

Transposon-associated small RNAs involved in plant defense in poplar

Sijia Liu

Beijing Advanced Innovation Center for Tree Breeding by Molecular Design, Beijing Forestry University,
No. 35, Qinghua East Road, Beijing 100083, P. R. China;

xian li

Beijing Advanced Innovation Center for Tree Breeding by Molecular Design, Beijing Forestry University,
No. 35, Qinghua East Road, Beijing 100083, P. R. China;

iyang zhao

Beijing Advanced Innovation Center for Tree Breeding by Molecular Design, Beijing Forestry University,
No. 35, Qinghua East Road, Beijing 100083, P. R. China;

jianbo xie (✉ jbxie@bjfu.edu.cn)

Beijing Advanced Innovation Center for Tree Breeding by Molecular Design, Beijing Forestry University,
No. 35, Qinghua East Road, Beijing 100083, P. R. China;

Research Article

Keywords: small RNAs, defense response, infection, poplar, *Melampsora larici-populina*

Posted Date: February 11th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-154563/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Transposon-associated small RNAs involved in plant defense in poplar

Running title: Possible large-scale regulation between plant and pathogens mediated by cross-kingdom siRNAs

Sijia Liu^{1,2,3}, Xian Li^{1,2,3}, Yiyang Zhao^{1,2,3}, and Jianbo Xie^{1,2,3,*}

¹Beijing Advanced Innovation Center for Tree Breeding by Molecular Design, Beijing Forestry University, No. 35, Qinghua East Road, Beijing 100083, P. R. China;

²National Engineering Laboratory for Tree Breeding, College of Biological Sciences and Technology, Beijing Forestry University, No. 35, Qinghua East Road, Beijing 100083, P. R. China;

³Key Laboratory of Genetics and Breeding in Forest Trees and Ornamental Plants, Ministry of Education, College of Biological Sciences and Technology, Beijing Forestry University, No. 35, Qinghua East Road, Beijing 100083, P. R. China;

*To whom correspondence should be addressed. Tel: +86-10-62336007; Fax: +86-10-62336164.

E-mail: jbxie@bjfu.edu.cn.

Abstract

Background: Small RNAs (sRNAs) are hypothesized to contribute to plant defense responses by increasing the overall genetic diversity and regulating gene expression; however, their origins and functional importance in plant defense remain unclear. Here, we use Illumina sequencing to assess how sRNA populations vary in the Chinese white poplar (*Populus tomentosa*) during a rust fungus (*Melampsora larici-populina*) infection. We sampled sRNAs from the biotrophic growth phase (T02; 48 h post infection) and the urediniospore formation and release phase (T03; 168 h), two essential stages associated with plant colonization and biotrophic growth in rust fungi.

Results: The proportion of siRNA clusters located in pseudogenes and transposons was significantly larger than would be expected by chance and infection-stage-specific differences in siRNAs primarily originated from those in the transposon regions. We also found that the abundance of clusters comprising 24-nt siRNAs located in the transposon and intergenic regions underwent more substantial changes as the infection progressed. A target analysis revealed that 95% of fungal genes were predicted to be targets of *Populus* sRNAs. Pathogen effector genes were targeted by more sRNAs identified during the biotrophic growth and urediniospore formation and release phases than in the control plants, suggesting a clear selection for sRNA-target interactions. Compared with the miRNAs conserved between different plant species, a significantly higher proportion of *Populus*-specific miRNAs appeared to target *NB-LRR* genes.

Conclusions: This integrated study on the plant colonization and biotrophic growth in rust fungi profiles could provide evolutionary insights into the origin and potential roles of the sRNAs in plant defense, coevolution with pathogens, and functional innovation.

Key words: small RNAs, defense response, infection, poplar, *Melampsora larici-populina*

Background

Small RNAs (sRNAs) are ~21-24-nucleotide (nt) RNAs that can regulate gene expression, maintain genome integrity and chromatin structure, and influence plant development, such as seed germination [1-4]. sRNAs typically have an inhibitory effect on gene expression, and the corresponding regulatory mechanisms are therefore subsumed under the heading of RNA silencing [5].

Two major classes of endogenous sRNAs have been identified in plants, namely microRNAs (miRNAs) and small interfering RNAs (siRNAs) [5, 6]. Plant miRNAs are transcribed from imperfect hairpin-shaped precursors and are typically 21 nt long, while siRNAs are 21-24 nt in length and are generated from double-stranded RNA duplexes [7]. miRNAs reduce the protein abundances of their targets through the direct cleavage or posttranscriptional regulation of target mRNAs to reduce their translation [8], which has important functions in a wide range of biological processes [5]. In *Arabidopsis*, over 100 miRNAs have been shown to be important for plant development [9] and abiotic stress tolerance [10].

siRNAs have been sequenced in many plant species, such as *Arabidopsis thaliana*, rice (*Oryza sativa*), and wheat (*Triticum aestivum*) [11-14], and have the ability to silence their targets by directly complementing their transcripts to prevent their translation. Studies in maize (*Zea mays*) have revealed that sRNA populations vary between parents and their progeny, contributing to the dramatic vigor of maize hybrids [15].

Innate immunity is an evolutionarily ancient mechanism that protects plants from

a wide range of pathogens [16, 17]. Many lines of evidence have confirmed that sRNAs contribute to plant defenses against pathogens [17-19]; for example, changes in the accumulation of specific miRNAs, such as miR393, miR160, and miR773, have been reported during the infection of *Arabidopsis* by the bacterial pathogen *Pseudomonas syringae* [20, 21] or after treatments with bacterial flagellin [22].

Evolutionary analyses revealed that a miRNA superfamily composed of the miR482 and miR2118 families targets the plant *NB-LRR* (nucleotide-binding leucine-rich-repeat) defense genes [23]. The best-known examples of plant defense sRNAs are the siRNAs, which target viral RNAs for degradation and thereby restrict viral infection [24]. The sRNAs also form part of the antibacterial defense mechanism in the eukaryote host cells [25]. Although thousands of endogenous siRNAs have been sequenced [26], relatively little is known about their biological roles in defense, and further examination is required. Association genetics involving the characterization of polymorphisms in genomic elements and the detection of associations with phenotypes could provide clues about the functions of these essential genomic elements.

Woody plants have distinct features that distinguish them from most herbaceous plants, including their large size, long lifespans, and perennial growth [27]. Plant growth and wood formation are complex and dynamic processes requiring the coordinated regulation of diverse metabolic pathways and trade-offs with immune responses [28, 29]. It is almost certain that a perennial plant will encounter a pathogen before it can reproduce, and the long generation intervals of trees makes it difficult for

them to match the evolutionary rates of microbes [30].

The *Populus* genus of poplars includes many important woody species, such as the western balsam poplar (*Populus trichocarpa*) [30], the desert poplar (*Populus euphratica*) [31], and the Chinese white poplar (*Populus tomentosa*) [32]. Several poplars have been selected as model tree species for their small genome size and rapid growth. The availability of reference genome sequences for *Populus* species thus make them important model systems for the investigation of sRNA functions during pathogen infections. Rust fungi (*Pucciniales*, *Basidiomycota*) are a diverse group of plant pathogens comprising more than 6,000 species. An economically important group of poplar pathogens, *Melampsora* spp., are the rust fungi, which represent a major constraint on the development of bioenergy programs using domesticated poplars because of the lack of durable host resistance [33].

Plants have the capacity to recognize pathogens through strategies involving both conserved and variable defense regulators, while pathogens suppress plant defense responses by delivering virulence effectors [34]. To learn more about the origin, evolution, and plant defense effects of the poplar sRNAs, we sequenced the sRNAome of Chinese white poplar during the two essential stages of rust fungus (*Melampsora larici-populina*) infection: plant colonization and biotrophic growth. These analyses reveal the evolutionary importance and functional novelty of the sRNAs.

