Gene 667 (2018) 25–33

Contents lists available at ScienceDirect

Gene

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

Research paper

Integrative analysis of long non-coding RNA acting as ceRNAs involved in chilling injury in tomato fruit

Yunxiang Wang[a,b], Lipu Gao[a], Benzhong Zhu[c], Hongliang Zhu[c], Yunbo Luo[c], Qing Wang[a,*], Jinhua Zuo[a,d,*]

[a] Key Laboratory of the Vegetable Postharvest Treatment of Ministry of Agriculture, Beijing Key Laboratory of Fruits and Vegetable Storage and Processing, Key Laboratory of Biology and Genetic Improvement of Horticultural Crops (North China) of Ministry of Agriculture, Key Laboratory of Urban Agriculture (North) of Ministry of Agriculture, Beijing Vegetable Research Center, Beijing Academy of Agriculture and Forestry Sciences, Beijing, China
[b] Beijing Academy of Forestry and Pomology Sciences, Beijing Academy of Agricultue and Forestry Sciences, Beijing, China
[c] Laboratory of Postharvest Molecular Biology of Fruits and Vegetables, Department of Food Biotechnology, College of Food Science and Nutritional Engineering, China Agricultural University, Beijing 100083, China
[d] Boyce Thompson Institute for Plant Research, Cornell University Campus, Ithaca, NY 14853, USA

ARTICLE INFO

Keywords:
Tomato fruit
Chilling injury
lncRNAs
ceRNAs
High-throughput sequencing
Regulatory network

ABSTRACT

Long-non-coding RNA (LncRNA) is a kind of non-coding endogenous RNA that plays essential roles in diverse biological processes and various stress responses. To identify and elucidate the intricate regulatory roles of lncRNAs in chilling injury in tomato fruit, deep sequencing and bioinformatics methods were performed here. After strict screening, a total of 1411 lncRNAs were identified. Among these lncRNAs, 239 of them were significantly differentially expressed. A large amount of target genes were identified and many of them were found to code chilling stress related proteins, including redox reaction related enzyme, important enzymes about cell wall degradation, membrane lipid peroxidation related enzymes, heat and cold shock protein, energy metabolism related enzymes, salicylic acid and abscisic acid metabolism related genes. Interestingly, 41 lncRNAs were found to be the precursor of 33 miRNAs, and 186 lncRNAs were targets of 45 miRNAs. These lncRNAs targeted by miRNAs might be potential ceRNAs. Particularly, a sophisticated regulatory model including miRNAs, lncRNAs and their targets was set up. This model revealed that some miRNAs and lncRNAs may be involved in chilling injury, which provided a new perspective of lncRNAs role.

1. Introduction

Non-coding RNAs (ncRNAs), which have no or little protein-coding potential, take up a large proportion of the eukaryotic genome and can work as regulatory molecules in various biological processes and responses to stress (Zhang and Chen, 2013; Liu et al., 2015; Wang et al., 2016). As part of ncRNAs, the understanding of functions and mechanisms of miRNAs and small interfering RNAs (ta-siRNAs and nat-siRNAs) in gene expression regulation has made great progress in the past few years (Cuperus et al., 2011; Chen et al., 2012). Recently, long non-coding RNA (lncRNA) with no predicted open reading frame and the length of above 200 nucleotides, which is a subgroup of non-coding RNA (ncRNA), has been widely identified as new regulatory elements that are involved in many biological processes in animals and plants (Cheetham et al., 2013; Leung et al., 2013; Ng et al., 2013). LncRNAs can be classified into three types based on their genomic origins: long intergenic ncRNAs (lincRNAs), intronic ncRNAs (incRNAs) and natural antisense transcripts (NATs). In plants, lncRNAs are mainly transcribed by RNA polymerase II (Pol II), and a small portion of them could also be transcribed by Pol IV and Pol V (Wierzbicki, 2012; Wu et al., 2012; Kim and Sung, 2012; Ulitsky and Bartel, 2013).

The application of next-generation sequencing technology greatly facilitated the discovery of lncRNAs in plants (Wang et al., 2015). Thousands of lncRNAs have been identified using deep sequencing and bioinformatics analyses in *Arabidopsis thaliana*, *Oryza sativa*, *Zea mays*, *Solanum lycopersicum*, *Medicago truncatula* and *Cucumis sativus* (Boerner and McGinnis, 2012; Li et al., 2014; W. Zhang et al., 2014; Y.C. Zhang et al., 2014; Wen et al., 2007; Hao et al., 2015; Zhu et al., 2015).

