Genomics xxx (2015) xxx–xxx

Contents lists available at ScienceDirect

Genomics

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

Profiling microRNA expression during multi-staged date palm (*Phoenix dactylifera* L.) fruit development

Chengqi Xin [a,b,c,1], Wanfei Liu [a,b,1], Qiang Lin [a,b,1], Xiaowei Zhang [a,b], Peng Cui [d], Fusen Li [a,b], Guangyu Zhang [a,b], Linlin Pan [a,b], Ali Al-Amer [a], Hailiang Mei [a,b,c], Ibrahim S. Al-Mssallem [a,e,*], Songnian Hu [a,b,**], Hasan Awad Al-Johi [a,***], Jun Yu [a,b,****]

[a] Joint Center for Genomics Research (JCGR), King Abdulaziz City for Science and Technology and Chinese Academy of Sciences, Prince Turki Road, Riyadh 11442, Saudi Arabia
[b] CAS Key Laboratory of Genome Sciences and Information, Beijing Institute of Genomics, Chinese Academy of Sciences, No. 1 Beichen West Road, Chaoyang District, Beijing 100101, China
[c] University of Chinese Academy of Sciences, No. 19A Yuquan Road, Shijingshan District, Beijing 100049, China
[d] Biological and Environmental Sciences and Engineering Division, King Abdullah University of Sciences and Technology, Thuwal 23955-6900, Saudi Arabia
[e] Department of Biotechnology, College of Agriculture and Food Sciences, King Faisal University, Hofuf, Saudi Arabia

ARTICLE INFO

Article history:
Received 7 November 2014
Accepted 20 January 2015
Available online xxxx

Keywords:
Phoenix dactylifera
MiRNA
Fruit development
MiRNA target
Gene regulation

ABSTRACT

MicroRNAs (miRNAs) play crucial roles in multiple stages of plant development and regulate gene expression at posttranscriptional and translational levels. In this study, we first identified 238 conserved miRNAs in date palm (*Phoenix dactylifera*) based on a high-quality genome assembly and defined 78 fruit-development-associated (FDA) miRNAs, whose expression profiles are variable at different fruit development stages. Using experimental data, we subsequently detected 276 novel *P. dactylifera*-specific FDA miRNAs and predicted their targets. We also revealed that FDA miRNAs function mainly in regulating genes involved in starch/sucrose metabolisms and other carbon metabolic pathways; among them, 221 FDA miRNAs exhibit negative correlation with their corresponding targets, which suggests their direct regulatory roles on mRNA targets. Our data define a comprehensive set of conserved and novel FDA miRNAs along with their expression profiles, which provide a basis for further experimentation in assigning discrete functions of these miRNAs in *P. dactylifera* fruit development.

1. Introduction

Date palm (*Phoenix dactylifera* L.), as a socio-cultural symbol for the Arabian Peninsula (Saudi Arabia and Gulf countries), has a long agriculture history [1] and is also an economically important food crop in tropical and subtropical regions. Its major edible part is its ripe fruit or the date that contains a large amount of monosaccharides and has high nutritional value [2]. The *P. dactylifera* date development and ripening is a complex process, including a long period from the initial pollination for fruit set to the terminal sugar-rich stage, during which its major metabolic mechanisms undergo a complex series of physiological events, such as carbon fixation, starch hydrolysis and synthesis, and glucose and fructose accumulation. The starch accumulated in the early stages eventually converts into easily-absorbed small sugar molecules, mostly monosaccharides, at the late stages, which include fructose, sucrose, and glucose [3]. It has been reported that several key enzyme-encoding genes, such as pyruvate kinase, glucosidase, synthase/phosphate synthase, and trehalose 6-phosphate synthase/phosphatase, play crucial roles in the *P. dactylifera* fruit development and ripening [3–5]. However, upstream regulators of these key genes, such as miRNAs, have yet to be studied in the species and in details.