Methods

Plant material and fungal treatments

The Chinese white poplar (*Populus tomentosa*) was planted in pots under natural light conditions (12 h of 1,250 $\mu\text{mol m}^{-2} \text{s}^{-1}$ photosynthetically active radiation) at $25\pm 1^\circ\text{C}$ (day and night) and $50\pm 1\%$ relative humidity (day and night) in an air-conditioned glasshouse using soilless culture technology. The infection experiments were performed using *M. larici-populina* strain G5. Six inoculation spots per poplar leaf were each inoculated with 5 μl of 10^6 conidia ml^{-1} , with three biological replicates per treatment. The inoculated leaves were incubated in an artificial climate incubator (LT36VL; Percival Scientific, Inc; Perry, U.S.A.) under 95% relative humidity and $25\pm 1^\circ\text{C}$, and were harvested at 48 (T02) and 168 (T03) hours post inoculation (hpi). In the control group (T01), the leaves were sprayed with sterile tap water and harvested at 48 hpi. The samples from all leaves exposed to the same treatment were pooled and treated as one biological repeat, and three independent experimental repeats were performed for each treatment (control, 48 hpi, 168 hpi). All samples were immediately frozen in liquid nitrogen and stored at -80°C for RNA extraction.

miRNA library construction

Total RNA was extracted with Trizol (Thermo Fisher Scientific) according to the manufacturer's protocol [35]. The total RNAs generated by Dicer processing are efficiently targeted by the included modified adapters. RNA-seq libraries and sRNA libraries were prepared using the TruSeq RNA Library Preparation Kit and TruSeq Small RNA Library Preparation Kit (Illumina), respectively [36]. sRNAs with adapters on both ends were used as templates to create cDNA constructs using reverse transcription PCR. After being purified and quantified using a Qubit dsDNA HS

(Qubit 2.0 Fluorometer) and Agilent 2100, the PCR products were used for cluster generation and sequencing on an Illumina HiSeq 2500, according to the cBot and HiSeq 2500 user guides, respectively. sRNA sequencing was performed with two biological replicates for each treatment and transcriptome sequencing was performed with three biological replicates for each treatment.

mRNA sequencing, alignment, and normalization

Quality control of raw reads was carried out using FastQC to remove reads containing >50% low-quality nucleotides (sQ \leq 5) and >10% ambiguous nucleotides. The qualified reads that passed the filters above were then aligned using TopHat2 [37] with the parameters --read-mismatches 2 -p -G to generate read alignments for each sample. Up to two mismatches were permitted in each read alignment. The transcript abundance was calculated, and differential transcript expression was computed using CuffDiff with the parameters -p -b [38]. Run bias detection and correction algorithms to improve the accuracy of transcript abundance calculations.

Effector prediction

Effector prediction was performed using a machine learning program, EffectorP, with default parameters [39]. In this analysis, all of the protein sequences in the *M. larici-populina* genome were used.

sRNAome analysis

The sequences generated from the leaves exposed to the three infection treatments were used to detect the transcript abundance of mature sRNAs. [Figure S1](#) shows the

procedures used for sRNA library construction. Over 16 million raw reads were produced per library, of which 33.5–44.0% aligned perfectly (no more than one mismatch) to version 3.0 of the *P. trichocarpa* genome. These perfectly aligned sequences were annotated by BLAST-searching them against the GenBank and Rfam databases (version 13; <http://rfam.xfam.org/>), allowing one mismatch. The tRNAs, rRNAs, snRNAs, snoRNAs, and scRNAs were removed from the sequencing reads. The remaining unannotated sRNAs were searched against the known miRNAs from miRBase version 22.0 (<http://www.mirbase.org/>), allowing a maximum of two mismatches. To predict the novel miRNAs, miRDeep2 [40] was used.

Detection of sRNA clusters

To identify the siRNA clusters, the strategy used by Johnson et al. (2009) [41] was applied, with some modifications. The 21-24-nt sRNAs from all the samples that did not match the miRNAs, ribosomal rRNA, or tRNAs and that could map to the *P. trichocarpa* genome were processed together. siRNAs within 100 bp of each other were merged into blocks referred to as siRNA clusters. The coordinates of the clusters were defined by the first and last siRNAs of the overlapping sequences. siRNA clusters with a transcription abundance of at least 5 tpm were included in the downstream analysis. The clusters were labeled according to the most abundant length of siRNA they contained (21, 22, or 24 nt), and the type of genomic region in which they were located.

Target prediction

The miRNA targets were annotated using the standard settings of psRNATarget [42]

with an expectation value of 2.0 and top targets of 200. For targets in poplar, the transcripts of *P. trichocarpa* were downloaded from Phytozome 12 (<https://phytozome.jgi.doe.gov/pz/portal.html>). For targets in *M. larici-populina*, transcripts from this species were downloaded from the Ensemble database (ftp://ftp.ensemblgenomes.org/pub/release-39/fungi/fasta/melampsora_laricipopulina/cdna/).

Genomic locations of miRNAs in the *P. trichocarpa* genome

The siRNA clusters were screened for their localization within the transposons, introns, pseudogenes, intergenic regions, 5' UTR, 3' UTR, and upstream and downstream (2 kb) of coding genes of the *P. trichocarpa* genome (Phytozome version 12), with an overlapping rate of above 80%. To assess the overrepresentation of siRNA clusters in each genomic element, deviations from expected proportions were assessed using randomization tests. The genomic locations containing the same length of siRNA clusters were randomly sampled 1,000 times, after which the overlap of the siRNA clusters with these known genomic elements was assessed. A one-tailed *P*-value was estimated based on the *Z* score of the difference between the actual count minus the mean count of 1,000 surrogates relative to the standard deviation from the 1,000 surrogates. The *P*-value was the probability to the right of the observed count, calculated using a one-tailed *z* test.

TE annotation

The TE annotations used in this study were obtained from the outputs of the RepeatMasker (RM) software version 4.0.6 (1) combined with the database

(Dfam_Consensus-20170127, RepBase-20170127; -species parameter: *Populus*).

These RM outputs were filtered to remove non-TE elements, such as satellites, simple repeats, low complexity sequences, and rRNA.

Pseudogene annotation

The intergenic sequences of the *P. trichocarpa* genome were used to identify the putative Pseudogenes. The overall pipeline used for this identification was generally based on the PlantPseudo workflow [43], and consisted of four major steps: (1) identify the masked intergenic regions with sequence similarity to known proteins using BLAST; (2) eliminate redundant and overlapping BLAST hits in places where a given chromosomal segment has multiple hits; (3) link homologous segments into contigs; and (4) realign sequences using tfasty to identify features that disrupt contiguous protein sequences.

Results

Poplar sRNAs expressed during a rust fungus infection

To study the expression profile of the Chinese white poplar sRNAs during the biotrophic growth (T02) and urediniospore formation and release (T03) stages of infection by *Melampsora larici-populina* [44], sRNA libraries were prepared and subjected to high-throughput sequencing. Specifically, we prepared a library of RNAs approximately 18-30 nt in length from the three samples and sequenced them using an Illumina HiSeq 2500, generating 81.3 million clean reads in total ([Table S1](#)).

After processing the sRNA data and combining the sequences across the libraries,

we identified a set of distinct sRNAs for the two infection stages and the mock-inoculated control plants. The resulting data were evaluated to ensure they were consistent with the expected size range of functional siRNAs and miRNAs, with 56.6–65.4% of the generated sRNAs having an expected sequence length of 21-24 nt (Figure S2). In total, 33.4–44.0% of the total reads could be aligned perfectly (with no more than one mismatch) to the *Populus trichocarpa* genome (version 3.0)[30].

Pseudogenes and transposons act as catalysts for the formation of siRNA clusters

We next studied the locations of the siRNAs to elucidate how they might have evolved. The sRNA reads matching miRNA, tRNA, or ribosomal DNA sequences were removed from the datasets, and only those that mapped to the genome and were of the characteristic siRNA length (21-24 nt) were retained. The vast majority of siRNAs expressed during the rust fungus infection were found to be unique to one stage, while few siRNAs were common to all three samples (Figure 1A). Thus, infections could therefore create stage-specific siRNAs, producing a more complex and targeted siRNA population.

We used an approach similar to that reported by Barber et al. (2012) [15] to identify the types of genetic features associated with the stage-specific differences in siRNAs. We grouped overlapping 21-24-nt siRNAs within a ≤ 100 -bp window on the *P. trichocarpa* genome into clusters based on their location within the genome. Totally, we identified 301,155 siRNA clusters. The relative abundances of the siRNAs clusters were calculated (tpm: transcripts per million mapped reads). We then classified and annotated the siRNA clusters according to their common siRNA length

(21, 22, or 24 nt) and their location in the genome. The 22-nt clusters accounted for less than 10% of all the clusters, and the 24-nt clusters accounted the largest proportion (> 50%). The siRNA clusters with an abundance of at least 5 tpm in at least one stage were retained for the downstream analyses ([Data S1](#)).

Next, we examined the types of locations of the siRNAs in relation to the regions of the poplar genome, including the 5' UTRs, 3' UTRs, introns, and upstream (2-kb) and downstream (2-kb) sequences of the coding genes, transposons, pseudogenes, and intergenic regions. Of the siRNA clusters, 625 (4%) were located within pseudogenes, 2,973 (18%) were in upstream regions, 2,586 (15%) were in downstream regions, 330 (2%) were in 5' UTR regions, 395 (2%) were in 3' UTR regions, 982 (6%) were in introns, 8,159 (48%) were in transposons, and 4,703 (19%) were in intergenic regions ([Figure 1](#)).

The siRNA clusters were enriched for TE (transposon elements)-associated sequences, as shown by the large percentages of siRNAs derived from the TE sequences ([Figure 1B](#)). Many of the TE sequences were annotated as high-copy retrotransposon families ([Data S2](#)); for example, 26.9% (2,556) were long-terminal-repeat (LTR) sequences. Notably, the proportions of siRNA clusters located in pseudogenes and transposons were significantly larger than would be expected by chance ($P < 0.001$, one-side z test; [Table 1](#)). This suggests that pseudogenes and retrotransposon TEs contributed to the origin of some *Populus* siRNAs.

