth and development associated with miRNAs in Chinese fir, as selfing often results in inbreeding depression, marked by reduced survival and fertility, stunted growth and poor stress tolerance and even death [23,24]. Given this, the research on the field of selfing related to plant growth and development has received extensive attention [25–27].

Chinese fir is generally regarded as an outcrossing species with an effective mechanism to encourage outcrossing and prevent selfing [28]. Previous studies have shown that Chinese fir selfing often leads to evident inbreeding depression, which is manifested in selfed progeny by lower seed yield, reduced germination rate, stunted seedling growth and even death [29–31]. However, our previous study showed that a special Chinese fir clone, name ‘cx569’, produced a selfed progeny population with lower inbreeding depression in the seedling stage [31]. The seedlings presented two distinct phenotypes, showing higher vigor and lower vigor in terms of seedling height, number of roots, and fresh weight, etc. Therefore, the selfed progeny of the special material ‘cx569’ could provide additional genetic basis to fill the gaps on miRNA related to selfing in Chinese fir. Given this, we used a selfed family from this special material for miRNA mining and aimed to expand the Chinese fir miRNA resource through small RNA sequencing. Furthermore, we integrated our datasets with the available Chinese fir miRNA information to profile their genetic architecture.

2. Materials and Methods

2.1. Plant Materials

The self-pollination of clone cx569 with the characteristic of self-fertility was conducted in a 2.5th generation seed orchard of Chinese fir in the Xiaokeng State Forest Farm (Dongguan, Guangdong Province, China; N24°70′, E113°81′, at an altitude of 328–339 m). The methods of pollination were adopted from a previous study [31]. The fresh seeds were collected from the self-pollination cones when they reached maturity. For collecting germinated seeds, the fresh seeds were surface sterilized and placed on seed germination boxes covered with gauzes moistened with sterile water, and the seed germination boxes were placed in a climatic incubator at 25 °C under a 24 h dark photoperiod with a humidity of 75%. Subsequently, the germinated seeds were then transferred to peat soil mixed with 20% expanded-perlite and maintained in a growth chamber at 25 °C under a 16/8 h light/dark photoperiod with a humidity of 75%. At 5 months old, the growth vigor of these progeny populations varied greatly with extremely depressed seedlings and normal seedlings (Figure S1). The normal seedlings were characterized by green stems and leaves, no visible pests or diseases and vigorous growth. Meanwhile, the depressed seedlings were shorter, dwarf, weak, chlorotic, reddish, leathery and subsequently withered compared to normal seedlings. To investigate whether there were differences in genetic basis between depressed seedlings and normal seedlings populations, we divided selfed progeny of cx569 into 2 groups according to seedling height. To reduce the error in the selection of depressed and normal seedlings, seven depressed seedlings (named ‘S21L’, ‘S38L’, ‘S42L’, ‘S189L’, ‘S218L’, ‘S220L’, and ‘S221L’) from the seedlings below the average seedling height and ten normal seedlings (named ‘S29L’, ‘S64L’, ‘S98L’, ‘S116L’, ‘S134L’, ‘S157L’, ‘S193L’, ‘S217L’, ‘S231L’, and ‘S256L’) from seedlings above the average seedling height were selected by using the random function of Microsoft Excel.

2.2. RNA Extraction, RNA Library Construction and Functional Annotation

Total RNA was extracted from the above-ground parts (stem and leaf) of seven depressed seedlings and ten normal seedlings using the RNeasy Plant Mini Kit (Qiagen, Hilden, Germany). The RNA concentration and integrity were evaluated using a Nanodrop 2000 spectrophotometer (Thermo Fisher Scientific, Wilmington, DE, USA) and an Agilent 2100 Bioanalyzer (Agilent Technologies, Santa Clara, CA, USA), respectively.

In order to obtain raw data, library construction and deep sequencing were conducted on an Illumina system HiSeq XTen platform provided by Biomarker Biotechnologies Corporation (Beijing, China, www.biocloud.net (accessed on 15 January 2022)). These raw data (17 libraries) were uploaded in the Sequence Read Archive (SRA) database of the NCBI with accession number PRJNA899478. After removal of low-quality reads, short reads, and adapters, the clean reads were generated for de novo assembly using Trinity [32]. Sequence alignment was performed using Bowtie software [33]. The mapped reads consisting of 48,755 unigenes were used for further analysis. The unigene function was annotated based on the following six databases: NCBI nonredundant protein sequences (nr), Kyoto Encyclopedia of Genes and Genomes (KEGG) and Gene Ontology (GO), Swiss-Prot, Protein family (Pfam) and Clusters of Orthologous Groups of proteins (KOG/COG). Additionally, the obtained Chinese fir transcriptome database was used for reference sequences in pre-miRNA prediction and identification of potential miRNA targets.

2.3. Identification of miRNAs and Prediction of Their Target Gene

The clean reads were mapped to various RNA databases including the Silva, GtRNAdb, Rfam, and Rfam databases for filtering ribosomal RNA (rRNA), transfer RNA (tRNA), small nuclear RNA (snRNA), small nucleolar RNA (snoRNA) and repeats using Bowtie. The remain reads were aligned to the Chinese fir mRNA transcriptome by Bowtie. To identify conserved miRNA, the drive reads were aligned with currently mature miRNA sequences from all plant species in the miRbase Sequence Database (<https://www.mirbase.org/>, (Release 22.1) accessed on 1 August 2022) using a BLASTn search with up to one mismatch [34]. Sequences with identity to other species were identified as known miRNA. The unannotated reads were processed for novel miRNA identification by miRDeep2 with adjusted parameters [35]. To further reveal known and novel miRNA precursors, the mapped reads were used to extract potential precursors using miRDeep2 with adjusted parameters [35]. The precursors were aligned to short reads, and statistical analyses of read distributions were generated for each miRNA gene for evaluation of miRNA. To quantify the abundance of miRNA, the transcripts per million (TPM) value was defined as ‘counts of reads mapped to miRNA \times 1,000,000 reads mapped to the reference genome’. The prediction of target genes of the mature miRNA from the known and novel precursor miRNAs (pre-miRNAs) was performed by TargetFinder with default parameters [36].

2.4. Bioinformatics Analysis

In order to integrate our datasets with the available Chinese fir miRNA information to profile the genetic architecture, we collected all Chinese fir miRNAs from miRbase and published small-RNA sequencing datasets for analysis. For the minimal folding free energies (MFEs) of secondary structures of miRNAs analysis, we derived all precursor sequences and their corresponding mature sequences from Chinese fir miRNA (previously reports included) and subjected them to Mfold web server (<http://www.unafold.org/> (accessed on 10 August 2022)) with default parameters [37].

