Gene 608 (2017) 66–72

Contents lists available at ScienceDirect

Gene

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

Identification and characterization of novel microRNAs for fruit development and quality in hot pepper (*Capsicum annuum* L.)

Zhoubin Liu [a,b,1], Yuping Zhang [c,1], Lijun Ou [b,1], Linyu Kang [a,b], Yuhua Liu [a,b], Junheng Lv [a,b], Ge Wei [b], Bozhi Yang [b], Sha Yang [b], Wenchao Chen [b], Xiongze Dai [b], Xuefeng Li [b], Shudong Zhou [b], Zhuqing Zhang [b,*], Yanqing Ma [b,*], Xuexiao Zou [a,b,**]

[a] Longping Branch, Graduate School of Hunan university, Changsha 410125, China
[b] Hunan Vegetable Research Institute, Changsha 410125, China
[c] College of Resources and Environment, Hunan Agricultural University, Changsha 410128, China

ARTICLE INFO

Article history:
Received 8 September 2016
Received in revised form 16 January 2017
Accepted 20 January 2017
Available online 22 January 2017

Keywords:
Capsicum annuum
miRNA
Fruit development
Fruit quality
Deep sequencing

ABSTRACT

MicroRNAs (miRNAs) are non-coding small RNAs which play an important regulatory role in various biological processes. Previous studies have reported that miRNAs are involved in fruit development in model plants. However, the miRNAs related to fruit development and quality in hot pepper (*Capsicum annuum* L.) remains unknown. In this study, small RNA populations from different fruit ripening stages and different varieties were compared using next-generation sequencing technology. Totally, 59 known miRNAs and 310 novel miRNAs were identified from four libraries using miRDeep2 software. For these novel miRNAs, 656 targets were predicted and 402 of them were annotated. GO analysis and KEGG pathways suggested that some of the predicted miRNAs targeted genes involved in starch sucrose metabolism and amino sugar as well as nucleotide sugar metabolism. Quantitative RT-PCR validated the contrasting expression patterns between several miRNAs and their target genes. These results will provide an important foundation for future studies on the regulation of miRNAs involved in fruit development and quality.

1. Introduction

Hot pepper (*Capsicum annuum* L.), belonging to the Solanaceae family, is one of the most important and widely cultivated vegetable crops. Due to their typical colour, pungency, taste and/or distinct aroma, fruits of pepper are commonly used in the diet (Govindarajan, 1985). The fruits of hot pepper can synthesize and accumulate capsaicinoids, pigments (anthocyanins and carotenoids) and vitamins A, B and C (Aza-González et al., 2011). For hot pepper breeding, yield and fruit quality were the important quality parameter during the past decade (Eggink et al., 2014). Fruit wall thickness and total soluble solids are important characteristics for fruit quality in many plants (do Rêgo et al., 2011). Hot pepper was recorded to have very high levels of vitamin C accumulation in the fruits (Isabelle et al., 2010). Previous studies revealed that matured (unripened) fruits contained more vitamin C and fewer reducing sugars, compared to ripened fruits (Perla et al., 2016). Recently, transcriptomes of *Capsicum* species have been studied for fruit development and biosynthesis of capsaicinoids. The transcriptomes of Serrano-type chili pepper ('Tampiqueño 74') fruits collected at 10, 20, 40 and 60 days after anthesis (DAA) have been analyzed (Martínez-López et al., 2014). Using RNA-seq, the studies revealed that genes involved in capsaicinoid and ascorbic acid biosynthesis were highly expressed during fruit maturation (Liu et al., 2013; Martínez-López et al., 2014).