Infection stage-specific differences in the siRNAs primarily originate from the

transposon regions

In the samples from the two infection stages, the proportions of 21-nt and 22-nt siRNAs included in any given cluster were significantly correlated (total $r = 0.96$; T01 $r = 0.93$; T02 $r = 0.93$; T03 $r = 0.94$; $P < 2.2e-16$), and their proportions were also positively correlated with the 24-nt siRNAs (total $r = 0.82$ and $r = 0.85$ for 21-nt and 22-nt siRNAs, respectively; T01 $r = 0.64$ and $r = 0.75$, respectively; T02 $r = 0.64$ and $r = 0.75$, respectively; T03 $r = 0.66$ and $r = 0.76$, respectively; $P < 2.2e-16$) ([Data S1](#)). The 21-nt clusters have a shorter mean length (205 bp) and were more abundant (80 tpm) than the 24-nt clusters (370 bp, 17 tpm; Wilcoxon rank-sum P -value $< 2.2e-16$). To investigate the clusters exhibiting large stage-specific differences, we ordered the clusters by their absolute value of fold change values (T02/T01 or T03/T01). The degree of difference in the 24-nt siRNA cluster abundances between stages increased for those located in the transposon and intergenic regions ([Figure 1C and 1D](#)), with the strongest trend observed in the transposon regions.

To further investigate the global differences in siRNA abundance in the retrotransposon families between the different infection stages, we mapped the 21-nt, 22-nt, and 24-nt siRNAs that perfectly matched the *Populus* genomes onto the retrotransposons. [Figure 2A](#) shows the retrotransposon families that showed a siRNA abundance of at least 100 tpm in at least one of the stages. The production of 21-nt siRNA lengths and the abundance of these siRNAs located in the retrotransposon families are more similar across the three stages than those of the 22-nt and 24-nt siRNAs.

The stage-specific differences in the abundance of siRNAs located in the transposons were examined using a X^2 test for each siRNA length. The differences in the siRNA abundance for most transposon families between the stages (T02 vs T01; T03 vs T01) was contingent on the 24-nt length sequences ($P < 0.05$; Chi-square test; Data S3; Figure 2A). The stage-specific differences in siRNA abundance for these retrotransposon families thus primarily resulted from the 24-nt siRNAs.

***Populus* siRNAs mediate plant–pathogen interactions, as revealed using a shotgun strategy**

To explore the potential role of the *Populus* siRNAs in plant–pathogen interactions, we profiled the siRNA sequences that mapped to the *P. trichocarpa* genome which could regulate the genes of the fungal pathogen *Melampsora larici-populina*. In the biotrophic growth (T02) and urediniospore formation and release (T03) stages, more than 95.0% of the fungal genes (T02: 15,954; T03: 15,472; total fungal genes: 16,372) were predicted to be targeted by siRNA sequences (Data S4). Nearly one-third of the expression of *Melampsora larici-populina* genes detected in our transcriptome were altered at least by twofold between sample pairs (T02/T01; T03/T01; T02/T01) (Data S5). We hypothesize that some siRNAs could suppress pathogen pathogenicity by targeting pathogen virulence genes using a shotgun strategy. Importantly, if the siRNA-pathogene targets not in a random match manner, we would expect to see the targets of siRNAs may bias some pathogen genes. As expected, compared with the control sample (T01), pathogen effectors were targeted by more siRNAs at the biotrophic growth and urediniospore formation and release phases ($P < 0.001$;

permutation test; [Figure 2B](#)), suggesting there is a clear selection for the siRNA-target interactions.

The *Populus*-specific miRNAs are more involved in the regulation of the disease-resistance (DR) genes

We next analyzed the presence of miRNAs in the sequencing data, as miRNAs have been shown to play critical roles in plant defense systems [45, 46]. We annotated the sRNAome by aligning the sRNA reads to Rfam (version 13). After removing the scRNAs, tRNAs, rRNAs, snRNAs, and snoRNAs, we identified 423 sRNAs precursors, which grouped into 211 miRNA families containing an average of 2.0 genes per family for each of the three stages. We found that 67.7% of the miRNAs could be annotated in miRBase (version 22); however, most of the putative novel miRNAs did not remain in the final dataset. These excluded miRNA sequences may not represent bona fide miRNAs or may be expressed at levels too low to reliably show that they have the characteristics of miRNAs (at least one read matching their star sequence; [Data S6](#)). The expression values of each miRNA gene were also calculated and normalized ([Data S7](#)).

Using the miRNAs identified in this study and the miRNAs of the closely related species annotated in miRBase (version 22), we first sought to identify the *Populus*-specific miRNAs. Of the 211 miRNA families, 181 were *Populus*-specific, and 70 were supported by the findings of a previous study ([Figure 3A](#))[47]. We then predicted the potential miRNA targets using psRNATarget [42]. A total of 65,777 predicted miRNA-target gene pair matches with an expectation score of 5 were

identified (Data S8). An examination of the expression of the miRNA/target genes revealed 1,940 pairs with negatively correlated expression patterns ($P < 0.05$; Pearson correlation; Data S9). Using the publicly available degradome library [48], we identified 553 miRNA/target pairs representing 165 miRNAs (Data S10), of which several conserved miRNA/target pairs (miR156-*SPL*, miR172-*AP2*, miR397-*LAC*) were observed in our dataset [49-51]. When the predicted DR gene targets of miRNAs in both the conserved and *Populus*-specific datasets were excluded, a total of 161 DR targets remained for the conserved miRNAs and 711 DR targets for *Populus*-specific miRNAs. More than 66% (120 of 181) of the *Populus*-specific miRNAs target DR genes, a significantly larger proportion than that of the conserved miRNAs (23.3%, $P = 0.009$; Fisher's exact test; Table 2, Data S11; Figure 3B).

To gain a better understanding of the functional roles of the *Populus*-specific miRNA families, we performed a pfam domain analysis of their predicted targets. The top 20 domains of the *Populus*-specific miRNA targets were highly correlated with the domains of the plant DR genes (Figure S1); for example, NB-ARC and LRR domains, which are known to dissociate upon effector-dependent activation [52], were overrepresented in the samples.

***Populus*-specific miR6579 and miR6590 target numerous NB-LRR genes**

Plant NB-LRRs (encoded by *NB-LRR* genes) are commonly targeted by miRNAs, typically generating phasiRNAs (phased small-interfering RNAs), which can reduce the levels of these targets in *cis* or in *trans* [48]. In the two stages of the rust fungus infection, a conserved miRNA superfamily composed of miR482 targeted the

NB-LRRs at the encoded and conserved P-loop motifs. This family was found to be highly conserved among divergent plant species, including rice, *Arabidopsis*, grape (*Vitis vinifera*), and soybean (*Glycine max*) [23, 53, 54]. A second conserved family, miR2111, was identified in the eudicotyledons (Figure 3C), indicating the existence of this miRNA family prior to the divergence of these species from a common ancestor around 103 Mya [55]. The highly conserved 20-nt miR482s were predicted to target 221 *NB-LRRs*. Two related *Populus*-specific families, miR6579 and miR6590, which both contain 21-nt miRNAs, were predicted to target 132 and 49 *NB-LRRs*, respectively (Figure 3D-F). These two *Populus*-specific miRNA families had the highest transcript abundance during the mock-inoculated control plants, decreasing to a low level during the biotrophic growth (T02) and urediniospore formation and release (T03) stages (Data S7). By contrast, the highly conserved 22-nt miR2111s were predicted to target only nine *NB-LRRs*, three of which were also predicted targets of miR482 while two were predicted to be targeted by miR6579 (Figure 3D-F). These results indicate that the *NB-LRRs* in *Populus* are primarily targeted by conserved miR482s and *Populus*-specific miR6579s.

A careful examination of their precursor sequences revealed that miR6579 and miR6590 were intron-derived and pseudogene-derived miRNAs, respectively (Figure 3E). The parent of the pseudogene Chr03|4847214-4847421 was a *NB-LRR* gene (Potri.001G426600), suggesting that pseudogenes could act as catalysts for the formation of miRNAs. In poplar, the 21-nt miR6590 family showed characteristics canonical of the typical miRNA triggers associated with ARGONAUTE1 (AGO1) for

the initiation of the PHAS loci, such as “U” in the 5’ position, “A” in position 10, and “C” in the 3’ position, indicating that this miRNA may trigger the biogenesis of secondary siRNAs [56]. It seems that species-specific miRNAs could become fixed in the regulatory network once a number of targets have evolved for which the presence of a miRNA-binding site would be beneficial.