Venn analysis was conducted to acquire unique miRNAs and overlapped miRNAs. For the phylogenetic analysis, the precursor miRNA sequences from Chinese fir and other plant species were subjected to MAFFT software (version 7) for multiple sequence alignment [38]. The phylogenetic tree was constructed on the PhyloSuite platform (version 1.2.2) using the maximum likelihood method with 1000 bootstrap replicates under the optimal model selected by IQ-TREE [39,40]. All phylogenetic tree modifications were completed using the ITOL online tool (<https://itol.embl.de/> (accessed on 15 August 2022)) [41].

3. Results

3.1. An Overview of High-Throughput Sequencing Datasets

A total of 290,752,309 raw reads were obtained from 17 small RNA libraries (Table S1). After we filtered out low-quality reads and reads <18 bp or >30 bp in length and removed adaptor sequences and reads with unknown base N content $\geq 10\%$, the clean reads accounted for 99.94% (290,566,075) of the total raw reads, and the Q30 values for all samples were greater than 95.31%, indicating the reliability of the data. In the seventeen clean data libraries, the length of the sRNAs (18–30 nt) presented a different distribution pattern, with a major peak at 21 nt, followed by 20 nt or 22 nt sRNAs (Table S2). This differs from earlier findings in Chinese fir and other conifer species, such as *Araucaria angustifolia*, in which the length of sRNAs had a highest abundance at 21 nt and a minor peak at 24 nt [13,20,21]. The result suggests that the sRNAs datasets obtained from the special materials used in this study are diverse compared to the datasets of sRNAs previously recorded in Chinese fir. Furthermore, the sRNA size distribution in most of conifer species differed from those of many angiosperms that are mainly clustered in 24 nt [8,42–44]. This implied that there was an unusual RNA silencing signature in gymnosperms. One reason for this may be the presence of a specific DCL3 enzyme, a member of the Dicer-like family (DCL) gene that cleaves double-strand RNAs, matures 24-nt RNAs in angiosperms [45]. The clean read sequences were then compared with the Silva, GtRNadb, Rfam, and Rепbase databases to further obtain rRNAs, tRNAs, snRNAs, snoRNAs, and repeat-associated sRNAs, and the unannotated reads containing small regulatory RNAs were used to identify miRNAs. More than 11 million sRNAs were obtained from each library (Figure 1; Table S3). Among them, unannotated reads were the most abundant RNAs with at least 7 million read counts. The large proportions of unannotated reads in this study agree with previous findings in Chinese fir and other gymnosperms such as *Cryptomeria fortunei* [21,46]. However, fewer unannotated sRNAs are recorded in many angiosperms [47]. Therefore, there may be abundant miRNA resources within coniferous species. The rRNAs were the second largest group with a range from 3,820,953 to 9,114,419 read counts. Other sRNAs possessed relatively low reads. In the next step, unannotated reads were subsequently mapped to the Chinese fir mRNA transcriptome database. Approximately 31.41–37.43% of the sRNAs were successfully matched to the reference transcriptome (Table S4). The mapped reads were further used to detect miRNAs.

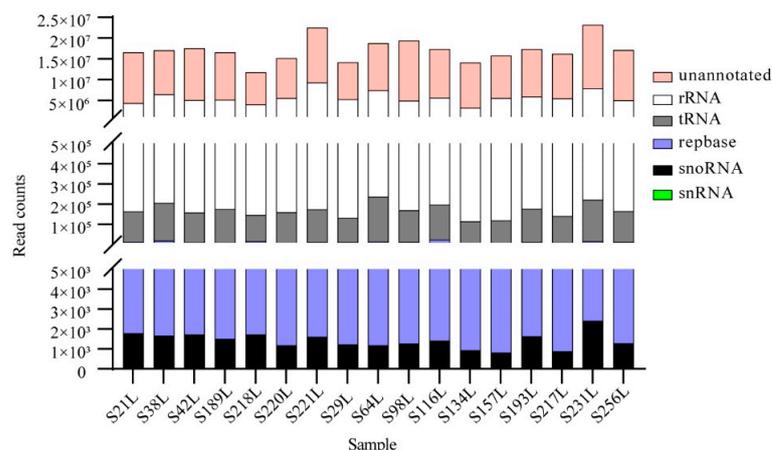


Figure 1. Sample read sequence types and attributes.

3.2. Classification of the miRNAs

It is well known that miRNAs can be divided into known and novel miRNAs. Known miRNAs are based on various databases matched to highly homologous sequences, while novel miRNAs are predicted by computerized algorithms under strict criteria. miRbase Database (Release 22.1) is the largest miRNA resource bank, which contains 38,589 miRNAs to date (1 August 2022). To identify known miRNA, mapped reads were aligned with

currently known mature miRNA sequences from all plant species in the miRbase Database using a BLASTn search with up to one mismatch. Sequences that were highly homologous to other species were identified as known miRNAs. As a result, 30 sequences successfully matched 16 conserved families (Figure 2). Within these miRNA families, the MIR159 and MIR396 families were the largest with 5 members, which comprised multiple species of different mature sequences. The second group were MIR319, MIR398, MIR399 and MIR529 family, consisting of 2–4 members. The remaining 10 families, including MIR156, MIR160, MIR162_1, MIR164, MIR166, MIR171_1, MIR395, MIR408, MIR536 and MIR6725, contained only one miRNA. For the length of miRNAs, the majority of the lengths of known mature miRNAs were 21 nt, and most of the lengths of their corresponding precursor sequences were 200–250 nt (Table S5). The prediction of target genes for miRNAs is a key step in understanding the regulation of miRNAs. In this study, we found that the identified known miRNAs harbored 1148 target genes with various functions by prediction and annotation (Table S6).

