Gene 536 (2014) 151–162

Contents lists available at ScienceDirect

Gene

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

Computational identification of microRNAs in the strawberry (*Fragaria* × *ananassa*) genome sequence and validation of their precise sequences by miR-RACE

Jian Han [a,1], Aying Li [a,1], Hong Liu [b], Xicheng Wen [a], Mizhen Zhao [b], Nadira Bilkish Korir [a], Nicholas Kibet Korir [c], Chen Wang [a], Jinggui Fang [a,*]

[a] College of Horticulture, Nanjing Agricultural University, Nanjing 210095, China
[b] Department of Horticulture, Nanjing Jinling Institute of Technology, Nanjing 210038, China
[c] Department of Agricultural Technology, Kenyatta University, P.O. Box, 43844-00100 Nairobi, Kenya

ARTICLE INFO

Article history:
Accepted 12 November 2013
Available online 11 December 2013

Keyword:
Strawberry
Bioinformatics
MicroRNAs
miRNA RACE
Genome database

ABSTRACT

In plants, microRNAs (miRNAs) play significant roles in post-transcriptional gene regulation and have been found to control many genes involved in different biological and metabolic processes. Extensive studies were carried out to discover miRNAs and analyze their functions in model plant species, such as in *Arabidopsis* and rice that have been reported. In this research, we used bioinformatics to predict microRNAs in an important strawberry rootstock cultivar to discover and validate precise sequences of microRNAs in strawberry. By adopting a range of filtering criteria, we obtained 59 potential miRNAs belonging to 40 miRNA families from the *Fragaria vesca* genome. Using two specific 5′ and 3′ miRNA RACE PCR reactions and a sequence-directed cloning method, we accurately determined 34 precise sequences of candidate miRNAs, while six other sequences exhibited some minor divergence in their termini nucleotides, and 19 miRNAs that could not be cloned owing to expression abundance may be too low or these mirRNAs predicted could not be existing in strawberry. Potential target genes were further predicted for the miRNAs above. The expression of the 16 miRNAs unreported and having exact sequences and their targets by experiment could be detected in different tissues of strawberry ranging from roots, stems, leaves, flowers and fruits by qRT-PCR and some of them showed differential expression in various tissues. The functional analysis of 16 miRNAs and their targets was carried out. Finally, we conclude that there are 34 mirRNAs in strawberry and their targets play vital roles not only in growth and development, but also in diverse physiological processes. These results show that regulatory miRNAs exist in agronomically important strawberry and might have an important function in strawberry growth and development.

1. Introduction

MicroRNAs are a series of non-coding RNAs that regulate gene expression at the post-transcriptional level, either by endonucleolytic cleavage or by translational inhibition, and they also play important roles in gene silencing (Chen, 2004; Lauter et al., 2005; Llave et al., 2002). In plants, miRNAs are transcribed by RNA polymerase II into long primary transcripts (pri-miRNAs) (B. Zhang et al., 2006; Chen et al., 2005), and they are cut into miRNA precursors (pre-miRNAs) with typical hairpin structure(s). Mature miRNAs are generated from the stem portion of single stranded stem–loop precursors by complexes containing the nuclear RNase III enzyme and the ribonuclease III-like enzyme Dicer (DCL1) (Kurihara and Watanabe, 2004), then the mature miRNAs are incorporated into the RNA-induced silencing complex (RISC) and guide RISC to complementary mRNA targets. Finally, the RISC inhibits translation elongation or triggers the degradation of target mRNAs (Lin et al., 2005). In recent years, the discovery of many small RNAs has aroused a great deal of interest in post-transcriptional regulation of gene expression during development and other biological processes. Increasing evidence indicates that miRNAs play major roles in key aspects of plant development and their responses to environmental stresses (Allen et al., 2005; Berezikov et al., 2006; Carthew and Sontheimer, 2009; Chen et al., 2005). The fact that a large number of known miRNAs in the plant kingdom ranging from mosses and ferns to higher flowering plants are vastly evolutionarily conserved has been used as a practical indicator for prediction of miRNAs by homology searches in other species (B. Zhang et al., 2006; Sunkar et al., 2008) and provides powerful evidence supporting prediction and validation of conserved miRNAs and their target genes from all plant species by bioinformatics and experimental methods (Axtell and Bartel, 2005; Fahlgren et al., 2007).