Discussion

Transposons and posttranscriptional regulation

Despite previously being referred to as junk DNA, transposons are now known to be essential elements of most eukaryotic genomes, making important contributions to their structure, diversity, capacity, and adaptation [57]. The widely distributed transposable components are the most rapidly evolving part of the genome because they are more susceptible to sequence change than coding sequences [58]. Here, while exploring the distribution of poplar siRNAs in different parts of genome regions, including the 5’ UTRs, 3’ UTRs, introns, upstream (2 kb) and downstream (2 kb) of the coding genes, transposons, pseudogenes, and intergenic regions, we determined that transposons contributed a large proportion of the siRNAs (48%). In particular, the siRNA clusters were enriched in genomic locations comprising high-copy retrotransposon families. We identified significant variation in the 24-nt siRNAs derived from specific retrotransposon families between the different infection stages. Based on their length and sequence similarity to the retrotransposon families from which they are derived, these siRNAs may act posttranscriptionally to mediate

repressive histone modification, thus contributing to the reinforcement of the silenced state of TE chromatin [59, 60].

Our findings further illustrate that it is difficult to assess the connection between siRNAs and transposons in the process of plant defense. In general, the mobilization of transposons can be mutagenic [61], so host genomes have evolved elaborate mechanisms to suppress their activities. Most of the 24-nt siRNAs are likely to be involved in the transcriptional regulation of transposons through RNA-directed DNA methylation [62], but many TE-associated siRNAs were also found to influence non-TE transcripts exerting certain regulatory roles [63]; for example, 24-nt siRNAs were found to target pathogen genes [61]. Transposons are rapidly activated by biotic and abiotic stresses however [64-66], potentially complicating the regulatory loop between the siRNAs and TEs.

siRNAs are an ideal tool for use in plant defense because of their target specificity and highly efficient responses [67]. Ohtsu et al. (2007) [68] proposed that the retrotransposon-associated siRNAs may target genes with homologous sequences. Also, direct evidence in *Arabidopsis* suggests that 21-nt siRNAs derived from an Athila retrotransposon could posttranscriptionally regulate the *UBP1b* (Upstream binding protein 1b) gene to mediate the stress response [69]. In rice, transposon-derived 24-nt siRNAs could target the genes involved in hormone homeostasis, suggesting that transposon-associated regulatory modules could play a variety of roles in plant development and defense [70]. It is therefore possible that the 21–24-nt siRNAs may have another purpose besides the maintenance of genome

stability, and may enhance plant defense by targeting the pathogen genes. Considering the high genetic diversity and mobilization of transposons, the regulatory system mediated by the TE-derived siRNAs could play a major role in the performance of plant defense.

Cross-kingdom RNAi and sRNAs mediate plant–pathogen interactions

An endless arms race drives host-pathogen coevolution [71]. In plants, pathogens can invoke multiple layers of immune responses, and many defense genes were found to play key roles in this process [72, 73]. Most potential pathogens are blocked by nonhost resistance and pathogen-associated molecular pattern (PAMP)-triggered immunity (PTI) [67]; however, many pathogens deliver a variety of effectors into plant cells to suppress host immunity. To counteract this attack, plants have evolved resistance proteins and immune mechanisms to sense these effectors and trigger a robust resistance response [74]. Plant sRNAs are essential regulatory noncoding RNAs that can posttranscriptionally silence their target genes [75, 76]. Recent progress in profiling sRNAs, especially advances in next-generation sequencing technology, has revealed their extensive and complicated involvement in interactions between plants and viruses, bacteria, and fungi [75].

Fungal pathogens deliver sRNAs into plant cells to induce cross-kingdom RNAi of the plant immunity genes [77]. A similar strategy was also observed in animal-parasite systems [78]. Furthermore, sRNAs generated from the host plant cells can also be transferred into pathogen cells [71], as fungal pathogens are known to be capable of taking up external sRNAs and long dsRNAs [71, 79]. Pathogen genomes

encode variable effector repertoires for suppressing host immunity [80]; specifically, these effectors are often encoded in genomic regions of high plasticity [81], such as TEs [82]. Given that effectors are often clustered in these regions, TEs must function as an evolutionary force shaping these virulence factors by contributing to their diversification [82]. Sequence diversification is particularly important for pathogen effectors, which are essential components facilitating their coevolution with, and outmaneuvering of, the host.

Considering the high diversification of virulence factors generated by pathogens, it is reasonable to assume that transposon-derived sRNAs are also the best sources of new regulators for plant defense resistance, facilitating cross-kingdom RNAi. TEs, the most abundant component of most eukaryote genomes, are hotspots for generating sRNAs, which is particularly true of the retrotransposons [83]. Here, we found that pathogen effectors were targeted by more poplar siRNAs in the biotrophic growth and urediniospore formation and release stages of infection by a rust fungus, in comparison with control. It is therefore reasonable to conclude that the activation of the siRNA repertoire does not occur randomly, and that their targets are instead involved in more central roles in the plant–pathogen interactions. Supporting this hypothesis, a previous study reported that the transfer of host sRNAs could silence virulence-related genes and suppress fungal pathogenicity in *Arabidopsis* [84].

Disease resistance associated with Populus-specific miRNAs

The plant *NB-LRR* genes mediate effector-triggered immunity by acting as key receptors during the innate immunity response against a wide range of pests and

diseases [85]. *NB-LRR* genes are generally grouped into two subclasses: the Toll/Interleukin-1 receptor-like group (TIR-NB-LRRs) and a coiled-coil domain-containing group (CC-NB-LRRs) [86]. Both classes can be triggered by miRNAs to generate phasiRNAs, which can reduce the levels of the transcripts of their targets in *cis* and in *trans* [87]. In contrast to low-copy genes, many *NB-LRR* genes have undergone dramatic duplications and losses, domain architecture variations, the partitioning of subfamilies, and copy number variation among species [85, 88, 89]. Thus, the *NB-LRR* genes are highly variable, lineage-specific, and associated with the plant immune response, providing material to allow rapid adaptive evolution.

The target sites of conserved miRNAs are often located within the highly conserved domains of the target genes [90]. Some of these conserved domains, such as the P-loop motifs, are targeted by several miRNA families [56]. Unlike conserved miRNAs, newly evolved miRNAs tend not to target these conserved functional domains [90], and may instead target mRNAs simply by chance [91]. The recent salicoid whole genome duplication (WGD) event, which was shared among the Salicaceae, resulting in overrepresentation of genes associated with special functional categories [92]. Indeed, the target genes of the newly emerged miRNAs in *Populus* were found to have various functions, including a large number of DR *NB-LRR* genes, ABC transporters, and a few transcription factors, consistent with the retained salicoid duplicates observed in this genus [30]. Our analysis demonstrated that more *Populus*-specific miRNAs than conserved miRNAs target the *NB-LRR* genes,

although the species-specific miRNAs have fewer targets than the conserved miRNAs. Considering that this *de novo* diversity may be associated with plant defense, these new miRNAs have a greater potential to target newly evolved plant DR genes, further contributing to the phenotypic innovation of the host. Once the newly emerged miRNAs become fixed in the regulatory modules, they may gradually evolve to target more genes linked to their specific function [48, 91]. Our study therefore provides insights into the origination of siRNAs and the regulation of siRNAs in the defense process.

Conclusions

In this study, we sequenced the sRNAome of Chinese white poplar during the two essential stages associated with plant colonization and biotrophic growth in rust fungi (*Melampsora larici-populina*). We identified a set of distinct sRNAs for the two infection stages and the mock-inoculated control plants. Compared with the control plants, more poplar siRNAs in the biotrophic growth and urediniospore formation and release stages targeted pathogen effector genes. Furthermore, there are more *Populus*-specific miRNAs than conserved miRNAs targeting the NB-LRR genes, although the species-specific miRNAs have fewer targets than the conserved miRNAs. The results reveal the evolutionary importance and functional novelty of the sRNAs, would provide valuable information on comprehensive understanding of the origin, evolution, and plant defense effects of the poplar sRNAs.

Declarations

Acknowledgements

Not applicable.

Author Contributions

LSJ, LX collected samples; LSJ analyzed the data and wrote the manuscript; ZYY, XJB revised the manuscript, and XJB obtained funding and is responsible for this article.

Funding

This work was supported by the Project of the National Natural Science Foundation of China (32022057 and 31972954), Young Elite Scientists Sponsorship Program by CAST (2018QNRC001) and Forestry and Grassland Science and Technology Innovation Youth Top Talent Project of China (No. 2020132607).

Availability of data and materials

Raw data of RNA-sequencing and small RNA sequencing are available for download at the BIGD Genome Sequence Archive under accession number CRA001965.

Ethics approval and consent to participate

Not applicable.

Consent for publication

All authors agreed to publish.

Competing interests

The authors declare that they have no competing interests.