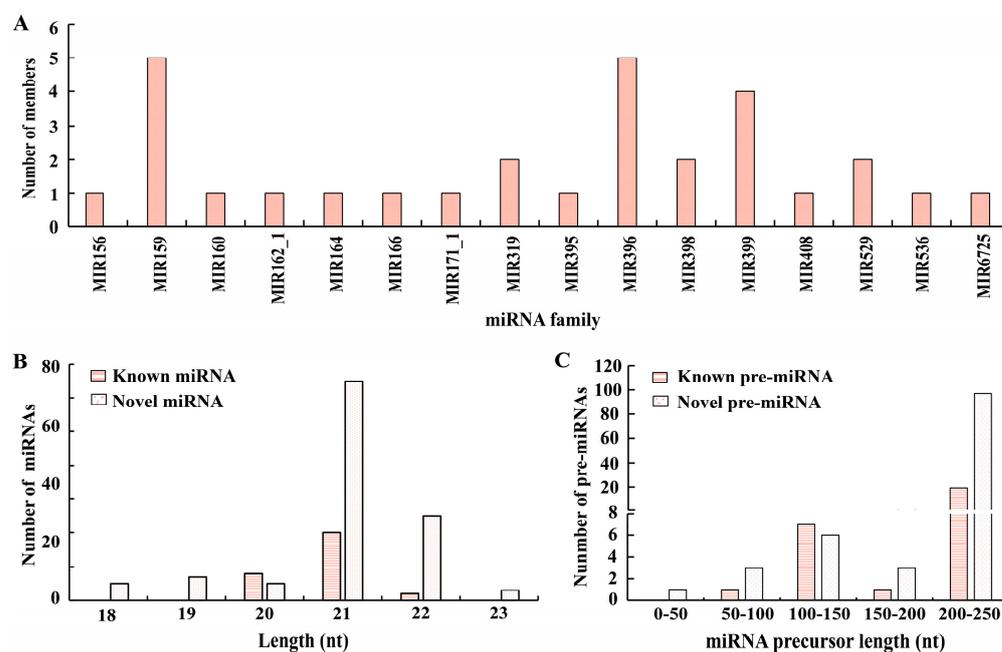


Figure 2. Characteristic of miRNA in Chinese fir. (A) Identification of miRNA family and its corresponding numbers from a selfed family in Chinese fir. (B) indicates the length distribution of known and novel miRNAs and its corresponding quantities. (C) indicates the length distribution of known and novel pre-miRNA and its corresponding quantities.

A typical stem-loop precursor is a prerequisite for the identification of newly developed miRNAs according to the criteria for the annotation of novel miRNAs [1]. For mapped reads that were not matched to known miRNAs, their sequences were processed for novel miRNA identification. In our study, 110 novel miRNAs and their precursor sequences were positively predicted with a computerized algorithm using miRDeep2 software with adjusted parameters [35]. These Chinese fir novel miRNAs were given names in the form of ‘cIn-miR plus number’, e.g., cIn-miR01 (Table S7). Similar to identified known miRNAs, 59% (n = 65) of the number of novel miRNA lengths primarily clustered at 21 nt, and the precursor length had a peak at 200–250 nt (Figure 2B,C). Additionally, the novel miRNAs targeted to 6927 hypothesis genes with extensive function indicating that these novel miRNAs may act as a pivotal genetic regulator during selfed seedling growth and development (Table S8). For example, some MYB transcription factors are involved in plant-form development of selfed progeny through associating with lignin biosynthesis-related genes [48]. Moreover, many novel miRNAs targeted to MYB transcription factor genes in this study, so there may be a similar mechanism here that regulates the plant-form development of selfed progeny.

To compare global expression profiles of detected miRNAs during seedling growth and normal development and depression, the expression analysis of miRNAs was performed with two different types of seedlings based on the normalized read count (transcripts per million, TPM) for known and novel miRNAs. In this study, we found that known and novel miRNAs did not show divergent expression patterns between normal and depressed seedlings (Figure S2 and Table S9), but presented different expression patterns between some individuals, indicating that these individuals may harbor differences in their regulatory processes in response to adaptive growth.

3.3. The Newly Developed miRNAs

Knowledge on available miRNA resources within a species is important for revealing the miRNA atlas in this species and will be very useful for further genetic and genomic use in this species. In light of this, we further integrated our datasets with the available Chinese fir miRNA information to profile the genetic architecture in this species, which mainly involves different biological processes, such as vascular cambium activity, vascular cambium development and seed dormancy, etc. [20–22]. A total of 899 known miRNAs were found in current reports based on statistical analysis (Figure 3A; Table S10). A larger proportion (72%, $n = 651$) of the known miRNAs were found by Liu et al. [20] compared to other Chinese fir studies. In contrast, Cao et al. [19] detected a smaller set of known miRNAs ($n = 19$). In this study, we also identified a relatively small set of known miRNAs ($n = 30$) in comparison with other records. Concerning the miRNA family, Venn analysis revealed four overlapped MIRNA families (MIR159, MIR164, MIR171_1 and MIR396) in all known Chinese fir miRNAs (Figure 3B). There were also a series of miRNA families in different overlapped combinations. Strikingly, one miRNA family belonging to MIR536 detected in this study had not been previously documented in Chinese fir. In addition, a high ratio (62%, $n = 110$) of novel miRNA numbers was found in our study when compared to all integrated novel miRNAs ($n = 178$) of Chinese fir (Figure 3C; Table S11). These results showed that the new set of miRNAs identified in this study was very different from other documents in Chinese fir. Similar to our findings, in *C. fortunei*, Hu et al. [49] revealed 80 conserved miRNAs belonging to 32 families and 145 unvalidated new miRNAs based on different developmental stages of the vascular cambium, whereas Zhang et al. [46] obtained more miRNAs, including 517 known and 212 novel miRNA mature/star sequences under the background of needle discoloration. Complicated biological processes of different developmental tissues within same species may contribute to the type and quantity of miRNAs. These results suggest that the further mining of miRNAs in this study is meaningful for clarifying the genetic architecture in this species.

For the sequence length of the known and novel mature miRNAs, the miRNA in every study also had a peak at 21 nt (Figure 3D,E). Meanwhile, 585 known pre-miRNAs with lengths ranging from 20 to 415 nt were detected in Liu et al. [20], followed by this study harboring 30 known pre-miRNAs (83–250 nt); other studies harbored fewer known pre-miRNAs, ranging from 4 to 19 known pre-miRNAs (Table S12). Interestingly, the largest number ($n = 110$) of novel pre-miRNAs was successfully predicted in this study in comparison with the intergraded data ($n = 178$). miRNAs regulate target gene expression post-transcriptionally through sequence complementarity, which is involved in many developmental processes in plants [50]. Although fewer known miRNAs ($n = 30$) were identified in this study, the number of target genes ($n = 1148$) was relatively high in our data when compared to that of target genes detected by Wan et al. [38] ($n = 209$), Qiu et al. [21] ($n = 240$), Cao et al. [22] and Liu et al. [20] ($n = 4077$) (Table S13). Additionally, we found a large number ($n = 6927$) of potential putative target genes in this study, while others recorded up to 660 target genes. These results indicated that the miRNAs detected in this study may play important roles in plant growth and development in the background of self-fertilization.