Abbreviations: ceRNAs, competing endogenous RNAs; incRNAs, intronic ncRNAs; lincRNAs, long intergenic ncRNAs; lncRNA, long non-coding RNA; NATs, natural antisense transcripts; nat-siRNAs, natural antisense short interfering RNAs; ncRNAs, non-coding RNAs; ta-siRNAs, trans-acting small interfering RNAs

* Corresponding authors at: Key Laboratory of the Vegetable Postharvest Treatment of Ministry of Agriculture, Beijing Key Laboratory of Fruits and Vegetable Storage and Processing, Key Laboratory of Biology and Genetic Improvement of Horticultural Crops (North China) of Ministry of Agriculture, Key Laboratory of Urban Agriculture (North) of Ministry of Agriculture, Beijing Vegetable Research Center, Beijing Academy of Agriculture and Forestry Sciences, Beijing, China.
E-mail addresses: wangqing@nercv.org (Q. Wang), zuojinhua@nercv.org (J. Zuo).

https://doi.org/10.1016/j.gene.2018.05.030
Received 6 March 2018; Received in revised form 21 April 2018; Accepted 10 May 2018

Previous reports have shown that these lncRNAs can participate in various biological processes and respond to biotic and abiotic stresses via regulating target gene expression (W. Zhang et al., 2014; Y.C. Zhang et al., 2014; Wang et al., 2014; Xin et al., 2011). For example, a lncRNA named COOLAIR, could regulate flowering by bounding to and repressing the expression of FLC (flowering locus C) after vernalization in *A. thaliana* (Liu et al., 2012). Acting as competing endogenous RNAs (ceRNAs), lncRNAs may also bind with miRNAs to regulate the mRNAs expression (W. Zhang et al., 2014; Y.C. Zhang et al., 2014; Shuai et al., 2014; Wu et al., 2013).

Recent reports have described an intricate interplay among diverse RNA species and the hypothesis that all RNA transcripts containing miRNA-binding sites can communicate with and regulate each other by competing for shared miRNAs have been built (Tay et al., 2014). These RNA transcripts binding to shared miRNAs could be called ceRNAs and lncRNAs, circRNAs and the protein coding genes all could act as ceRNAs (Ebert and Sharp, 2010; Salmena et al., 2011; Seitz, 2009). The first examples of ceRNAs was found in *Arabidopsis* and the study showed that the non-coding RNA IPS1 could influence the expression level of PHO2 by binding to miR399 (Franco-Zorrilla et al., 2007). Recent studies also showed that lncRNAs, circular RNAs (circRNAs), pseudogene and mRNAs all could act as ceRNAs (Cesana et al., 2011; Wang et al., 2010; Hansen et al., 2013; Memczak et al., 2013; Karreth et al., 2011; Johnsson et al., 2013).

Cold storage is one of the most effective postharvest technologies to control quality of fruit and vegetables from the time of harvest until final preparation for human consumption (Kondo et al., 2005; Zou et al., 2014). However, there is a great risk of cold storage for postharvest produces to suffer chilling injury (CI), especially the cold sensitive crops. Tomato fruit is susceptible to chilling injury when exposed to low temperatures. Chilled tomato fruits may show several symptoms such as sunken areas on the fruits (blemishes), diseases caused by pathogen, and inability to develop full color, which will lead to substantial degradation of produce quality of tomato fruit (Zhao et al., 2009; Soleimani, 2013). Studies were carried out on the fruit ripening and virus stress related lncRNAs in tomato (Wang et al., 2015; Zhu et al., 2015), however, the relation between lncRNAs and the chilling injury in tomato fruit is unknown.

To investigate the molecular basis of the relationship between lncRNAs and chilling injury in tomato fruits, deep sequencing and bioinformatic analysis were performed. LncRNAs were characterized at a whole genome level and their expression patterns were parsed simultaneously. The target genes were parsed and screened and a regulation model was set up. This research provided new perspective of lncRNAs role in chilling tolerance in tomato fruits.