Author details

¹Beijing Advanced Innovation Center for Tree Breeding by Molecular Design, Beijing Forestry University, No. 35, Qinghua East Road, Beijing 100083, P. R. China;

²National Engineering Laboratory for Tree Breeding, College of Biological Sciences and Technology, Beijing Forestry University, No. 35, Qinghua East Road, Beijing 100083, P. R. China;

³Key Laboratory of Genetics and Breeding in Forest Trees and Ornamental Plants, Ministry of Education, College of Biological Sciences and Technology, Beijing Forestry University, No. 35, Qinghua East Road, Beijing 100083, P. R. China;

Emails

LSJ: liusijia_1@163.com ; LX: lixianwzx@163.com; ZYY: yiyangzhao@bjfu.edu.cn;

XJB: E-mail: jbxie@bjfu.edu.cn;

Reference:

1. Matzke MA, Birchler JA: **RNAi-mediated pathways in the nucleus.** *Nat Rev Genet* 2005, **6**(1):24-35.
2. Po-Pu L, Montgomery TA, Noah F, Kasschau KD, Hiroyuki N, Carrington JC: **Repression of auxin response factor10 by microRNA160 is critical for seed germination and post-germination stages.** *Plant Journal for Cell & Molecular Biology* 2010, **52**(1):133-146.
3. Ghildiyal M, Zamore P: **Small silencing RNAs: an expanding universe.** *Nat Rev Genet* 2009, **10**(2):94-108.
4. Ivan O, Aravin AA, Katalin FT: **Small RNA in the nucleus: the RNA-chromatin ping-pong.** *Curr Opin Genet Dev* 2012, **22**(2):164-171.
5. Carthew RW, Sontheimer EJ: **Origins and Mechanisms of miRNAs and siRNAs.** *Cell* 2009, **136**(4):642-655.
6. Nykänen A, Haley B, Zamore PD: **ATP Requirements and Small Interfering RNA Structure in the RNA Interference Pathway.** *Cell* 2001, **107**(3):309-321.
7. Harding JL, Horswell S, Heliot C, Armisen J, Zimmerman LB, Luscombe NM, Miska EA, Hill CS: **Small RNA profiling of *Xenopus* embryos reveals novel miRNAs and a new class of small RNAs derived from intronic transposable elements.** *Genome Res* 2014, **24**(1):96-106.
8. Johnson W, Jameson JL: **Transcriptional control of gene expression.** *Cell* 2010, **140**(1):111-122.
9. Mallory AC, Vaucheret H: **Functions of microRNAs and related small RNAs in plants.** *Nat Genet* 2006, **38**(Suppl):S31.
10. Ramanjulu S, Jian-Kang Z: **Novel and stress-regulated microRNAs and other small RNAs from *Arabidopsis*.** *Plant Cell* 2004, **16**(8):2001-2019.
11. Hamilton AJ, Baulcombe DC: **A species of small antisense RNA in posttranscriptional gene silencing in plants.** *Science* 1999, **286**(5441):950-952.
12. Misook H, Jie L, Lu T, Vanitharani R, Kasschau KD, Chapman EJ, Carrington JC, Xuemei C, Xiu-Jie W, Z Jeffrey C: **Small RNAs serve as a genetic buffer against genomic shock in *Arabidopsis* interspecific hybrids and allopolyploids.** *P Natl Acad Sci USA* 2009, **106**(42):17835-17840.
13. Guangming H, Xiaopeng Z, Elling AA, Liangbi C, Xiangfeng W, Lan G, Manzhong L, Hang H, Huiyong Z, Fangfang C: **Global epigenetic and transcriptional trends among two rice subspecies and their reciprocal hybrids.** *Plant Cell* 2010, **22**(1):17-33.
14. Michal KE, Dena L, Lior T, Elad N, Cathy MB, Moshe F, Levy AA: **Wheat hybridization and polyploidization results in deregulation of small RNAs.** *Genetics* 2011, **188**(2):263-272.
15. Barber WT, Zhang W, Win H, Varala KK, Dorweiler JE, Hudson ME, Moose SP: **Repeat associated small RNAs vary among parents and following hybridization in maize.** *Proceedings of the National Academy of Sciences* 2012, **109**(26):10444-10449.
16. Peláez P, Sanchez F: **Small RNAs in plant defense responses during viral and bacterial interactions: similarities and differences.** *Front Plant Sci* 2013, **4**(343):343.

17. James W, Lei G, Yang Y, Jixian Z, Siwaret A, Yu Y, Shuyi D, Vicky C, Qin X, Jun Y: **Roles of small RNAs in soybean defense against *Phytophthora sojae* infection.** *Plant Journal for Cell & Molecular Biology* 2015, **79**(6):928-940.
18. Pablo P, Federico S: **Small RNAs in plant defense responses during viral and bacterial interactions: similarities and differences.** *Front Plant Sci* 2013, **4**(343):343.
19. Nathan P, Olivier V: **RNA silencing suppression by plant pathogens: defence, counter-defence and counter-counter-defence.** *Nature Reviews Microbiology* 2013, **11**(11):745-760.
20. Navarro L, Dunoyer P, Jay F, Arnold B, Dharmasiri N, Estelle M, Voinnet O, Jones JD: **A plant miRNA contributes to antibacterial resistance by repressing auxin signaling.** *Science* 2006, **312**(5772):436-439.
21. Zhang W, Gao S, Zhou X, Chellappan P, Chen Z, Zhou X, Zhang X, Fromuth N, Coutino G, Coffey M: **Bacteria-responsive microRNAs regulate plant innate immunity by modulating plant hormone networks.** *Plant Mol Biol* 2011, **75**(1-2):93-105.
22. Yan L, Qingqing Z, Jiangguang Z, Liang W, Yijun Q, Jian-Min Z: **Identification of microRNAs involved in pathogen-associated molecular pattern-triggered plant innate immunity.** *Plant Physiol* 2010, **152**(4):2222-2231.
23. Zhao M, Cai C, Zhai J, Lin F, Li L, Shreve J, Thimmapuram J, Hughes TJ, Meyers BC, Ma J: **Coordination of microRNAs, phasiRNAs, and NB-LRR genes in response to a plant pathogen: insights from analyses of a set of soybean Rps gene near-isogenic lines.** *The Plant Genome* 2015, **8**(1).
24. Diaz-Pendon JA, Li F, Li W-X, Ding S-W: **Suppression of antiviral silencing by cucumber mosaic virus 2b protein in *Arabidopsis* is associated with drastically reduced accumulation of three classes of viral small interfering RNAs.** *Plant Cell* 2007, **19**(6):2053-2063.
25. Katiyar-Agarwal S, Morgan R, Dahlbeck D, Borsani O, Villegas A, Zhu J-K, Staskawicz BJ, Jin H: **A pathogen-inducible endogenous siRNA in plant immunity.** *Proceedings of the National Academy of Sciences* 2006, **103**(47):18002-18007.
26. Xie Z, Johansen LK, Gustafson AM, Kasschau KD, Lellis AD, Zilberman D, Jacobsen SE, Carrington JC: **Small RNA Pathways in Plants.** 2004.
27. Xie JB, Tian JX, Du QZ, Chen JH, Li Y, Yang XH, Li BL, Zhang DQ: **Association genetics and transcriptome analysis reveal a gibberellin-responsive pathway involved in regulating photosynthesis.** *J Exp Bot* 2016, **67**(11):3325-3338.
28. Persson S, Wei H, Milne J, Page GP, Somerville CR: **Identification of genes required for cellulose synthesis by regression analysis of public microarray data sets.** *Proceedings of the National Academy of Sciences* 2005, **102**(24):8633-8638.
29. Denancé N, Sánchez-Vallet A, Goffner D, Molina A: **Disease resistance or growth: the role of plant hormones in balancing immune responses and fitness costs.** *Front Plant Sci* 2013, **4**:155.
30. Tuskan GA, DiFazio S, Jansson S, Bohlmann J, Grigoriev I, Hellsten U, Putnam N, Ralph S, Rombauts S, Salamov A *et al*: **The genome of black cottonwood, *Populus trichocarpa* (Torr. & Gray).** *Science* 2006, **313**(5793):1596-1604.
31. Ma T, Wang J, Zhou G, Yue Z, Hu Q, Chen Y, Liu B, Qiu Q, Wang Z, Zhang J: **Genomic insights into salt adaptation in a desert poplar.** *Nature communications* 2013, **4**:2797.