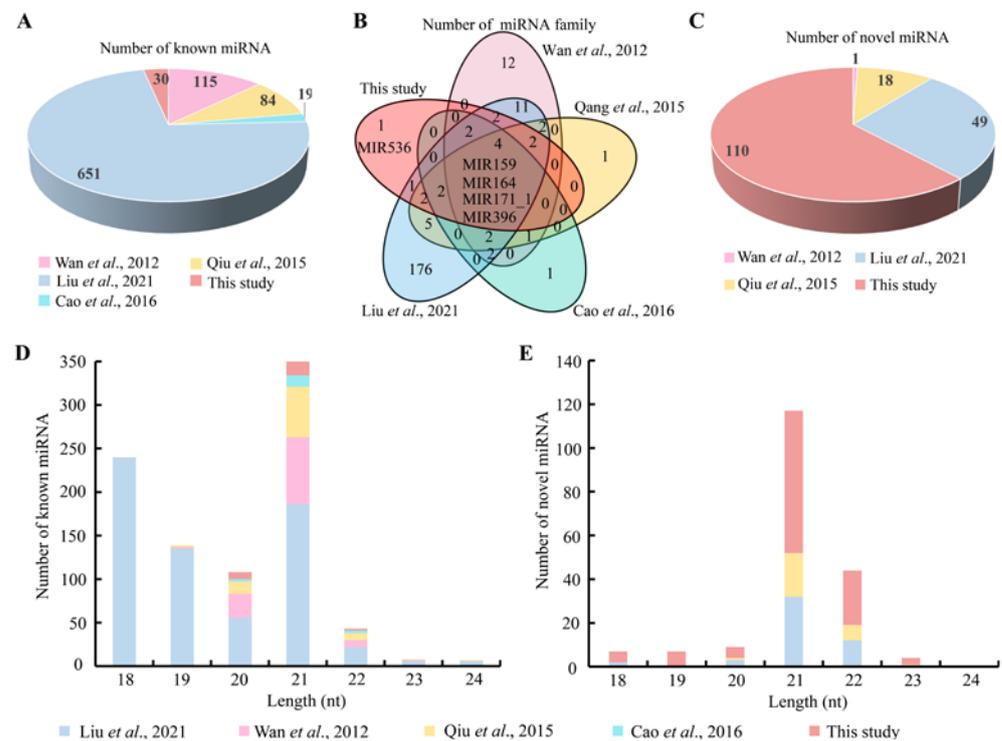


Figure 3. The number of miRNAs and the known miRNA families from published Chinese fir miRNA datasets. These Chinese fir miRNA datasets were collected from Wan et al., 2012 [16]; Qiu et al., 2015 [20]; Cao et al., 2016 [21]; Liu et al., 2021 [22]; and this study. (A) Pie chart showing the number of known miRNAs from different Chinese fir miRNA datasets. (B) Venn diagram presenting the number of miRNA families from different Chinese fir miRNA datasets and revealing four overlapped miRNA families (MIR 159, MIR164, MIR 171_1 and MIR 396) in all miRNA datasets and one unique miRNA family (MIR 536) in this study. (C) Pie chart showing the number of novel miRNAs from different Chinese fir miRNA datasets. (D) indicates the length distribution of known miRNAs and its corresponding quantities. (E) indicates the length distributions of novel miRNAs and the corresponding quantities.

In order to investigate the number of newly developed miRNAs in this study, all integrated known and novel mature miRNA sequences were blasted by ClustalX2 software, respectively. As a result, 19 known and 5 novel mature miRNA sequences in this study were found to be identical to those previously documented (Table S14). These results showed that 11 known and 105 novel miRNAs with 6858 target genes in this study were newly expanded miRNAs sets for Chinese fir (Table S15). The detected overlapped miRNAs and unique miRNAs in this study may help to understand the genetic variation in trees.

3.4. An Integrated Analysis of the Stability of Chinese Fir miRNAs in Terms of Energy

Precursor miRNA sequences do not have tight interactions with proteins before their cleavage by Dicer (or Dicer-like) enzymes [51]. Therefore, a stable secondary structure with low minimal folding free energy (MFE) might be necessary for miRNAs to avoid early degradation. Regarding the stability, we calculated the MFE for both the pre-miRNA and mature miRNA with Mfold web server under the default parameters [37]. Only the results of the optimal folding with minimum MFE value were used (Table S16). In this study, the MFE of known pre-miRNAs ranged from -39.75 to -97.4 kcal/mol with an average negative folding value of -72.4 kcal/mol (Table S5). In contrast, the absolute MFE value of their mature miRNAs was less than 5. For the MFE of novel miRNAs, their MFE of precursor sequence spanned from -11.1 to -115.50 kcal/mol with an average negative folding value of -67.6 kcal/mol. The absolute MFE value of the corresponding mature miRNAs was still less than 5. Regression analysis further revealed a relatively low R^2 value

($R^2 < 0.22$) between the MFE of pre-miRNAs and the MFE of the corresponding mature miRNAs in most studies (Figure 4). This result suggests that the thermodynamic stability of the secondary structure does not seem to be of major importance for mature miRNAs. One of the important reasons for this is that mature miRNAs typically combine with Argonaute proteins in order to form RNA-induced silencing complexes (RISCs) with stronger stable structure, which can silence the expression of complementary mRNAs [3].