Abbreviations: miRNAs, microRNAs; qRT-PCR, quantitative real-time PCR; DCL1, the ribonuclease III-like enzyme Dicer; RISC, the RNA-induced silencing complex; UPE, unpaired energy; LMW, low molecular weight; EST, expressed sequence tag; Fve-mirRNA, *Fragaria vesca* microRNA; CDS, coding sequence.

* Corresponding author.
E-mail address: fanggg@njau.edu.cn (J. Fang).

[1] These authors contributed equally to this study.

0378-1119/$ – see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.gene.2013.11.023

With the advance of bioinformatics, many computational methods have been employed for the prediction and identification of miRNAs (BH. Zhang et al., 2006; Rhoades and Bartel, 2004; Sunkar and Zhu, 2004; Yin et al., 2008). A great number of plant miRNAs predicted computationally have been deposited in miRBase (http://www.mirbase.org/) and they included *Arabidopsis thaliana* (190), *Populus trichocarpa* (234), *Vitis vinifera* (137), *Oryza sativa* (414) and *Zea mays* (109). Theoretically, computational prediction of miRNAs can lead to discovery of all the potential miRNAs based on the specific criteria for characterization of miRNAs. However, due to the evolution of miRNAs and the non-availability of miRNA end characteristics that exactly define both ends of mature miRNAs and false predictions are unavoidable to some extent. This is one of the major disadvantages of computational prediction compared to direct cloning of miRNAs. Therefore, employment of experimental methods to verify the precise sequences of miRNA is indispensable. In earlier studies, the experimental methods used for validating expression of miRNAs predicted by bioinformatics were mainly focused on Northern blotting and/or RT-PCR techniques (Song et al., 2009). However, Northern blotting does not reveal the actual miRNA sequences and is not sensitive enough to detect miRNAs with low abundance, on the other hand PCR-based amplification is practically difficult when the actual mature miRNA region is not known. Northern blotting and PCR-based amplification can only confirm the sizes and existence of miRNAs, but could not determine the full precise sequence of the computer-predicted miRNAs. MiR-RACE (Song et al., 2010) can overcome the shortcomings mentioned above, as it can verify the changes of 3–4 nucleotide sequences at 5′ and 3′ ends of miRNAs. This makes miR-RACE technology a suitable option for validation of precise sequences, particularly at both ends, of computationally predicted miRNAs in an organism, which were exemplified by successful utilization in previous studies on identification of grape and apple miRNA precise sequences (Song et al., 2010; Yu et al., 2010).

Strawberry is one of the most economically and nutritionally important fruit crop in the world and is used as a model species for molecular biology and genomic studies since the genome of woodland strawberry was obtained (Shulaev et al., 2011). Although miRNAs have been extensively studied in the last five years, limited systematic studies of miRNAs have been performed in the genus *Fragaria*. The identification of sets of small RNAs in any organism has fundamental importance in understanding diversity of small RNAs and their functions. In this research, we combined bioinformatics prediction and miR-RACE technology to identify miRNAs in strawberry. Findings of this study make a substantial contribution and provide a foundation for further research on miRNAs particularly in strawberry.

2. Materials and methods

2.1. Plant material

Research materials were collected from the different parts of the plant such as roots, stems, leaves, flowers as well as fruits at 5 different developmental stages (young green, big green, white, pink and mature red) from strawberry 'Sweet Charlie' were collected from the Beijing University of Agriculture, China in 2013. After collection, all the samples were immediately frozen in liquid Nitrogen and stored at −80 °C until use for RNA extraction.

