Gene 666 (2018) 72–82

Contents lists available at ScienceDirect

Gene

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

Research paper

# Identification of microRNAs and long non-coding RNAs involved in fatty acid biosynthesis in tree peony seeds

Dan-Dan Yin[a,b], Shan-Shan Li[a], Qing-Yan Shu[a], Zhao-Yu Gu[a], Qian Wu[a,b], Cheng-Yong Feng[a,b], Wen-Zhong Xu[a,*], Liang-Sheng Wang[a,b,**]

[a] Key Laboratory of Plant Resources and Beijing Botanical Garden, Institute of Botany, Chinese Academy of Sciences, Beijing, 100093, China
[b] University of Chinese Academy of Sciences, Beijing 100049, China

ARTICLE INFO

*Keywords:*
miRNA
lncRNA
α-Linolenic acid
Target gene
Tree peony seed
*Paeonia*

ABSTRACT

MicroRNAs (miRNAs) and long noncoding RNAs (lncRNAs) act as important molecular regulators in a wide range of biological processes during plant development and seed formation, including oil production. Tree peony seeds contain > 90% unsaturated fatty acids (UFAs) and high proportions of α-linolenic acid (ALA, > 40%). To dissect the non-coding RNAs (ncRNAs) pathway involved in fatty acids synthesis in tree peony seeds, we construct six small RNA libraries and six transcriptome libraries from developing seeds of two cultivars (J and S) containing different content of fatty acid compositions. After deep sequencing the RNA libraries, the ncRNA expression profiles of tree peony seeds in two cultivars were systematically and comparatively analyzed. A total of 318 known and 153 new miRNAs and 22,430 lncRNAs were identified, among which 106 conserved and 9 novel miRNAs and 2785 lncRNAs were differentially expressed between the two cultivars. In addition, potential target genes of the microRNA and lncRNAs were also predicted and annotated. Among them, 9 miRNAs and 39 lncRNAs were predicted to target lipid related genes. Results showed that all of miR414, miR156b, miR2673b, miR7826, novel-m0027-5p, TR24651|c0_g1, TR24544|c0_g15, and TR27305|c0_g1 were up-regulated and expressed at a higher level in high-ALA cultivar J when compared to low-ALA cultivar S, suggesting that these ncRNAs and target genes are possibly involved in different fatty acid synthesis and lipid metabolism through post-transcriptional regulation. These results provide a better understanding of the roles of ncRNAs during fatty acid biosynthesis and metabolism in tree peony seeds.

## 1. Introduction

Tree peony (*Paeonia* section *Moutan*) is valued for medical and aesthetic uses, and is thus one of the most important horticultural crops in China (Li et al., 2012; Li et al., 2017). Recently, it has been considered a newly developing woody oil crop to produce edible oils containing substantial amounts of unsaturated fatty acids (UFAs) in seeds, such as oleic acid (OA, 18:1$\Delta^{9}$), linoleic acid (LA, 18:2$\Delta^{9,12}$), and α-linolenic acid (ALA, 18:3$\Delta^{9,12,15}$). Tree peony seeds contain abundant UFAs (> 90%) and high proportions of ALA, which accounts for 45% (Li et al., 2015b; Han et al., 2016). ALA is an essential FA that cannot be biosynthesized in mammals, and the recommended ratio of omega-6 relative to omega-3 (n-6:n-3) FA in the human diet is approximately 2:1 to 6:1 (Simopoulos, 2000; Wijendran and Hayes, 2004). Several studies have strongly indicated that ALA not only reduces cholesterol when used as a supplement, it also plays an important role in preventing and curing coronary heart disease and hypertension during pregnancy and lactation (De Lorgeril and Salen, 2004; Ramaprasad et al., 2006). However, traditional oilseed crops (e.g., soybean, peanut, maize, sunflower, and rapeseed) have a low ALA content (0–10%) with a high n-6:n-3 ratio (15:1 to 20:1) (Lee et al., 1998; Erdemoglu et al., 2003; Wang et al., 2011; Kim et al., 2015). The high n-6:n-3 ratio in our modern diet has been thought to be a major contributor to lipid metabolism disorder, cardiovascular disease and autoimmune diseases