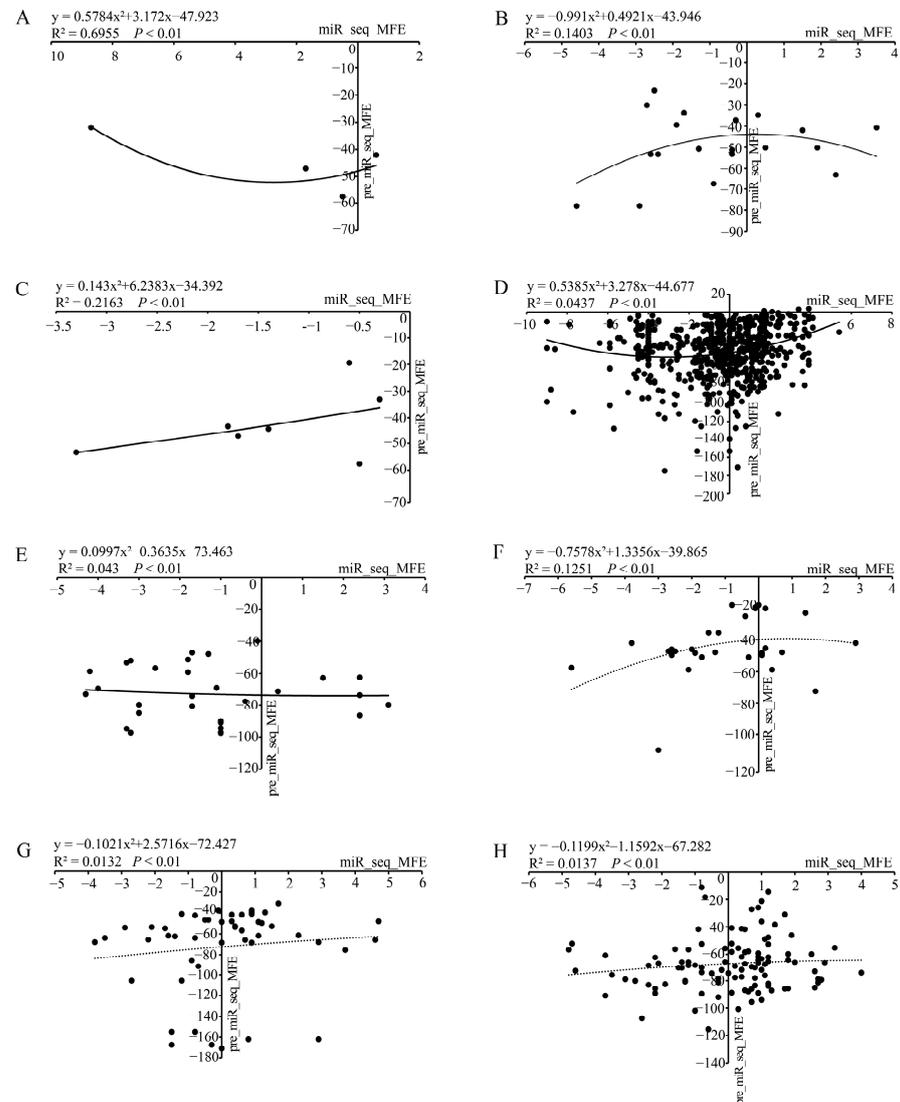


Figure 4. The relationship of minimal folding free energy (MFE) between pre-miRNA and the corresponding mature miRNA. Pre_miR_seq_MFE (precursor miRNA sequence MFE) and the corresponding miR_seq_MFE (mature miRNA sequence MFE) of known miRNAs in (A–E) from Wan et al., 2012 [16]; Cao et al., 2016 [20]; Qiu et al., 2015 [21]; Liu et al., 2021 [22]; and this study, respectively. Pre_miR_seq_MFE and the corresponding miR_seq_MFE of novel miRNAs in (F–H) from Qiu et al., 2015 [21]; Liu et al., 2021 [22]; and this study, respectively.

3.5. Phylogenetic Analysis of the Conserved miRNAs between Chinese Fir and Other Plant Species

Based on the one unique miRNA family (MIR536) and the four overlapped (conserved) miRNA families (MIR159, MIR164, MIR171_1, and MIR396) in Chinese fir miRNA datasets in this study, we attempted to elucidate the evolutionary history of these conserved family miRNAs among plant species. Phylogenetic analysis of precursor sequences contributes to the understanding of the evolutionary relationships of individual MIRNA genes in a family [52]. For the MIR536 family members in most plant species, we selected all sequences

belonging to MIR536 family member from the current miRbase database, which comprised one monocotyledon (*Asparagus officinalis* L.), two eudicotyledons (*Populus trichocarpa* Torr. & Gray and *Citrus sinensis* (L.) Osbeck), one gymnosperm (*Picea abies* (L.) Karst.), one moss (*Physcomitrella patens* Bruch & Schimp.) and one lycophyte (*Selaginella moellendorffii* Hieron.). Meanwhile, we searched for all pre-miRNA sequences of MIR159, MIR164, MIR171_1 and MIR396 families in 11 land plant species, including *Oryza sativa* L., *Zea mays* L., *Populus trichocarpa*, *Arabidopsis thaliana* (L.) Heynh., *Brassica napus* L., *Medicago truncatula* Gaertn., *Glycine max* (Linn.) Merr., *Pinus taeda* L., *Pinus densata* Mast. and *Picea abies* from the current miRbase database (Release 22.1) and also included the sequences of *Platykladus orientalis* (L.) Franco, a close relative of Chinese fir [14].

In our analysis of the phylogenetic relationships of the MIR536 family, it is clear that the Chinese fir MIR536 clustered with other gymnosperms, indicating its high evolutionary conservation (Figure 5). Despite a lack of empirical evidence regarding the presence of cln-MIR536, it is undeniable that the MIR536 families are predominant in gymnosperms [53]. Most of the MIR159 gathered with other gymnosperms, suggesting that Chinese fir MIR159 is also deeply conserved in gymnosperms during evolution (Figure S3A). This is consistent with early findings that miR159 families in gymnosperms formed a separate distinct clade, suggesting a highly conserved relationship of these families in plant species [54]. In contrast, some MIR164 and MIR396 precursor sequences clustered in monocotyledon and eudicotyledons species, suggesting their diversification of evolutionary divergence (Figures S3B and S4A). In the phylogenetic tree of the MIR171_1 family precursor sequence (Figure S4B), cln-MIR171 was distributed in all three pedigree lines including lycophytes, monocotyledon, and eudicotyledons, while the distribution of MIR171 in other gymnosperms was relatively scattered, implying the diversification of cln-miR171. However, Qiu et al. [55] revealed pde-MIR171 to be clustered with other gymnosperms, in which angiosperms form a separate cluster indicating a deep evolutionary conservation of the miR171 families. The reason for this may be that all miR171 family sequences of selected species were used for phylogenetic analysis in this study, while Qiu et al. [55] only used homologous sequences from different plant species. The result suggests that the different homologous sequences of MIR171_1 family have interspecific differences in evolutionary relationships.

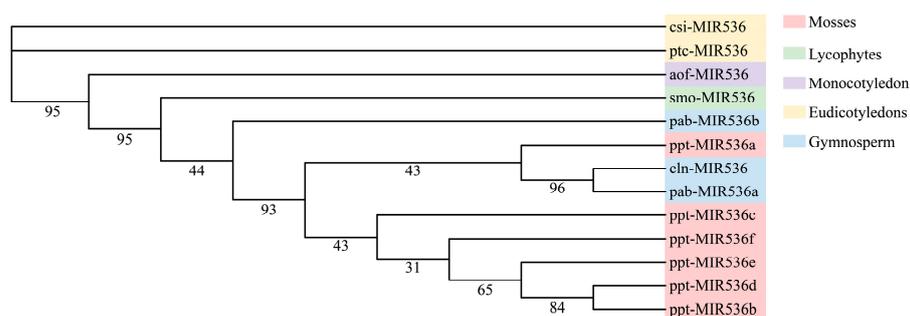


Figure 5. Phylogenetic tree of pre-miR536 in mosses, lycophytes, monocotyledons, eudicotyledons, and gymnosperms obtained from the IQ-TREE using the maximum likelihood method with 1000 bootstrap replicates. Labels: pab—*P. abies*; cln—*C. lanceolata*; smo—*Selaginella moellendorffii*; csi—*Citrus sinensis*; ppt—*Physcomitrella patens*; aof—*Asparagus officinalis*.