MicroRNAs (miRNAs) are a class of endogenous small non-coding RNAs (20 to 24 nt in size), which are derived from single-stranded precursors that form a stem–loop secondary structure and common among animals, plants, and even viruses [6,7]. MiRNAs regulate gene expression by binding to their complementary sequences, leading to either cleavage-induced degradation or translational repression of their target transcripts [8]. MiRNAs have been studied in plants since 2002 and known to play important regulatory roles in plant hormone

* Correspondence to: I. Al-Mssallem, Joint Center for Genomics Research (JCGR), King Abdulaziz City for Science and Technology and Chinese Academy of Sciences, Prince Turki Road, Riyadh 11442, Saudi Arabia. Fax: +966 0135808411.
** Correspondence to: S. Hu, CAS Key Laboratory of Genome Sciences and Information, Beijing Institute of Genomics, Chinese Academy of Sciences, No. 1 Beichen West Road, Chaoyang District, Beijing 100101, China. Fax: +86 10 84097458.
*** Correspondence to: H. Al-Johi, Joint Center for Genomics Research (JCGR), King Abdulaziz City for Science and Technology and Chinese Academy of Sciences, Prince Turki Road, Riyadh 11442, Saudi Arabia. Fax: +966 4813787.
**** Correspondence to: J. Yu, CAS Key Laboratory of Genome Sciences and Information, Beijing Institute of Genomics, Chinese Academy of Sciences, No. 1 Beichen West Road, Chaoyang District, Beijing 100101, China. Fax: +86 10 84097540.
E-mail addresses: imssallem@gmail.com (I.S. Al-Mssallem), husn@big.ac.cn (S. Hu), haljohi@kacst.edu.sa (H.A. Al-Johi), junyu@big.ac.cn (J. Yu).
[1] These authors contributed equally to this work.

http://dx.doi.org/10.1016/j.ygeno.2015.01.004
0888-7543/

homeostasis [9], stress responses [10], and diverse developmental processes, including seed development, meristem and lateral organ development, root initiation, flowering and sex determination, fruit development, timing and phase transitions [11,12]. Since the first discovery of miRNAs from *Arabidopsis* in 2002 [13], plant miRNAs from various species have been intensely studied through both experimental and computational approaches. A recent update has pointed out that there have been 21,264 miRNAs discovered from 193 species and 5943 of them are plant in origin (miRBase, Release 19.0, August 2012) [14]. These plant-specific miRNAs include 713 from *Arabidopsis*, 721 from *Oryza*, 374 from *Populus*, and 719 from *Medicago truncatula*, and the rest are mostly from a dozen or so other plant species.

The *P. dactylifera* (*Pd*) genome has been sequenced [4,15], and its high-coverage sequences and the genome assemblies facilitate many aspects of genome-wide studies that include miRNA identification. Furthermore, the study of transcriptomic profiles for its fruit development and ripening allows direct comparison between miRNAs and their target genes expression [3,4]. In this study, we use a combined strategy—computational prediction and experiment identification—to study *P. dactylifera* or *Pd* miRNAs. In the computational approach, we use several filters based on known characteristics of plant miRNAs [13,16]. In the experimental approach, we sequence 6 small RNA libraries from different *P. dactylifera* fruit development stages, using the next-generation sequencing platforms. We correlate the miRNA expression profiles to their corresponding transcriptome from our previous studies of the *P. dactylifera* fruit development stages and find a clear negative correlation between most miRNAs and their target gene transcripts. Several important miRNAs are up-regulated during fruit development and ripening, and the mRNA targets of these fruit-development-associated (FDA) miRNAs are found involved in carbon metabolism, such as starch and sucrose metabolisms, and our results suggest that miRNA may play critical roles in regulating sugar accumulation and conversion, which are key transitional events to fruit ripening. Our genome-wide *Pd* miRNA characterization provides novel insights into the dynamics of miRNAs in regulating fruit development and evidence for understanding gene regulation in fruit development and ripening.