*Abbreviations:* ncRNA, non-coding RNA; ALA, α-linolenic acid; FA, fatty acid; LA, linoleic acid; OA, oleic acid; KEGG, Kyoto Encyclopedia of Genes and Genomes; GO, gene ontology; rRNA, ribosomal RNA; snoRNA, small nucleolar RNA; snRNA, small nuclear RNA; tRNA, transfer RNA; qRT-PCR, quantitative real-time PCR; Q30 percentage, percentage of bases with sequencing error rate lower than 1‰; ACCase, acetyl-CoA carboxylase; ACP, acyl carrier protein; HAD, hydroxyacyl-ACP dehydrase; SPL, squamosa promoter binding protein-like; CCT, choline-phosphate cytidylyltransferase; PDHC, pyruvate dehydrogenase; PLC, phospholipase C; LPAAT, lysophosphatidic acid acyltransferase; PHS, very long chain (*3R*) 3-hydroxyacyl-CoA dehydratase; AP2-EREBP, APETALA2/ethylene-responsive element binding protein; MCAAT, malonyl-CoA: ACP transacylase; KAS, β-ketoacyl-ACP synthase; KAR, β-ketoacyl-ACP reductase; EAR, enoyl-ACP reductase; TAG, triacylglycerol

* Corresponding author.
** Correspondence to: L.-S. Wang, University of Chinese Academy of Sciences, Beijing 100049, China.
*E-mail addresses:* xuwzh@ibcas.ac.cn (W.-Z. Xu), wanglsh@ibcas.ac.cn (L.-S. Wang).

https://doi.org/10.1016/j.gene.2018.05.011
Received 2 February 2018; Received in revised form 20 April 2018; Accepted 2 May 2018
Available online 05 May 2018
0378-1119/ 

(Simopoulos, 2002; Simopoulos, 2006).

As the benefits of tree peony seed oil for human health have become clear, tree peony has been increasingly cultivated on a massive scale for the production of oils. However, the oil compositions and fatty acids content in tree peony seed vary significantly among cultivars or when growing under different environments (Li et al., 2015a; Li et al., 2015b; Han et al., 2016). Therefore, it is necessary and interesting to discover the underlying molecular mechanisms on the biosynthesis of abundant unsaturated fatty acids with high proportions of ALA in tree peony seeds. In our previous study (Li et al., 2015b), we found that in most cultivars, ALA was the dominant compound, and LA and OA were subdominant, with some exceptions such as in "Saiguifei" (Cultivar S). Cultivar S is unique because it has higher content of LA than ALA and the ALA content was only 26.1%. In addition, "Jingshenhuanfa" (Cultivar J) has been screened for oil production with high yield and high quality. The ALA content in Cultivar J could reach 53.5%. Therefore, these two cultivars are ideal materials used for elucidating the mechanism of fatty acid and ALA synthesis.

In oilseeds, pyruvate is the primary precursor for FA synthesis in the plastid. Firstly, the pyruvate dehydrogenase complex (PDHC) is used to catalyze the oxidative decarboxylation of pyruvate to produce acetyl-CoA, which is in turn used by acetyl-CoA carboxylase (ACCase) to form malonyl-CoA. ACCase has long been considered as a rate-limiting step for the FA biosynthetic process (Baud, 2018). Before entering the process, the malonyl group is transferred from CoA to a protein cofactor named acyl carrier protein (ACP) by a malonyl-CoA: ACP transacylase (MCAAT). Production of saturated FAs is performed in a stepwise manner by the fatty acid synthase (FAS) complex, which includes a series of condensation reactions catalyzed by β-ketoacyl-ACP synthase (KAS), β-ketoacyl-ACP reductase (KAR), hydroxyacyl-ACP dehydrase (HAD) and enoyl-ACP reductase (EAR). Then in the first step of desaturation in plastid, saturated fatty acid (18:0-ACP) is efficiently catalyzed by a stromal Δ9 stearoyl-ACP desaturase (SAD), so that OA is preferentially exported to the endoplasmic reticulum (ER). The main pathway leading to the production of LA and ALA in oilseeds starts in the ER with the incorporation of oleoyl-CoA into membrane phosphatidylcholine (PC) where it is desaturated to LA by the Δ12 FA desaturase (FAD2) (Kang et al., 2011; Zeng et al., 2017). Further desaturation of LA to ALA is accomplished by the Δ15 FA desaturase (FAD3) (O'Neill et al., 2011).

A non-coding RNA (ncRNA) is a class of small RNA molecules with specific functions that can be transcribed but do not code for proteins. A well-characterized class of ncRNAs is microRNAs (miRNAs), approximately 22 nucleotides (nt) small ncRNAs that are processed from larger precursors. MicroRNA database (miRBase 21.0) contains 28,645 hairpin precursor miRNAs (pre-miRNAs) from 223 species. A total of 35,828 mature miRNA products are expressed by these pre-miRNAs (Kozomara and Griffiths-Jones, 2014). Most of these miRNAs play important roles in a wide range of physiological and metabolic processes, including growth, development, and responses to various stresses in plants (Tang et al., 2017; Tang and Chu, 2017; Yang et al., 2017). miRNAs are involved at the post-transcriptional level by targeting mRNA degradation and repressing translation (Bartel, 2004; Gregory and Shiekhattar, 2005). The rapid development of next-generation sequencing technology and bioinformatics tools provides a broad platform for the analysis of miRNAs in more and more plant species.