4. Conclusions

In the present study, 140 miRNAs were obtained from a Chinese fir selfing family harboring vigor and depressed progeny with rather different phenotypes. Remarkably, 75.7% (n = 106) of these miRNAs have not been recorded previously, and most (n = 105) of them belong to the novel set with 6858 putative target genes. Based on published small-RNA sequencing datasets from Chinese fir, we found a high ratio (62%, n = 110) of novel miRNA numbers in our study when compared to all integrated novel miRNAs (n = 178) of Chinese fir. Furthermore, we identify one MIR536 family that had not been

previously documented in this species and four overlapped miRNA families (MIR159, MIR164, MIR171_1 and MIR396) from Chinese fir miRNA datasets. Additionally, we also revealed a relatively low R^2 value ($R^2 < 0.22$) between the low minimal folding free energy (MFE) of pre-miRNAs and the MFE of its corresponding mature miRNAs in most datasets. From the conservation aspect, phylogenetic analysis shows that both MIR536 and MIR159 are conserved in gymnosperms, whereas the other miRNA families had evolutionary diversification. The present data provide potential value regarding Chinese fir miRNA biology and will be very useful for further genetic and genomic use in this species.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/genes13112137/s1>, Figure S1: The contrast photos of the morphological features in selfed progeny of Chinese fir clone cx569; Figure S2: Transcript profiles of known and novel miRNAs; Figure S3: Phylogenetic analysis of pre-miR159 and pre-miR164 between Chinese fir and other plant species; Figure S4: Phylogenetic analysis of pre-miR396 and pre-miR171_1 between Chinese fir and other plant species; Table S1: Statistics of clean reads; Table S2: Length distribution of small RNAs in *C. lanceolata*; Table S3: Distribution of *C. lanceolata* sRNAs among different categories; Table S4: Reference transcriptome comparison statistics; Table S5: Identification of known miRNAs in Chinese fir; Table S6: Known miRNA targets and their putative function; Table S7: Identification of novel miRNAs in Chinese fir; Table S8: Novel miRNA targets and their putative function; Table S9: The expression of novel and known miRNAs; Table S10: The number of miRNAs and the family in known MIRNA from Chinese fir; Table S11: The number of novel miRNAs in Chinese fir; Table S12: The length of miRNA precursor from Chinese fir; Table S13: The number of miRNAs target in Chinese fir; Table S14: The same mature miRNA sequence in all Chinese fir; Table S15: The novel miRNA (n = 105) targets and their putative function in this study; Table S16: The relationship of MFE between pre-miRNA and its corresponding mature miRNA.

Author Contributions: Conceptualization, Y.L. and H.Z.; methodology, H.Z. and H.D.; software, H.D.; for analysis H.D.; resources, H.Z.; materials, H.Z.; collection and/or processing, D.H., G.W. and H.D.; data curation, R.H.; writing—original draft preparation, H.D.; writing—review and editing, Y.L., Y.S., R.H., R.W. (Runhui Wang), R.W. (Ruping Wei), S.Y. and H.Z.; funding acquisition, H.Z. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported by the National Natural Science Foundation of China (no. 31972956) and the Key-Area Research and Development Program of Guangdong Province (no. 2020B020215001).

Data Availability Statement: The dataset(s) supporting the conclusions of this article are available in the NCBI's Short Read Archive (SRA) with the BioProject accession numbers PRJNA899478.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Meyers, B.C.; Axtell, M.J.; Bartel, B.; Bartel, D.P.; Baulcombe, D.; Bowman, J.L.; Cao, X.; Carrington, J.C.; Chen, X.; Green, P.J.; et al. Criteria for annotation of plant MicroRNAs. *Plant Cell* **2008**, *20*, 3186–3190. [[CrossRef](#)]
2. Bonnet, E.; Wuyts, J.; Rouzé, P.; Van de Peer, Y. Evidence that microRNA precursors, unlike other non-coding RNAs, have lower folding free energies than random sequences. *Bioinformatics* **2004**, *20*, 2911–2917. [[CrossRef](#)]
3. Iwakawa, H.O.; Tomari, Y. Molecular insights into microRNA-mediated translational repression in plants. *Mol. Cell* **2013**, *52*, 591–601. [[CrossRef](#)]
4. Samad, A.F.A.; Kamaroddin, M.F.; Sajad, M. Cross-kingdom regulation by plant micrnas provides novel insight into gene regulation. *Adv. Nutr.* **2021**, *12*, 197–211. [[CrossRef](#)]
5. Yates, L.A.; Norbury, C.J.; Gilbert, R.J.C. The long and short of microRNA. *Cell* **2013**, *153*, 516–519. [[CrossRef](#)]
6. Shi, T.; Wang, K.; Yang, P. The evolution of plant microRNAs: Insights from a basal eudicot sacred lotus. *Plant J.* **2017**, *89*, 442–457. [[CrossRef](#)]
7. Niu, S.; Li, J.; Bo, W.; Yang, W.; Zuccolo, A.; Giacomello, S.; Chen, X.; Han, F.; Yang, J.; Song, Y.; et al. The Chinese pine genome and methylome unveil key features of conifer evolution. *Cell* **2022**, *185*, 204–217. [[CrossRef](#)]
8. Morin, R.D.; Aksay, G.; Dolgosheina, E.; Ehardt, H.A.; Magrini, V.; Mardis, E.R.; Sahinalp, S.C.; Unrau, P.J. Comparative analysis of the small RNA transcriptomes of *Pinus contorta* and *Oryza sativa*. *Genome Res.* **2008**, *18*, 571–584. [[CrossRef](#)] [[PubMed](#)]
9. Lu, S.; Sun, Y.H.; Amerson, H.; Chiang, V.L. MicroRNAs in loblolly pine (*Pinus taeda* L.) and their association with fusiform rust gall development. *Plant J.* **2007**, *51*, 1077–1098. [[CrossRef](#)] [[PubMed](#)]