MicroRNAs (miRNAs) are a class of non-coding small RNAs with approximately 21 nucleotides (nt) in length, play a fundamental role in plant biological processes, such as development, signal transduction and biotic and abiotic stress responses (Hu et al., 2013a). In most cases, miRNAs are processed from single-stranded small RNAs digested by DICER-LIKE1 (DCL1) enzymes, resulting in the biogenesis of a mature miRNA duplex (Chen, 2009). After unwinding of the duplexes, the single stranded mature miRNAs are released. Finally, the mature miRNA enters the RNA-induced silencing complex (RISC), either by cleavage or translational repression, of complementary transcripts (Bartel, 2004). Although many miRNAs have been identified in a few model plants and the discovery of these miRNAs and their targets makes it possible to further understand the gene regulation network and biological mechanism mediated by miRNAs. Recently, the genome sequences of

Abbreviations: DAA, days after anthesis; RISC, RNA-induced silencing complex; KOG, Clusters of Orthologous Groups of proteins; KEGG, Kyoto Encyclopedia of Genes and Genomes; GO, Gene Ontology; AGO, Argonaute.

* Corresponding authors.
** Correspondence to: X. Zou, Longping Branch, Graduate School of Hunan University, Hunan Vegetable Research Institute, Changsha 410125, China
E-mail addresses: cszzq@126.com (Z. Zhang), yanqingmahn@163.com (Y. Ma), zouxuexiao428@163.com (X. Zou).

[1] The authors contrubited equally.

http://dx.doi.org/10.1016/j.gene.2017.01.020
0378-1119/

hot pepper have been sequenced (Qin et al., 2014; Kim et al., 2014) and made available for further identifying miRNAs in *Capsicum annuum*. Using high-throughput sequencing technology, 29 conserved and 35 novel miRNA families s were identified in hot pepper (Hwang et al., 2013). However, less is known about miRNAs in fruit development and quality of hot pepper.

In the current study, we investigated the miRNA profiles of hot pepper fruits with different quality during the fruit development using an Illumina HiSeq 2500 platform. The differentially expressed miRNAs involved in fruit ripening were identified, and the corresponding target genes were also predicted. Subsequently, the potential functions of the differentially expressed miRNAs and their target genes were discussed. In addition, the expression patterns of several miRNAs and their target genes were examined by qRT-PCR methods. These data provide novel insights into the molecular mechanisms of hot pepper fruit development and quality. The identification and characterization of known and novel miRNAs will enable better understanding the roles of these miRNAs in the fruit development of pepper.

2. Materials and methods

2.1. Plant materials

Two hot pepper varieties 'Luosijiao' and '06J19-1-1-1-2', were grown in greenhouse in the Hunan vegetable research institute under normal farming conditions (16 h of light at 26 °C and 8 h of darkness at 20 °C). The variety 'Luosijiao' is widely planted in China, with better flavor and more vitamin C than the variety '06J19-1-1-1-2'. To collect the fruits, they were tagged at anthesis, and harvested at the mature green at 30 days after anthesis (30DAA, *A*1), colour-changed period (40DAA, A2) and red-ripening (50DAA, A3) of variety 'Luosijiao'. And for the variety '06J19-1-1-1-2', fruit at the mature green at 30 days after anthesis (30DAA, *B*1) was collected. The four samples from different plants were pooled and immediately store at −80 °C until further use.

2.2. RNA extraction and deep sequencing

Total RNA was extracted from pepper fruits at different stages using the Trizol (Invitrogen) according to the manufacturer's instructions. For each sample, 5 μg of total RNA was used as the input material for constructing a small RNA library. Sequencing libraries were generated using NEBNext Multiplex Small RNA Library Prep Set for Illumina (NEB, USA) according to the manufacturer's recommendations. Briefly, the small RNAs were ligated with 3′ and 5′ adapters using T4 RNA ligase. The RNAs were subsequently transcribed to single-stranded cDNA using Super-Script II Reverse Transcriptase (RNase H-) (Invitrogen). Thereafter, PCR amplification was performed using LongAmp Taq2X Master Mix and primers that anneal to adapters. PCR products were purified on an 8% polyacrylamide gel (100 V, 80 min). After quality assessment using DNA High Sensitivity Chips with the Agilent 2100 Bioanalyzer, DNA fragments 140–160 bp in length were recovered and dissolved in 8 μl of elution buffer for sequencing. Then, the four libraries were sequenced by Biomarker Technologies (Beijing, China) using the Illumina HiSeq 2500 platform with single-end (SE) 50 nt.