Recently, another class of ncRNAs, the long noncoding RNAs (lncRNAs) have gained increased attention (Mercer et al., 2009; Hung and Chang, 2010). LncRNAs, longer than 200 nucleotides, have been found to play a key role in transcriptional, post transcriptional and epigenetic gene regulation in numerous organisms. However, the regulation mechanism and function of lncRNAs is still poorly understood in plants, especially in woody plants. What' more, lncRNAs may play widespread roles in many biological processes (Ulitsky and Bartel, 2013; Komiya et al., 2014; Zhang et al., 2014), but the regulation of lncRNAs related to lipid and ALA accumulation is poorly discovered.

Above all, it is important to understand the molecular mechanism underlying fatty acid synthesis, especially ALA biosynthesis, between different tree peony. In this study, we performed high-throughput sequencing technologies for extracting ncRNAs (microRNAs and lncRNAs) and analyzed the profiles of ncRNA differential expression in tree peony seeds between high-ALA-content and low-ALA-content cultivars. Conserved and novel miRNAs were identified, target genes for miRNAs were predicted, and expression patterns of selected miRNAs and their potential targets were examined by qRT-PCR. In addition, lncRNAs and their potential target genes were identified and characterized.

## 2. Materials and methods

### 2.1. Plant materials, RNA extraction, and small RNA sequencing

Two tree peony cultivars were grown in Beijing Botanical Garden, Institute of Botany, Chinese Academy of Sciences (Lat. 39°48′ N, 116°28′ E, Alt. 76 m), China. Seeds of both cultivars were collected at 14, 35, 56, 70 and 91 days after pollination separately. At each time point, seeds from three trees were collected, flash frozen in liquid nitrogen and stored at −80 °C until further use. The seeds collected from five stages of each cultivar were then equally mixed to prepare for sequencing. Total RNA was isolated using a TIANGEN RNA Prep Pure Plant kit (Tiangen Biotech Co. Ltd., Beijing, China) according to the manufacturer's instructions, and the RNA quality was evaluated using a NanoDrop ®ND-1000 spectrophotometer. Small RNA sequences of Cultivar S and J seeds were generated by Illumina HiSeq® X-Ten paired-end sequencing system with three replicates for each cultivar. The library preparation and sequencing were constructed by Ori-Gene Science and Technology Corporation (Beijing, China).

### 2.2. Small-RNA data analysis

After removing low-quality sequences (reads with 5′ end primer contaminants, reads without 3′ end primer, reads without the insert tag, reads with poly A and reads shorter than 15 nt), the clean reads were mapped to the Rfam (http://rfam.sanger.ac.uk) to remove ribosomal RNA (rRNA), transfer RNA (tRNA), small nuclear RNA (snRNA), and small nucleolar RNA (snoRNA). Merged unique reads were also screened against the miRNA database (miRBase 21.0) using a nucleotide-nucleotide Basic Local Alignment Search Tool (BLASTn) to identify the conserved miRNAs. In order to identify the novel miRNAs, all candidate precursors with hairpin-like structures were obtained using the Mireap program (http://sourceforge.net/projects/mireap). Potential targets for the miRNAs were predicted using the TargetFinder software. After that we annotated the predicted target transcripts with gene ontology (GO) terms. Expression levels of miRNAs were calculated according to the value of Reads Per Million (RPM). Differentially expressed miRNAs were defined based on $p$ value $< 0.05$ and differential expression fold $> 2$ or $< 0.5$.

### 2.3. qRT-PCR validation

An ABI Step One Plus™ Real-Time PCR System (Applied Bio-systems, USA) was used for the qRT-PCR of miRNAs and their targets. In a miRNA qRT-PCR experiment, 1 μg of sRNA (TIANGEN miRcute miRNA extraction and separation kit) was used in each reaction with U6 as the reference gene; miRcute miRNA First-Strand cDNA synthesis Kit (TIANGEN) was used to generate single-stranded miRNA cDNA by reverse transcription; the expression levels of miRNAs were analyzed using qRT-PCR (SYBR Green Premix, TIANGEN). In a mRNA qRT-PCR experiment, 1 μg of total RNAs (TIANGEN RNA Prep Pure Plant kit) was translated to cDNA using TIANGEN FastQuant RT Kit, with *ubiquitin* gene (Wang et al., 2012) as the internal control; predicted target genes were validated by qRT-PCR using specific primers designed with the