10. Niu, S.; Liu, C.; Yuan, H.W.; Li, P.; Li, Y.; Li, W. Identification and expression profiles of sRNAs and their biogenesis and action-related genes in male and female cones of *Pinus tabuliformis*. *BMC Genom.* **2015**, *16*, 693. [[CrossRef](#)]
11. Shen, T.; Xu, M.; Qi, H.; Feng, Y.; Yang, Z.; Xu, M. Uncovering miRNA-mRNA regulatory modules in developing xylem of *Pinus massoniana* via small RNA and degradome sequencing. *Int. J. Mol. Sci.* **2021**, *22*, 10154. [[CrossRef](#)] [[PubMed](#)]
12. Xia, R.; Xu, J.; Arikiti, S.; Meyers, B.C. Extensive families of miRNAs and PHAS loci in Norway Spruce demonstrate the origins of complex phasiRNA networks in seed plants. *Mol. Biol. Evol.* **2015**, *32*, 2905–2918. [[CrossRef](#)] [[PubMed](#)]
13. Galdino, J.H.; Eguiluz, M.; Guzman, F.; Margis, R. Novel and conserved miRNAs among brazilian pine and other gymnosperms. *Front. Genet.* **2019**, *10*, 222. [[CrossRef](#)]
14. Jia, Z.; Zhao, B.; Liu, S.; Lu, Z.; Chang, B.; Jiang, H.; Cui, H.; He, Q.; Li, W.; Jin, B.; et al. Embryo transcriptome and miRNA analyses reveal the regulatory network of seed dormancy in *Ginkgo biloba*. *Tree Physiol.* **2021**, *41*, 571–588. [[CrossRef](#)]
15. Yang, Y.; Hu, X.G.; Zheng, B.; Li, Y.; Wang, T.; Sharma, A.; Yuan, H.; Mao, J.F. Transcriptome-wide identification and characterization of microRNAs and their Targets in a highly adaptable conifer *platycladus orientalis*. *J. Am. Soc. Hortic. Sci.* **2022**, *147*, 7–17. [[CrossRef](#)]
16. Wan, L.C.; Wang, F.; Guo, X.; Lu, S.; Qiu, Z.; Zhao, Y.; Zhang, H.; Lin, J. Identification and characterization of small non-coding RNAs from Chinese fir by high throughput sequencing. *BMC Plant Biol.* **2012**, *12*, 146. [[CrossRef](#)]
17. Yang, Y.; Ferguson, D.K.; Liu, B.; Mao, K.; Gao, L.; Zhang, S.; Wan, T.; Rushforth, K.; Zhang, Z. Recent advances on phylogenomics of gymnosperms and a new classification. *Plant Divers.* **2022**, *44*, 340–350. [[CrossRef](#)]
18. Shi, J.; Zhen, Y.; Zheng, R.H. Proteome profiling of early seed development in *Cunninghamia lanceolata* (Lamb.) Hook. *J. Exp. Bot.* **2010**, *61*, 2367–2381. [[CrossRef](#)]
19. Zheng, H.; Hu, D.; Wang, R.; Wei, R.; Yan, S. Assessing 62 Chinese Fir (*Cunninghamia lanceolata*) breeding parents in a 12-year grafted clone test. *Forests* **2015**, *6*, 3799–3808. [[CrossRef](#)]
20. Qiu, Z.; Li, X.; Zhao, Y.; Zhang, M.; Wan, Y.; Cao, D.; Lu, S.; Lin, J. Genome-wide analysis reveals dynamic changes in expression of microRNAs during vascular cambium development in Chinese fir, *Cunninghamia lanceolata*. *J. Exp. Bot.* **2015**, *66*, 3041–3054. [[CrossRef](#)] [[PubMed](#)]
21. Cao, D.; Xu, H.; Zhao, Y.; Deng, X.; Liu, Y.; Soppe, W.J.; Lin, J. Transcriptome and degradome sequencing reveals dormancy mechanisms of *Cunninghamia lanceolata* Seeds. *Plant Physiol.* **2016**, *172*, 2347–2362. [[CrossRef](#)] [[PubMed](#)]
22. Liu, G.; Xue, X.; Feng, J.; Cao, D.; Lin, J.; Xu, H. Age-dependent microRNAs in regulation of vascular cambium activity in Chinese fir (*Cunninghamia lanceolata*). *Trees* **2021**, *35*, 1451–1466. [[CrossRef](#)]
23. Charlesworth, D.; Willis, J.H. The genetics of inbreeding depression. *Nat. Rev. Genet.* **2019**, *10*, 783–796. [[CrossRef](#)] [[PubMed](#)]
24. Rymer, P.D.; Sandiford, M.; Harris, S.A.; Billingham, M.R.; Boshier, D.H. Remnant Pachi-ra quinate pasture trees have greater opportunities to self and suffer reduced reproductive success due to inbreeding depression. *Heredity* **2015**, *115*, 115–124. [[CrossRef](#)] [[PubMed](#)]
25. Roessler, K.; Muyle, A.; Diez, C.M.; Gaut, G.R.J.; Bousios, A.; Stitzer, M.C.; Seymour, D.K.; Doebley, J.F.; Liu, Q.; Gaut, B.S. The genome-wide dynamics of purging during selfing in *maize*. *Nat. Plants* **2019**, *5*, 980–990. [[CrossRef](#)]
26. Tsuchimatsu, T.; Kakui, H.; Yamazaki, M.; Marona, C.; Tsutsui, H.; Hedhly, A.; Meng, D.; Sato, Y.; Städler, T.; Grossniklaus, U.; et al. Adaptive reduction of male gamete number in the selfing plant *Arabidopsis thaliana*. *Nat. Commun.* **2020**, *11*, 2885. [[CrossRef](#)] [[PubMed](#)]
27. Yi, H.; Wang, J.; Wang, J.; Rausher, M.; Kang, M. Genomic insights into inter- and intra-specific mating system shifts in *Primulina*. *Mol. Ecol.* **2022**, *31*, 5699–5713. [[CrossRef](#)]
28. Wang, R.H.; Hu, D.H.; Zheng, H.Q.; Liu, W.X.; Liang, R.Y.; Yang, B.Q. Genetic Variation Analysis of Flowering Phenology in the 2.5 Generation Seed Orchard of *Cunninghamia lanceolata*. *J. Southwest For. Univ. (Nat. Sci.)* **2013**, *33*, 25–29. [[CrossRef](#)]
29. Wang, Z.M.; Chen, Y.T. An analysis on the combining ability of main growth character in Chinese fir and the application of its heterosis. *For. Res.* **1988**, *1*, 614–624.
30. Chen, Y.T.; He, G.P.; Li, G.X. The effect of inbreeding on seed germination and seedling height growth of Chinese fir. *For. Res.* **1989**, *2*, 420–426.
31. Deng, H.; Hu, D.; Wei, R.; Yan, S.; Wang, R.; Zheng, H. Global transcriptome analysis reveals genes associated with seedling advance growth traits in a selfed family of Chinese fir (*Cunninghamia lanceolata*). *Dendrobiology* **2022**, *87*, 27–46. [[CrossRef](#)]
32. Haas, B.J.; Papanicolaou, A.; Yassour, M.; Grabherr, M.; Blood, P.D.; Bowden, J.; Couger, M.B.; Eccles, D.; Li, B.; Lieber, M.; et al. De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. *Nat. Protoc.* **2013**, *8*, 1494–1512. [[CrossRef](#)]
33. Langmead, B.; Salzberg, S.L. Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **2012**, *9*, 357–359. [[CrossRef](#)]
34. Griffiths-Jones, S.; Saini, H.K.; van Dongen, S.; Enright, A.J. miRBase: Tools for microRNA genomics. *Nucleic Acids Res.* **2008**, *36*, D154–D158. [[CrossRef](#)]
35. Friedlander, M.R.; Mackowiak, S.D.; Li, N.; Chen, W.; Rajewsky, N. miRDeep2 accurately identifies known and hundreds of novel microRNA genes in seven animal clades. *Nucleic Acids Res.* **2012**, *40*, 37–52. [[CrossRef](#)]
36. Fahlgren, N.; Carrington, J.C. miRNA target prediction in plants. *Methods Mol. Biol.* **2010**, *592*, 51–57. [[CrossRef](#)]
37. Zuker, M. Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res.* **2003**, *31*, 3406–3415. [[CrossRef](#)]