2. Materials and methods

2.1. Sample collection, RNA quantification and qualification

The *Solanum lycopersicum* 'Rui Xin' (heterozygous) tomatoes were grown in the greenhouse at standard greenhouse conditions (24 °C, 75% relative humidity, 16/8-h light/dark cycle). Tomato fruits at mature green stage (40 days postanthesis) were harvested to prepare for the chilling injury experiments. The tomato fruits were divided into two groups. The control group was put in the refrigeration house for 72 h at 10 °C, the chilling injury group was put in the refrigeration house for 72 h at 0 °C. After storage, all the tomatoes of the chilling injury group showed pitting symptom while the control group had none. Pooled mesocarp tissues from two groups (two biological replicates, each biological sample is combined from ten fruits) were collected and immediately frozen in liquid nitrogen. All the samples were stored at −80 °C until further analysis. The RNA samples (about 1 g tissue combined from ten fruits) were extracted with Trizol. RNA degradation and contamination, especially DNA contamination, was monitored on 1.5% agarose gels. RNA concentration and purity was measured using the NanoDrop 2000 Spectrophotometer (Thermo Fisher Scientific, Wilmington, DE). RNA integrity was assessed using the RNA Nano 6000 Assay Kit of the Agilent Bioanalyzer 2100 System (Agilent Technologies, CA, USA).

2.2. Library preparation for lncRNA-seq

A total amount of 1.5 μg RNA per sample was used as input material for rRNA removal using the Ribo-Zero rRNA Removal Kit (Epicentre, Madison, WI, USA). Sequencing libraries were generated using NEBNext® Ultra™ Directional RNA Library Prep Kit for Illumina® (NEB, USA) following manufacturer's recommendations and index codes were added to attribute sequences to each sample. Briefly, fragmentation was carried out using divalent cations under elevated temperature in NEBNext First Strand Synthesis Reaction Buffer (5×). First strand cDNA was synthesized using random hexamer primer and Reverse Transcriptase. Second-strand cDNA synthesis was subsequently performed using DNA Polymerase I and RNase H. Remaining overhangs were converted into blunt ends via exonuclease/polymerase activities. After adenylation of 3′ ends of DNA fragments, NEBNext Adaptor with hairpin loop structure were ligated to prepare for hybridization. In order to select insert fragments of preferentially 150–200 bp in length, the library fragments were purified with AMPure XP Beads (Beckman Coulter, Beverly, USA). Then 3 μl USER Enzyme (NEB, USA) was used with size-selected, adaptor-ligated cDNA at 37 °C for 15 min before PCR. Then PCR was performed with Phusion High-Fidelity DNA polymerase, Universal PCR primers and Index(X) Primer. At last, PCR products were purified (AMPure XP system) and library quality was assessed on the Agilent Bioanalyzer 2100 and qPCR.

2.3. Clustering and sequencing

The clustering of the index-coded samples was performed on acBot Cluster Generation System using TruSeq PE Cluster Kitv3-cBot-HS (Illumia) according to the manufacturer's instructions. After cluster generation, the library preparations were sequenced on an Illumina Hiseq platform and paired-end reads (125 bp) were generated. The sequencing results were deposited in the Sequence Read Archive (SRA) at the NCBI database (accession number: SRP094091).

2.4. Quality control

Raw data (raw reads) of fastq format were firstly processed through in-house perl scripts. In this step, clean data (clean reads) were obtained by removing reads containing adapter, reads containing ploy-N and low quality reads (reads that base quality scores < 10 counting for 50% or more) from raw data. The base quality scores (Q-score) were the integer mapping of the error probability of base calling. Q10 indicated one in ten bases will calling error. At the same time, Q20, Q30, GC-content and sequence duplication level of the clean data were calculated. The Q30 in all the samples were > 85.01%. All the downstream analyses were based on clean data with high quality.

2.5. LncRNA identification and analysis

The transcriptome was assembled using the Cufflinks and Scripture based on the reads mapped to the reference genome. Briefly, each RNA-seq dataset was aligned to the tomato genome (SL2.50) independently using the TopHat 2.0 program. The parameters used were "mismatch 2 (-N 2), Insert_size 40 (-r 40)," and other parameters were default. The transcriptome was assembled independently using Cufflinks (version 2.2.1, the parameters used were "operation core number 4, library-type fr-unstranded," and other parameters were default) (Trapnell et al., 2012) and Scripture programs (versions VPaperR3) (Guttman et al., 2010) and the assemblies were then merged to produce a final transcriptome. After the final transcriptome was produced, the long non-