2. Results

2.1. Prediction of conserved miRNAs in P. dactylifera genome

In the plant kingdom, a substantial number of miRNAs are highly conserved in different lineages, ranging from mosses and gymnosperms to angiosperms. Such homologous miRNA families are also typically function-conserved and play essential regulatory roles across plant taxa [17]. Sequence and structure homologies are the central theory of computing-based approaches for miRNA prediction, and many computational methods and protocols have been developed to date [18,19]. We obtained 5943 known mature miRNA sequences from the miRBase (Release 19.0 August 2012) [14] and searched them against the *P. dactylifera* genome assembly put together by our group. The effort yielded 238 potential *Pd* miRNAs, which belong to 54 miRNA families (Supplementary Table S1 in Ref. [20]). Of the 54 miRNA families, the miR169 family is the largest with 24 members, similar to what has been observed in other plant species [21,22]. We have 6 other families, miR171, miR156, miR164, miR167, miR172 and miR529, whose members per family are all greater than 10 (Fig. 1A). Among the 54 miRNA families, 19 are highly conserved (more than 10 species including *Arabidopsis* and rice) among many plant species, of which 8 families (miR156, miR159, miR160, miR166, miR167, miR171, miR172, and miR396) are present in more than 30 monocots and dicots, whereas the other 35 show less evolutionary conservation (less than 10 species) (Fig. 2). Interestingly, several dicot-associated miRNAs (including miR391, miR479, miR828, miR845, miR856, miR1511, miR1918, miR2673, miR5225, miR5645, miR6034, and miR6150) are identified in *P. dactylifera* which is actually a monocotyledon plant but a rather primitive one according to plant systematics. Moreover, miR536, which is only identified in the lower plants, *Physcomitrella patens* (bryophyte) and *Selaginella moellendorffii* (lycophyte) (Fig. 2), is found for the first time in an angiosperm. These data suggest that the *Pd* miRNAs are both specific and complex.

The length of the 238 predicted conserved miRNAs ranges from 18 to 24 nt, where 21-nt miRNA takes the majority (43.28%), followed by the 19-nt (16.80%), 22-nt (15.96%), 20-nt (13.86%), 18-nt (6.30%), 23-nt

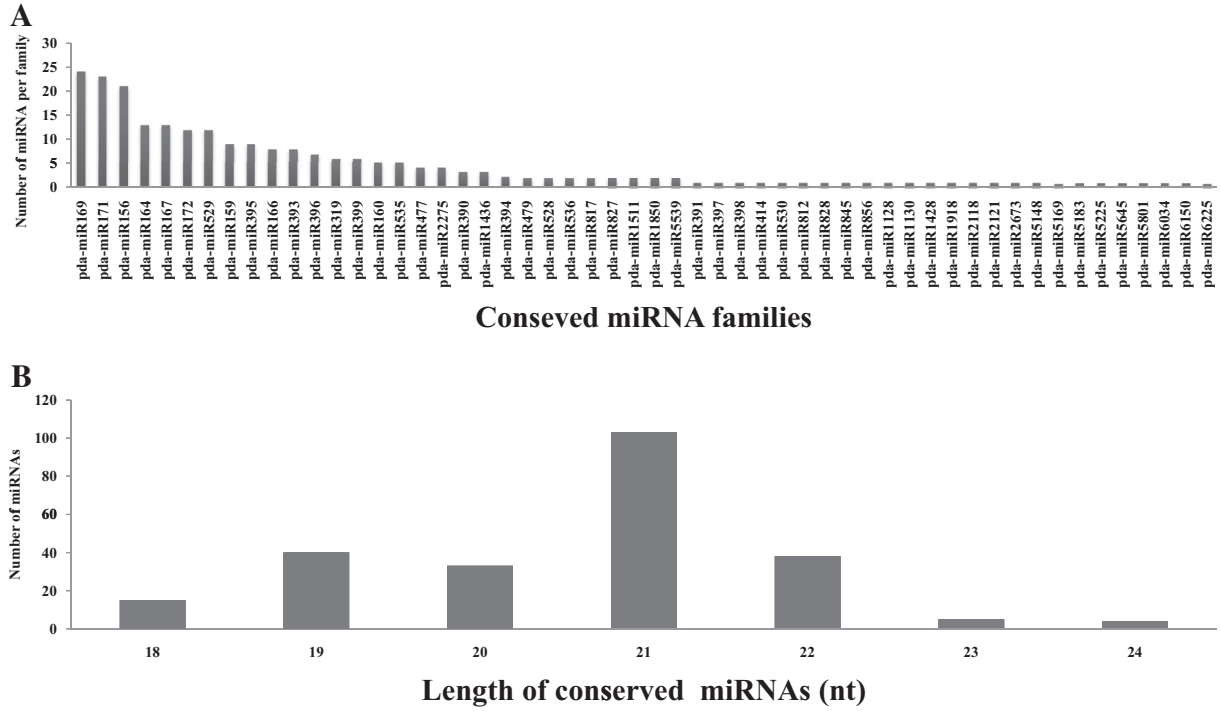

Fig. 1. The distribution of conserved miRNAs in *P. dactylifera*. A, MiRNA members in different miRNA families. B, The length distribution of conserved miRNAs.