38. Katoh, K.; Standley, D.M. MAFFT multiple sequence alignment software version 7: Improvements in performance and usability. *Mol. Biol. Evol.* **2013**, *30*, 772–780. [[CrossRef](#)]
39. Nguyen, L.T.; Schmidt, H.A.; Von Haeseler, A.; Minh, B.Q. IQ-TREE: A fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol. Biol. Evol.* **2015**, *32*, 268–274. [[CrossRef](#)]
40. Zhang, D.; Gao, F.; Jakovlić, I.; Zou, H.; Zhang, J.; Li, W.X.; Wang, G.T. PhyloSuite: An integrated and scalable desktop platform for streamlined molecular sequence data management and evolutionary phylogenetics studies. *Mol. Ecol. Resour.* **2020**, *20*, 348–355. [[CrossRef](#)]
41. Letunic, I.; Bork, P. Interactive Tree of Life (iTOL) v5: An online tool for phylogenetic tree display and annotation. *Nucleic Acids Res.* **2021**, *49*, 293–296. [[CrossRef](#)] [[PubMed](#)]
42. Zhou, Z.S.; Zeng, H.Q.; Liu, Z.P.; Yang, Z.M. Genome-wide identification of *Medicago truncatula* microRNAs and their targets reveals their differential regulation by heavy metal. *Plant Cell Environ.* **2012**, *35*, 86–99. [[CrossRef](#)] [[PubMed](#)]
43. Wu, Y.; Yang, L.; Yu, M.; Wang, J. Identification and expression analysis of microRNAs during ovule development in rice (*Oryza sativa*) by deep sequencing. *Plant Cell Rep.* **2017**, *36*, 1815–1827. [[CrossRef](#)]
44. Chen, S.Y.; Su, M.H.; Kremling, K.A.; Lepak, N.K.; Romay, M.C.; Sun, Q.; Bradbury, P.J.; Buckler, E.S.; Ku, H.M. Identification of miRNA-eQTLs in maize mature leaf by GWAS. *BMC Genom.* **2020**, *21*, 689. [[CrossRef](#)]
45. Dolgosheina, E.V.; Morin, R.D.; Aksay, G.; Sahinalp, S.C.; Magrini, V.; Mardis, E.R.; Mattsson, J.; Unrau, P.J. Conifers have a unique small RNA silencing signature. *RNA* **2008**, *14*, 1508–1515. [[CrossRef](#)]
46. Zhang, Y.; Yang, J.; Zhu, L.; Xue, J.; Hu, H.; Cui, J.; Xu, J. Identification of microRNAs and their target genes related to needle discoloration of evergreen tree Chinese cedar (*Cryptomeria fortunei*) in cold winters. *Planta* **2021**, *254*, 31. [[CrossRef](#)] [[PubMed](#)]
47. Dou, X.; Zhou, Z.; Zhao, L. Identification and expression analysis of miRNAs in germination and seedling growth of Tibetan hulless barley. *Genomics* **2021**, *113*, 3735–3749. [[CrossRef](#)]
48. Arro, J.; Cuenca, J.; Yang, Y.; Liang, Z.; Cousins, P.; Zhong, G.Y. A transcriptome analysis of two grapevine populations segregating for tendril phyllotaxy. *Hortic. Res.* **2017**, *4*, 17032. [[CrossRef](#)]
49. Hu, H.; Guo, Z.; Yang, J.; Cui, J.; Zhang, Y.; Xu, J. Transcriptome and microRNA sequencing identified miRNAs and target genes in different developmental stages of the vascular cambium in *Cryptomeria fortunei* Hooibrenk. *Front. Plant Sci.* **2021**, *12*, 751–771. [[CrossRef](#)]
50. Yu, Y.; Jia, T.; Chen, X. The ‘how’ and ‘where’ of plant microRNAs. *New Phytol.* **2017**, *216*, 1002–1017. [[CrossRef](#)]
51. Lee, Y.; Jeon, K.; Lee, J.T.; Kim, S.; Kim, V.N. MicroRNA maturation: Stepwise processing and subcellular localization. *EMBO J.* **2002**, *21*, 4663–4670. [[CrossRef](#)] [[PubMed](#)]
52. Barik, S.; SarkarDas, S.; Singh, A.; Gautam, V.; Kumar, P.; Majee, M.; Sarkar, A.K. Phylogenetic analysis reveals conservation and diversification of micro RNA166 genes among diverse plant species. *Genomics* **2014**, *103*, 114–121. [[CrossRef](#)] [[PubMed](#)]
53. Cha’vez Montes, R.A.; Rosas-Cárdenas, D.F.F.; De Paoli, E.; Accerbi, M.; Rymarquis, L.A.; Mahalingam, G.; Marsch-Martínez, M.; Meyers, B.C.; Green, P.J.; De Folter, S. Sample sequencing of vascular plants demonstrates widespread conservation and divergence of microRNAs. *Nat. Commun.* **2014**, *5*, 3722. [[CrossRef](#)] [[PubMed](#)]
54. Anand, S.; Lal, M.; Das, S. Comparative genomics reveals origin of MIR159A-MIR159B paralogy, and complexities of PTGS interaction between miR159 and target GA-MYBs in *Brassicaceae*. *Mol. Genet. Genom.* **2019**, *294*, 693–714. [[CrossRef](#)]
55. Qiu, Z.B.; Yuan, M.M.; Hai, B.Z.; Wang, L.; Zhang, L. Characterization and expression analysis of conserved miRNAs and their targets in *Pinus densata*. *Biol. Plant.* **2016**, *60*, 427–434. [[CrossRef](#)]