32. Du Q, Wang B, Wei Z, Zhang D, Li B: **Genetic diversity and population structure of Chinese white poplar (*Populus tomentosa*) revealed by SSR markers.** *J Hered* 2012, **103**(6):853-862.
33. Hacquard S, Joly DL, Lin Y-C, Tisserant E, Feau N, Delaruelle C, Legué V, Kohler A, Tanguay P, Petre B: **A comprehensive analysis of genes encoding small secreted proteins identifies candidate effectors in *Melampsora larici-populina* (poplar leaf rust).** *Mol Plant Microbe In* 2012, **25**(3):279-293.
34. Dodds PN, Rathjen JP: **Plant immunity: towards an integrated view of plant–pathogen interactions.** *Nat Rev Genet* 2010, **11**(8):539.
35. Huang J, Wang C, Li X, Fang X, Huang N, Wang Y, Ma H, Wang Y, Copenhaver GP: **Conservation and Divergence in the Meiocyte sRNAomes of Arabidopsis, Soybean, and Cucumber.** *Plant Physiology* 2020, **182**(1):301-317.
36. He J, Jiang Z, Gao L, You C, Ma X, Wang X, Xu X, Mo B, Chen X, Liu L: **Genome-Wide Transcript and Small RNA Profiling Reveals Transcriptomic Responses to Heat Stress.** *Plant Physiology* 2019, **181**(2):609-629.
37. Kim D, Pertea G, Trapnell C, Pimentel H, Kelley R, Salzberg SL: **TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions.** *Genome Biology* 2013, **14**(4):R36.
38. Ghosh S, Chan C-KK: **Analysis of RNA-Seq Data Using TopHat and Cufflinks.** 2016, **1374**:339-361.
39. Sperschneider J, Gardiner DM, Dodds PN, Tini F, Covarelli L, Singh KB, Manners JM, Taylor JM: **EffectorP: predicting fungal effector proteins from secretomes using machine learning.** *New Phytol* 2016, **210**(2):743-761.
40. Friedländer MR, Mackowiak SD, Li N, Chen W, Rajewsky N: **miRDeep2 accurately identifies known and hundreds of novel microRNA genes in seven animal clades.** *Nucleic Acids Research* 2012, **40**(1):37-52.
41. Johnson C, Kasprzewska A, Tennessen K, Fernandes J, Nan G-L, Walbot V, Sundaresan V, Vance V, Bowman LH: **Clusters and superclusters of phased small RNAs in the developing inflorescence of rice.** *Genome Res* 2009, **19**(8):1429-1440.
42. Dai X, Zhuang Z, Zhao PX: **psRNATarget: a plant small RNA target analysis server (2017 release).** *Nucleic Acids Res* 2018, **46**(W1):W49-W54.
43. Xie J, Li Y, Liu X, Zhao Y, Li B, Ingvarsson PK, Zhang D: **Evolutionary origins of pseudogenes and their association with regulatory sequences in plants.** *Plant Cell* 2019:tpc.00601.02018.
44. Duplessis S, Hacquard S, Delaruelle C, Tisserant E, Frey P, Martin F, Kohler A: ***Melampsora larici-populina* transcript profiling during germination and timecourse infection of poplar leaves reveals dynamic expression patterns associated with virulence and biotrophy.** *Molecular plant-microbe interactions : MPMI* 2011, **24**(7):808.
45. Katiyaragarwal S, Morgan R, Dahlbeck D, Borsani O, Villegas A, Zhu JK, Staskawicz BJ, Jin H: **A pathogen-inducible endogenous siRNA in plant immunity.** *P Natl Acad Sci USA* 2006, **103**(47):18002-18007.
46. Weiberg A, Wang M, Lin FM, Zhao H, Zhang Z, Kaloshian I, Huang HD, Jin H: **Fungal Small RNAs Suppress Plant Immunity by Hijacking Host RNA Interference Pathways.** *Science* 2013, **342**(6154):118-123.

47. Puzey JR, Karger A, Axtell M, Kramer EM: **Deep Annotation of Populus trichocarpa microRNAs from Diverse Tissue Sets.** *PLoS ONE* 2012, **7**(3):e33034-.
48. Xie J, Yang X, Song Y, Du Q, Li Y, Chen J, Zhang D: **Adaptive evolution and functional innovation of Populus-specific recently evolved microRNAs.** *New Phytol* 2017, **213**(1):206-219.
49. Lu S, Li Q, Wei H, Chang M-J, Tunlaya-Anukit S, Kim H, Liu J, Song J, Sun Y-H, Yuan L: **Ptr-miR397a is a negative regulator of laccase genes affecting lignin content in Populus trichocarpa.** *Proceedings of the National Academy of Sciences* 2013, **110**(26):10848-10853.
50. Nova-Franco B, Íñiguez LP, Valdés-López O, Alvarado-Affantranger X, Leija A, Fuentes SI, Ramírez M, Paul S, Reyes JL, Girard L: **The micro-RNA172c-APETALA2-1 node as a key regulator of the common bean-Rhizobium etli nitrogen fixation symbiosis.** *Plant Physiol* 2015, **168**(1):273-291.
51. Wang H, Wang H: **The miR156/SPL module, a regulatory hub and versatile toolbox, gears up crops for enhanced agronomic traits.** *Molecular plant* 2015, **8**(5):677-688.
52. Sloatweg EJ, Spiridon LN, Roosien J, Butterbach P, Pomp R, Westerhof L, Wilbers R, Bakker E, Bakker J, Petrescu AJ: **Structural Determinants at the Interface of the ARC2 and Leucine-Rich Repeat Domains Control the Activation of the Plant Immune Receptors Rx1 and Gpa2.** *Plant Physiol* 2013, **162**(3):1510-1528.
53. Lu C, Kulkarni K, Souret FF, MuthuValliappan R, Tej SS, Poethig RS, Henderson IR, Jacobsen SE, Wang W, Green PJ: **MicroRNAs and other small RNAs enriched in the Arabidopsis RNA-dependent RNA polymerase-2 mutant.** *Genome Res* 2006, **16**(10):1276-1288.
54. Subramanian S, Fu Y, Sunkar R, Barbazuk WB, Zhu J-K, Yu O: **Novel and nodulation-regulated microRNAs in soybean roots.** *Bmc Genomics* 2008, **9**(1):160.
55. Massoni J, Couvreur TL, Sauquet H: **Five major shifts of diversification through the long evolutionary history of Magnoliidae (angiosperms).** *BMC Evol Biol* 2015, **15**(1):49.
56. Zhao M, Meyers BC, Cai C, Xu W, Ma J: **Evolutionary patterns and coevolutionary consequences of MIRNA genes and microRNA targets triggered by multiple mechanisms of genomic duplications in soybean.** *Plant Cell* 2015, **27**(3):546-562.
57. Le QH, Wright S, Yu Z, Bureau T: **Transposon diversity in Arabidopsis thaliana.** *Proceedings of the National Academy of Sciences* 2000, **97**(13):7376-7381.
58. Tomita M: **Revolver and superior: novel transposon-like gene families of the plant kingdom.** *Current genomics* 2010, **11**(1):62-69.
59. Zilberman D, Cao X, Johansen LK, Xie Z, Carrington JC, Jacobsen SE: **Role of Arabidopsis ARGONAUTE4 in RNA-directed DNA methylation triggered by inverted repeats.** *Curr Biol* 2004, **14**(13):1214-1220.
60. Zhong X, Du J, Hale CJ, Gallego-Bartolome J, Feng S, Vashisht AA, Chory J, Wohlschlegel JA, Patel DJ, Jacobsen SE: **Molecular mechanism of action of plant DRM de novo DNA methyltransferases.** *Cell* 2014, **157**(5):1050-1060.
61. Lisch D: **Epigenetic Regulation of Transposable Elements in Plants.** *Annu Rev Plant Biol* 2009, **60**(1):43.
62. Matzke MA, Birchler JA: **RNAi-mediated pathways in the nucleus.** *Nat Rev Genet* 2005, **6**(1):24.
63. Cho J: **Transposon-Derived Non-coding RNAs and Their Function in Plants.** *Front Plant*

- Sci* 2018, **9**.
64. Grandbastien M-A, Audeon C, Bonnivard E, Casacuberta J, Chalhoub B, Costa A-P, Le Q, Melayah D, Petit M, Poncet C: **Stress activation and genomic impact of Tnt1 retrotransposons in Solanaceae**. *Cytogenetic and genome research* 2005, **110**(1-4):229-241.
 65. Cavrak VV, Lettner N, Jamge S, Kosarewicz A, Bayer LM, Scheid OM: **How a retrotransposon exploits the plant's heat stress response for its activation**. *PLoS Genet* 2014, **10**(1):e1004115.
 66. Matsunaga W, Ohama N, Tanabe N, Masuta Y, Masuda S, Mitani N, Yamaguchi-Shinozaki K, Ma JF, Kato A, Ito H: **A small RNA mediated regulation of a stress-activated retrotransposon and the tissue specific transposition during the reproductive period in Arabidopsis**. *Front Plant Sci* 2015, **6**:48.
 67. Ito H: **Small RNAs and regulation of transposons in plants**. *Genes Genet Syst* 2013, **88**(1):3-7.
 68. Ohtsu K, Smith MB, Emrich SJ, Borsuk LA, Zhou R, Chen T, Zhang X, Timmermans MC, Beck J, Buckner B: **Global gene expression analysis of the shoot apical meristem of maize (Zea mays L.)**. *The Plant Journal* 2007, **52**(3):391-404.
 69. McCue AD, Nuthikattu S, Reeder SH, Slotkin RK: **Gene expression and stress response mediated by the epigenetic regulation of a transposable element small RNA**. *PLoS Genet* 2012, **8**(2):e1002474.
 70. Wei L, Gu L, Song X, Cui X, Lu Z, Zhou M, Wang L, Hu F, Zhai J, Meyers BC: **Dicer-like 3 produces transposable element-associated 24-nt siRNAs that control agricultural traits in rice**. *Proceedings of the National Academy of Sciences* 2014, **111**(10):3877-3882.
 71. Wang M, Weiberg A, Lin F-M, Thomma BP, Huang H-D, Jin H: **Bidirectional cross-kingdom RNAi and fungal uptake of external RNAs confer plant protection**. *Nature plants* 2016, **2**(10):16151.
 72. Mackintosh CA, Lewis J, Radmer LE, Shin S, Heinen SJ, Smith LA, Wyckoff MN, Dill-Macky R, Evans CK, Kravchenko S: **Overexpression of defense response genes in transgenic wheat enhances resistance to Fusarium head blight**. *Plant cell reports* 2007, **26**(4):479-488.
 73. Weiberg A, Wang M, Bellinger M, Jin H: **Small RNAs: a new paradigm in plant-microbe interactions**. *Annu Rev Phytopathol* 2014, **52**:495-516.
 74. Sin YW, Annavi G, Dugdale HL, Newman C, Burke T, MacDonald DW: **Pathogen burden, co - infection and major histocompatibility complex variability in the European badger (Meles meles)**. *Mol Ecol* 2014, **23**(20):5072-5088.
 75. Zhai J, Jeong D-H, De Paoli E, Park S, Rosen BD, Li Y, González AJ, Yan Z, Kitto SL, Grusak MA: **MicroRNAs as master regulators of the plant NB-LRR defense gene family via the production of phased, trans-acting siRNAs**. *Genes & development* 2011, **25**(23):2540-2553.
 76. Hidetaka I: **Small RNAs and regulation of transposons in plants**. *Genes Genet Syst* 2013, **88**(1):3-7.
 77. Weiberg A, Wang M, Lin F-M, Zhao H, Zhang Z, Kaloshian I, Huang H-D, Jin H: **Fungal small RNAs suppress plant immunity by hijacking host RNA interference pathways**. *Science* 2013, **342**(6154):118-123.
 78. Buck AH, Coakley G, Simbari F, McSorley HJ, Quintana JF, Le Bihan T, Kumar S,

