Biochimie xxx (2014) 1–10

Contents lists available at ScienceDirect

Biochimie

journal homepage: www.elsevier.com/locate/biochi

Research paper

Dynamic expression of novel and conserved microRNAs and their targets in diploid and tetraploid of *Paulownia tomentosa*

Guoqiang Fan [a,*], Xiaoqiao Zhai [b], Suyan Niu [a], Yuanyuan Ren [b]

[a] Institute of Paulownia, Henan Agricultural University, No 95 Wenhua Road, 450002 Zhengzhou, People's Republic of China
[b] Forestry Academy of Henan, 450008 Zhengzhou, People's Republic of China

ARTICLE INFO

Article history:
Received 21 May 2013
Accepted 14 February 2014
Available online xxx

Keywords:
Paulownia tomentosa
Diploid
Tetraploid
MicroRNA
Targets

ABSTRACT

MicroRNAs (miRNAs) play profound roles in plant growth and development by regulating gene expression. Tetraploid plants often have better physical characteristics and stress tolerance than their diploid progenitors, but the role of miRNAs in this superiority is unclear. *Paulownia tomentosa*, (Paulowniaceae) is attracting research attention in China because of its rapid development, wide distribution, and potential economic uses. To identify miRNAs at the transcriptional level in *P. tomentosa*, Illumina sequencing was used to sequence the libraries of diploid and tetraploid plants. Sequence analysis identified 37 conserved miRNAs belonging to 14 miRNA families and 14 novel miRNAs belonging to seven miRNA families. Among the miRNAs, 16 conserved miRNAs from 11 families and five novel miRNAs were differentially expressed in the tetraploid and diploid; most were more strongly expressed in the former. The miRNA target genes and their functions were identified and discussed. The results showed that several *P. tomentosa* miRNAs may play important roles in the improved traits seen in tetraploids. This study provides a foundation for understanding the regulatory mechanisms of miRNAs in tetraploid trees.

1. Introduction

MicroRNAs (miRNAs) are a major type of endogenous 21–24 nucleotide (nt) single-stranded non-coding RNAs that are predominately derived from intergenic regions in both prokaryotes and eukaryotes [1–3]. They play important regulatory roles at the transcriptional and post-transcriptional levels during many growth and developmental processes, such as developmental timing, hormone responses, and reactions to environmental stress response [4–6]. MiRNAs were initially discovered in *Caenorhabditis elegans* [7].

In recent years, hundreds of small RNAs, especially miRNAs with low abundance, have been isolated in various higher plant species using next-generation sequencing technology and experimental and/or bioinformatical approaches, but most of these studies have focused on crop and annual plants; very few studies have involved forest trees [8–17]. For example, little information on miRNAs from *Paulownia* (Paulowniaceae) is available. *Paulownia tomentosa* is one of the most important indigenous fast-growing tree species in China, where it has a very wide distribution and is intercropped on 2.5 million ha of farmland. For effective breeding and improvement of *P. tomentosa*, autotetraploids were derived from diploid parent plants using colchicine [18]. Recently, the ecological characteristics and timber quality of the autotetraploid trees were found to be better than those of the corresponding diploids [19,20]. To clarify the underlying molecular mechanisms of prominent characteristics of the autotetraploid trees, we here adopted high-throughput sequencing technology to identify conserved and novel miRNAs of *P. tomentosa*. Differences in expression levels of these miRNAs between the diploids and autotetraploids were analyzed, and the potential roles of their target genes were investigated.

2. Materials and methods

2.1. Materials

Diploid (PT2) and autotetraploid (PT4) *P. tomentosa* were grown for 30 days in vitro on MS media containing 20 g L^{-1} sucrose, 1.0 mg L^{-1} NAA and 180 mg L^{-1}6-BA at 25 °C under a 16/8 h (light/dark) photoperiod, respectively. Leaves from plants were collected, frozen immediately in liquid nitrogen, and stored at −86 °C for total RNA extraction.

2.2. Small RNA library construction and sequencing

Total RNA was extracted from PT2 and PT4 leaves with Trizol reagent (Invitrogen, Carlsbad, CA, USA) following the

* Corresponding author. Tel.: +86 (0)37163558605.
E-mail address: zlxx64@126.com (G. Fan).

http://dx.doi.org/10.1016/j.biochi.2014.02.008

manufacturer's instructions. Two sRNA libraries were constructed and sequenced using the GAIIx platform. Briefly, 4 μg of total RNA were ligated with the 5′ and 3′ adapters successively. The reverse transcription reaction is used to create single stranded cDNA, and then amplified by 12 cycles PCR. The cDNA library were purified by polyacrylamide gel electrophoresis (PAGE) to select the fragments sized from 140 to 160 bp to produce the library for cluster generation and sequenced on the GAIIx platform were performed following the manufacturer's standard cBot and sequencing protocols.

2.3. Identification of miRNAs

The raw reads were produced with Illumina sequencing, and the low quality reads, adapters and contaminated reads were removed. The unique reads were then used to analyze the length distribution and mapped onto the *Paulownia* unigenes using miRDeep2. The perfectly matched reads were retained for further analysis. The reads matching non-coding RNA (including tRNA, rRNA, snoRNA, and other ncRNA, except microRNA) in non-coding RNA database [21] (Release 10) were deleted.