2.3. Bioinformatic analysis of miRNAs

After sequencing, raw reads were filtered to remove low quality sequences, adapter sequences, reads <18 nt in length and reads with poly N from the raw data. The clean reads were blasted against the RepeatMasker and Rfam database (http://www.sanger.ac.uk/software/Rfam) to exclude rRNAs, tRNAs, snRNAs, snoRNAs and other non-coding RNAs. The remaining sequences were searched against the miRBase 21.0 database (http://www.mirbase.org) to identify putative known miRNAs (Kozomara and Griffiths-Jones, 2011). The remaining unannotated sRNA sequences were analyzed by miRDeep2 (Friedlander et al., 2012) software to predict potential novel miRNAs by exploring hairpin structure, Dicer cleavage sites and the minimum free energy. The criteria used for novel miRNA were based on the work of Meyers et al. (2008).

2.4. Analysis of differentially expressed miRNAs

To define the expression level of miRNAs, miRNA count was normalized as transcripts per million (TPM) with the following formula: normalized expression = mapped read count/total reads $\times 10^6$. Differential expression analysis of two samples was performed using the DEGseq R package (Wang et al., 2010) and *p*-values were adjusted using *q*-value. The criteria of $q < 0.01$ and $|\log2(\text{fold change})| > 1$ were set as the threshold for statistically significant differential expression.

2.5. Target gene predication and enrichment analysis

Putative target genes of miRNAs identified in the present study were predicated in silico using psRNATarget (Dai and Zhao, 2011) based on sequence similarity to the transcripts of hot pepper (Qin et al., 2014). The functional classification and pathways were based on the three databases: KOG (Clusters of Orthologous Groups of proteins), KEGG (Kyoto Encyclopedia of Genes and Genomes) and GO (Gene Ontology), using BLAST with a cutoff *E*-value of 10–5. GO enrichment analysis was performed using the TopGO R package; and GO terms with corrected FDR < 0.05 were considered significantly enriched (Alexa et al., 2006). Enriched GO terms were plotted using WEGO (http://wego.genomics.org.cn/cgi-bin/wego/index/pl).

2.6. Quantitative RT-PCR (qRT-PCR) and stem-loop qRT-PCR

Total RNA was extracted from the samples using Trizol (Takara, Dalian, China) and treated with RNase-free Dnase I (Promega, USA). For each miRNA, 2 μg of DNase I-treated total RNA was used in a reverse transcription reaction with the Prime-Script RT reagent Kit (Takara, Dalian, China). Reverse transcription was performed with the following conditions: the reactions were incubated for 30 min at 16 °C, followed by 60 cycles of pulsed RT at 30 °C for 30 s, 42 °C for 30 s and 50 °C for 1 s and finally the reactions were terminated at 70 °C for 5 min (Varkonyi-Gasic et al., 2007). For target genes, 1 μg of Dnase I-treated total RNA was used for synthesis with the Prime-Script RT reagent Kit (Takara, Dalian, China) and an oligo (dT)18 primer. Reverse transcription was performed using the following conditions: 37 °C for 15 min, 42 °C for 1 h, and then holding at 4 °C. All the primers are listed in Table S1.

Real time qRT-PCR analysis of the miRNA and their targets was performed using the FastStart Universal SYBR Green Master Mix (Roche) on the StepOne plus PCR platform (Applied Biosystems). The qRT-PCRs reactions were conducted with the following protocol: 95 °C for 10 min, followed by 40 cycles of 95 °C for 15 s and 56 °C for 30s and 72 °C for 15 s. *CpAction* was used as the internal control for miRNA and their targets. A melting curve analysis was performed to determine

Table 1
Distribution of small RNAs among different categories.

	A1	A2	A3	B1
Raw reads	24,048,077	20,929,299	19,362,716	19,299,815
Clean reads	22,799,181	19,120,746	16,365,419	18,270,760
Uniq Mapped reads	16,075,461	11,662,314	9,236,660	12,481,144
rRNA	1,436,622	2,180,510	3,371,931	1,278,345
snRNA	1587	2174	2436	950
snoRNA	54	58	58	71
tRNA	120,971	123,760	178,480	81,848
Repbase	31,706	25,474	21,955	26,624
Unannotated	21,208,241	16,788,770	12,790,559	16,882,922