- Abreu-Goodger C, Lear M, Harcus Y: **Exosomes secreted by nematode parasites transfer small RNAs to mammalian cells and modulate innate immunity.** *Nature communications* 2014, **5**:5488.
79. Hinas A, Wright AJ, Hunter CP: **SID-5 is an endosome-associated protein required for efficient systemic RNAi in *C. elegans*.** *Curr Biol* 2012, **22**(20):1938-1943.
80. Sébastien C, Suma C, Kvitko BH, Russell AB, Martin GB, Alan C: **Genetic disassembly and combinatorial reassembly identify a minimal functional repertoire of type III effectors in *Pseudomonas syringae*.** *P Natl Acad Sci USA* 2011, **108**(7):2975-2980.
81. Ninio S, Celli J, Roy CR: **A *Legionella pneumophila* effector protein encoded in a region of genomic plasticity binds to Dot/Icm-modified vacuoles.** *PLoS pathogens* 2009, **5**(1):e1000278.
82. Meile L, Croll D, Brunner PC, Plissonneau C, Hartmann FE, McDonald BA, Sánchez - Vallet A: **A fungal avirulence factor encoded in a highly plastic genomic region triggers partial resistance to septoria tritici blotch.** *New Phytol* 2018, **219**(3):1048-1061.
83. Lipardi C, Paterson BM, Singer MF: **Identification of an RNA-Dependent RNA Polymerase in *Drosophila* Involved in RNAi and Transposon Suppression.** *Proceedings of the National Academy of Sciences of the United States of America* 2009, **106**(37):15645-15650.
84. Cai Q, Qiao L, Wang M, He B, Lin F-M, Palmquist J, Huang S-D, Jin H: **Plants send small RNAs in extracellular vesicles to fungal pathogen to silence virulence genes.** *Science* 2018, **360**(6393):1126-1129.
85. Innes RW, Ameline-Torregrosa C, Ashfield T, Cannon E, Cannon SB, Chacko B, Chen NW, Couloux A, Dalwani A, Denny R: **Differential accumulation of retroelements and diversification of NB-LRR disease resistance genes in duplicated regions following polyploidy in the ancestor of soybean.** *Plant Physiol* 2008, **148**(4):1740-1759.
86. Jones JD, Dangl JL: **The plant immune system.** *Nature* 2006, **444**(7117):323.
87. Fei Q, Xia R, Meyers BC: **Phased, secondary, small interfering RNAs in posttranscriptional regulatory networks.** *Plant Cell* 2013, **25**(7):2400-2415.
88. Karasov TL, Horton MW, Bergelson J: **Genomic variability as a driver of plant-pathogen coevolution?** *Curr Opin Plant Biol* 2014, **18**:24-30.
89. Li Y-h, Zhou G, Ma J, Jiang W, Jin L-g, Zhang Z, Guo Y, Zhang J, Sui Y, Zheng L: **De novo assembly of soybean wild relatives for pan-genome analysis of diversity and agronomic traits.** *Nat Biotechnol* 2014, **32**(10):1045.
90. Rhoades MW, Reinhart BJ, Lim LP, Burge CB, Bartel B, Bartel DP: **Prediction of plant microRNA targets.** *Cell* 2002, **110**(4):513-520.
91. Chen K, Rajewsky N: **The evolution of gene regulation by transcription factors and microRNAs.** *Nat Rev Genet* 2007, **8**(2):93.
92. Rodgers-Melnick E, Mane SP, Dharmawardhana P, Slavov GT, Crasta OR, Strauss SH, Brunner AM, DiFazio SP: **Contrasting patterns of evolution following whole genome versus tandem duplication events in *Populus*.** *Genome Res* 2012, **22**(1):95-105.

Figure legends

Figure 1. The stage differences in siRNA abundance primarily originate from transposons in response to infection. Shown are the 21 and 24-nt siRNA clusters with at least 5 tpm (transcripts per million mapped reads). Clusters are arranged in ascending order of fold change. (A) The common and stage-specific siRNA clusters. (B) The relative abundance of the siRNA clusters in the various types of physical locations in the *Populus* genome. (C, D) Degree of stage-specific differences for the siRNA clusters in the different locations. The \log_{10} values of their abundance in T01 (The control group; 48 h) divided by their abundance in (C) T02 (biotrophic growth phase; 48 h post infection) or (D) T03 (urediniospore formation and release phase; 168 h post infection). Up indicate the 2 kb upstream region of genes and down indicates the 2 kb downstream region of genes. The 22-nt and 24-nt siRNA clusters with at least 5 tpm in the three samples are shown. The clusters below the horizontal line were negatively regulated, while those above the horizontal line were positively regulated. The clusters are arranged in ascending order of fold change.

Figure 2. The difference in transposon siRNAs activity are driven by 24-nt sRNAs. (A) The abundance of siRNAs in the three samples (T01; T02; T03). The total abundance column displays the relative abundance of 21-, 22-, and 24-nt siRNAs. (B) The number of pathogen effectors targeted by the siRNAs in the three infection samples. The effectors were predicted by EffectorP and the target prediction were performed by using psRNATarget. T01: The control group; 48 h. T02: biotrophic growth phase; 48 h post infection. T03: urediniospore formation and release phase;

168 h post infection. Error bars, bootstrap-based 95% confidence intervals on the mean estimates. Statistical inference is conducted with a permutation test on the mean (perm.test in the R package exact RankTests).

Figure 3. The *Populus NB-LRR* genes are targeted by *Populus*-specific miRNAs. (A) The species-specific miRNAs of nine plant species. ptc: *Populus trichocarpa*; mtr: *Medicago truncatula*; bra: *Brassica rapa*; sbi: *Sorghum bicolor*; osa: *Oryza sativa*; vvi: *Vitis vinifera*; ath: *Arabidopsis thaliana*; bdi: *Brachypodium distachyon*; gma: *Glycine max*. (B) Comparison of the number of predicted disease-resistance target genes of the conserved and *Populus*-specific miRNAs. (C) Conservation profile of miR482, miR2118, and miR2111 in diverse plant species. The mature miRNA sequences were retrieved from miRBase. (D) Comparison of the number of predicted *NB-LRR* genes targeted by miR482, miR2111, miR6579, and miR6590. (E) The genome locations of two *Populus*-specific miRNAs. (F) The consensus coding sequences of the targets of four *Populus* miRNAs.

Table legends

Table 1. Enrichment of siRNA clusters in various regions of *Populus trichocarpa* genome.

Table 2. Comparison of the targets of *Populus*-specific and conserved miRNAs.