2.2. Analysis of miRNAs by 5′miR-RACE and 3′miR-RACE

The cDNA was amplified with the mirRacer 5′ primer (5′-GGACACTG ACATG GACTGAAGGAGTA-3′) and the mirRacer 3′ primer (5′-ATTCTA GAGGCCGAGGCGGCCGAC ATG-3′) to generate a pool of non-gene-specific product. 5′ miR-RACE reactions were performed with the mirRacer 5′ primer and miRNA-gene-specific forward primers (GSP1) (Table 1), and 3′ miR-RACE reactions were carried out with the mirRacer 3′ primer and miRNA-gene-specific reverse primers (GSP2) (Table 1) (Song et al., 2010). The PCR cycling conditions for amplification were carried out for 25 cycles at a final annealing temperature of 60 °C. In each case, unique gene-specific DNA fragments were amplified. After the amplification, the 5′and 3′ PCR products were separated in a 2.5% agarose gel with ethidium bromide (EtBr) staining. The gel slices containing DNA with a size of about 56 bp (5′-end product) and 87 bp (3′-end product) were excised and the DNA fragments were purified using an agarose gel DNA purification kit (Takara, Otsu, Japan), following the manufacturer's instructions. The DNA fragment was directly sub-cloned with the TOPO TA cloning Kit (Invitrogen) and Colony PCR was performed using the PCR-specific primer pairs just as mentioned above. The 5′ RACE and 3′ RACE clones having PCR products of about 56 bp and 87 bp, respectively, were sequenced (Invitrogen).

2.3. Prediction of potential target genes for strawberry miRNAs

Since plant miRNAs have a high degree of sequence complementarity with their target genes, an assortment of software used for predicting the target genes of plant miRNA was developed. These tools include MIRU, miRNAassist, PatScan and psRNATarget. Prediction of target genes was performed based on methods described by Zhang et al. (2011), with psRNATarget software being incorporated. The maximum expected value was set at three while the other parameters were at default; psRNATarget screens target genes through a penalty point mechanism scoring three points per 20 nt, each swing of G:U pairing, deleting or inserting, as well as other mismatching penalties of 0.5, two and one point respectively (because sequences in the 5′ends of miRNAs are very important for the complementary sequences of the target genes, 2 to 7 nt at the 5′end of miRNA base having other than G:U a mismatch penalty one point is awarded), sequences with a total score of less than three and UPE less than 25.0 kcal/mol are potential target genes of miRNA.

2.4. Oligonucleotide synthesis and preparation

All the oligonucleotides used were purchased from Invitrogen Technologies, and were purified by desalting. The primers used are as listed in Table 1.

2.5. Low molecular weight RNA extraction

Using the CTAB method (Chang et al., 1993), total RNA was isolated from the strawberry tissues and the low and larger molecular weight RNA separated by 10 M LiCl following procedures reported earlier (Song et al., 2009). The low molecular weight (LMW) RNA fraction was dissolved in 30 μl of RNase free water and the concentration of the RNA samples was measured by a UV-1800 spectrophotometer (Shimadzu, Japan) and visually ascertained in a 2.5% agarose gel.

2.6. Construction and screening of small RNA cDNA libraries

We generated a miRNA-enriched library and used it to clone miRNAs as well as measure the expression of miRNAs via RT-PCR (Elbashir et al., 2001; Lewis et al., 2005). For this 5′- and 3′-end adaptors were linked to the miRNA molecules (Song et al., 2010), which were then further reverse transcribed using Superscript III reverse transcriptase (Invitrogen) in the presence of random nonamers (Sigma), according to manufacturer protocols. After preparation of the miRNA libraries from the different tissues, we pooled uniform quantities of these library samples for further PCR amplification reactions.

2.7. Bioinformatic prediction of miRNAs potential in strawberry

A flowchart illustrated the search for potential miRNAs as shown in Fig. 1. The Patscan program was used to identify matches to known miRNAs in the target genomic sequence with 0–2 substitutions and no