The remaining reads were searched against the plant mature microRNA of Sanger miRBase (Release 19.0) to identify the conserved miRNAs using the program Blastall, allowing two mismatches. The potential novel miRNAs were identified by using MIREAP and RNAfold [22] to fold flanking sequences and predict secondary structures. If the sRNA had a perfect stem loop structure and followed the other criteria described by Meyers et al., it was considered to be a novel miRNA [23].

2.4. Differential expression analysis of miRNAs in the PT2 and PT4

In the miRNA expression analysis, the abundance of miRNAs in the two libraries was normalized to one million, regardless of the total number of miRNAs in each sample. The fold change between the PT4 and PT2 was calculated as follows:

Fold change = miRNA normalized read counts in PT4 library/miRNA normalized read counts in PT2 library.

The *P*-value was obtained according to the calculations as follows:

$$P(x|y) = \left(\frac{N_2}{N_1}\right) \frac{(x+y)!}{x!y!\left(1+\frac{N_2}{N_1}\right)^{(x+y+1)}}$$

$$C(y \le y\min|x) = \sum_{y=0}^{y\le y\min} p(y|x)$$

$$D(y \ge y\max|x) = \sum_{y\ge y\max}^{\infty} p(y|x)$$

$$\log 2 \text{ ratio} = \log 2(\text{fold change})$$

For direct comparison, *Paulownia* plantlets used to construct the libraries were grown under similar conditions.

2.5. Identification of miRNAs targets by degradome sequencing

To dissect miRNA-guided genes regulation in the PT2 and PT4 plants, two degradome libraries suitable for miRNA target identification were constructed following a previously described protocol [24,25]. In brief, poly(A) RNA was isolated and ligated to a 5′RNA adapter containing a MmeI recognition site. After reverse transcription using oligod(T) and PCR enrichment, the PCR products were purified and digested with *MmeI* (NEB, Ipswich, MA, USA). A double-stranded DNA adapter was then ligated to the digested products using T4 DNA ligase (NEB, Ipswich, MA, USA). The products were amplified using 20 PCR cycles and the final cDNA library was purified and sequenced on Illumina HiSeqTM 2000 system.

After initial processing, the unique sequence signatures were mapped to the database of *P. tomentosa* transcriptome sequences using SOAP software (http://soap.genomics.org.cn/) to define the coverage rate. The perfect matching sequences were retained and extend to 31 nt by adding approximately 15 nt of upstream of the sequence. All resulting reads (t-signature) were reverse-complemented and aligned to the miRNA identified in this study. Alignments with scores not exceeding 4 and having the 5′ end of the degradome sequence coincident with the tenth and eleventh nucleotides of complementarity to the sRNA were considered potential targets. Furthermore, t-plots were built according to the distribution of signatures (and abundances) along these transcripts. To better understand the functions of these targets, the Iprscan (http://www.ebi.ac.uk/tools/pfa/iprscan/) program was employed to gain the GO annotations and the pathway from the unigene database through Blastall hits against the available Pfam database and the Kyoto Encyclopedia of Genes and Genomes (KEGG) database in NCBI by an *E*-value threshold of less than 10^{-5}. The GO categorization results were expressed as three independent hierarchies for biological process, cellular component, and molecular function [26].

2.6. Quantitative real-time PCR

Identified *P. tomentosa* miRNAs and their target genes were experimentally validated using quantitative real-time PCR (qRT-PCR). For the experiment samples, diploid and tetraploid *P. tomentosa* tissue culture plants grown for 30 days, six months and one year were used. RNA from two biological replications were used for qRT-PCR, and total RNA was isolated using plant RNA extraction KIT (Aidlab Biotechnologies Co.,Ltd., Beijing, China). The stem-loop primers were designed for qRT-PCR as described previously [27]. The forward primers were designed based on the mature miRNA sequences and the reverse primer was the universal reverse primer, with U6 as the endogenous reference. The primers for target genes were designed with Beacon Designer, version 7.7 (Premier Biosoft International, Ltd., Palo Alto, CA, USA), and the 18S rRNA of Paulownia was chosen as an endogenous reference gene for normalization. All reactions were run in triplicate for each sample. The sequences of the primers are listed in Table S1. SuperScriptIII platinum SYBR Green one-step qRT-PCR kit (Invitrogen, Carlsbad, CA, USA) and a CFX96 real time PCR system (Bio-Rad) were employed to detect and compare the expression levels. For each reaction, 500 ng of total RNA was mixed with 10 mL of SuperScriptIII platinum SYBR green PCR master mix and 4 pmol each of the reverse transcription, the forward and reverse primers in a final volume of 20 mL. The conditions for PCR amplification were as follows: 40 cycles at 95 °C for 15 s and 55 °C for 30 s. The $2^{-\Delta\Delta CT}$ relative quantization method was used to analyze relative changes in gene expression during the qRT-PCR experiments.

3. Results

3.1. Statistical analysis of sRNAs

The two sRNA libraries generated a total of 14,520,461 (PT2) and 13,109,201 (PT4) reads, respectively, by Illumina sequencing. After discarding the low-quality tags, adaptors, contaminants, sequences shorter than 18 nt, and sequences with poly-A tails, there remained

Please cite this article in press as: G. Fan, et al., Dynamic expression of novel and conserved microRNAs and their targets in diploid and tetraploid of *Paulownia tomentosa*, Biochimie (2014), http://dx.doi.org/10.1016/j.biochi.2014.02.008