Figures

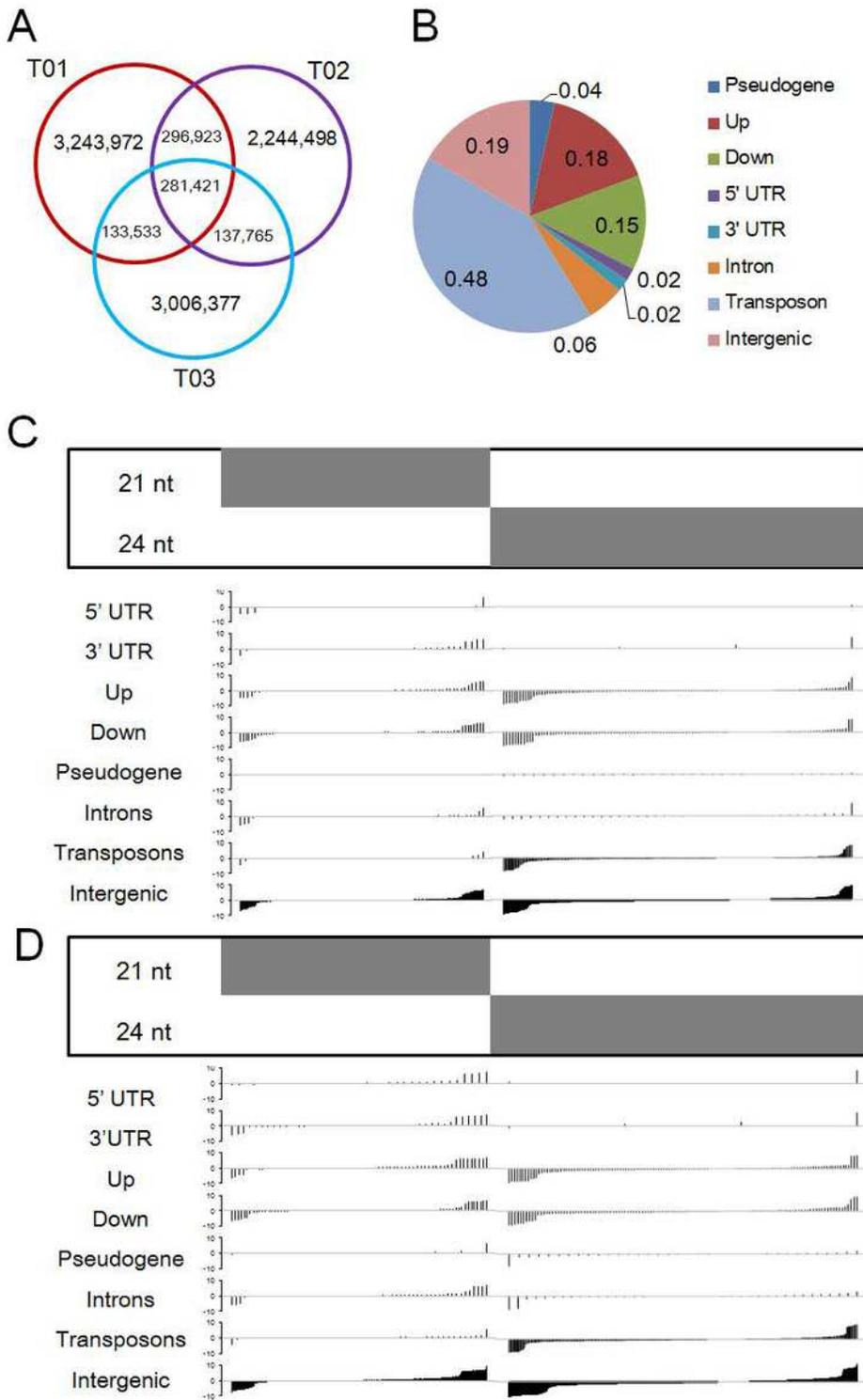


Figure 1

The stage differences in siRNA abundance primarily originate from transposons in response to infection. Shown are the 21 and 24-nt siRNA clusters with at least 5 tpm (transcripts per million mapped reads). Clusters are arranged in ascending order of fold change. (A) The common and stage-specific siRNA

clusters. (B) The relative abundance of the siRNA clusters in the various types of physical locations in the *Populus* genome. (C, D) Degree of stage-specific differences for the siRNA clusters in the different locations. The log₁₀ values of their abundance in T01 (The control group; 48 h) divided by their abundance in (C) T02 (biotrophic growth phase; 48 h post infection) or (D) T03 (urediniospore formation and release phase; 168 h post infection). Up indicate the 2 kb upstream region of genes and down indicates the 2 kb downstream region of genes. The 22-nt and 24-nt siRNA clusters with at least 5 tpm in the three samples are shown. The clusters below the horizontal line were negatively regulated, while those above the horizontal line were positively regulated. The clusters are arranged in ascending order of fold change.

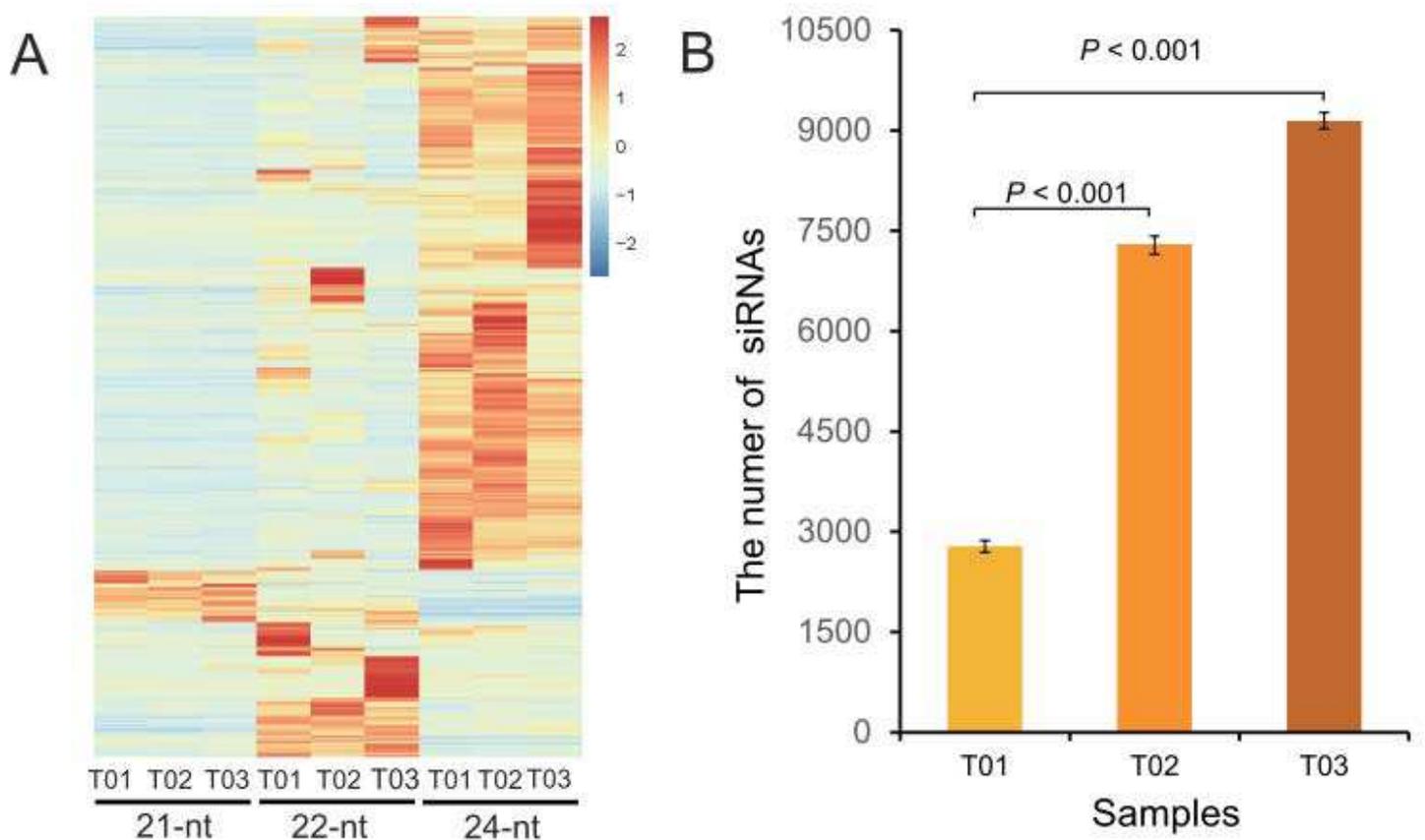


Figure 2

The difference in transposon siRNAs activity are driven by 24-nt sRNAs. (A) The abundance of siRNAs in the three samples (T01; T02; T03). The total abundance column displays the relative abundance of 21-, 22-, and 24-nt siRNAs. (B) The number of pathogen effectors targeted by the siRNAs in the three infection samples. The effectors were predicted by EffectorP and the target prediction were performed by using

This is a list of supplementary files associated with this preprint. Click to download.

- [DataS1.xlsx](#)
- [DataS2.xlsx](#)
- [DataS3.xlsx](#)
- [DataS4.xlsx](#)
- [DataS5.xlsx](#)
- [DataS6.xlsx](#)
- [DataS7.xlsx](#)
- [DataS8.xlsx](#)
- [DataS9.xlsx](#)
- [DataS10.xlsx](#)
- [DataS11.xlsx](#)
- [FigureS1.docx](#)
- [FigureS2.docx](#)
- [FigureS3.docx](#)
- [TableS1.docx](#)