Biochemical and Biophysical Research Communications xxx (2015) 1–7

Contents lists available at ScienceDirect

Biochemical and Biophysical Research Communications

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

Comparative transcriptome analysis revealed the genotype specific cold response mechanism in tobacco

Risheng Hu [a], Xianxin Zhu [b], Shipeng Xiang [a, b], Youguo Zhan [c], Mingdong Zhu [d], Hanqi Yin [e], Qingming Zhou [b], Lieshu Zhu [a, b], Xianwen Zhang [f, *], Zhi Liu [f, **]

[a] Central South Agricultural Experiment Station of China Tobacco, Changsha, 410004, China
[b] College of Agronomy, Hunan Agricultural University, Changsha, 410128, China
[c] Kunming Municipal Tobacco Company, Kunming, 650000, China
[d] Hunan Rice Research Institute, Changsha, 410125, China
[e] Shanghai Biotechnology Corporation, Shanghai, 201203, China
[f] College of Bioscience and Biotechnology, Hunan Agricultural University, Changsha, 410128, China

ARTICLE INFO

Article history:
Received 8 December 2015
Accepted 10 December 2015
Available online xxx

Keywords:
Cold stress
Transcriptome
Tobacco
Circadian clock

ABSTRACT

Cold stress is a major adverse environmental factor that affects plant growth, development, productivity and quality. In the present study, comparative genome-wide transcriptome analysis on two tobacco (*Nicotiana tobacum* L.) cultivars, cold-tolerant NC567 and cold-sensitive Taiyan8, was performed using RNA-seq technology. After the first assembly, total length of unigenes is from 101,308,644 to 123,781,795 bp, the N50 length is from 1357 to 1475 bp, and 152,688 unigenes in NC567 and 144,160 unigenes in Taiyan8 were identified, respectively. Functional classification of cold-responsive (COR) genes showed that the genes involved in cell wall metabolism, transcription factors, ubiquitin-proteasome system (UPS) and signaling are over-represented, and the COR genes are specifically induced during cold stress in NC567. Pathway analysis revealed the significant enrichment of the COR genes in plant circadian clock. Taken together, the present study suggested the positive roles of the highly induced expression of the COR genes and the conserved mechanism of circadian clock related genes in tobacco response to cold stress, and provided some valuable genes for crop improvement to cope with cold stress.

1. Introduction

Cold stress is a major environmental factor that affects plant growth, development, productivity and distribution. In higher plants, the well-known major cold signaling pathway is the C-repeat (CRT)-binding factor/dehydration-responsive element (DRE) binding factor (CBF/DREB)-mediated transcriptional regulatory cascade, which is essential for the induction of a set of cold responsive (COR) genes. Overexpression of ICE1 (inducer of CBF expression 1), an upstream transcription factor that regulates the transcription of *CBF* genes in the cold, enhances the expression of the *CBF* regulon in the cold and improves freezing tolerance of the transgenic plants [1], and the ICE1-CBF pathway model critical for acquired freezing tolerance was then postulated [2]. It's well-documented that transcription factors play important roles in stress response. A study has elucidated the significant role of ethylene response factor (ERF) proteins in plant adaptation to drought, freezing, and salt in tobacco [3]. In the *Anthurium* transcriptome study, 4363 differentially expressed genes were cold-responsive, and a total of 39 cold-inducible transcription factors were identified, including subsets of *AP2/ERF*, *Zinc finger*, *NAC*, *MYB* and *bZIP* family members [4].

The early *Arabidopsis* transcriptome analysis indicated that multiple regulatory pathways are activated upon cold acclimation in addition to CBF transcriptional activators cold response pathway [5]. An *Arabidopsis thaliana* genome expression profile identified a large number of early cold-responsive genes for transcription factors [6]. The transcriptome analyses showed that the genome-wide responses of transcription to the temperature shift may be associated with chromatin dynamics [7].

* Corresponding author.
** Corresponding author.
E-mail addresses: xianwenzhang@hunau.edu.cn (X. Zhang), tigerzhiliu@126.com (Z. Liu).

http://dx.doi.org/10.1016/j.bbrc.2015.12.040
0006-291X/

The tobacco is one of economically important plants, and cold stress at seedling stage shows serious effects on tobacco growth and production. However, the molecular mechanism of tobacco response to cold stress remains limited. Here, we reported the time course gene expression profile of cold-tolerant tobacco cultivar NC567 and cold-sensitive Taiyan8 from 0 h to 48 h after 6 °C treatment using Illumina RNA-seq technology. The comprehensive analysis revealed the significant expression difference between two tobacco genotypes, and the important roles of transcription factors, ubiquitin-proteasome system (UPS), signaling, cell wall metabolism, and circadian clock related genes in tobacco acclimation to cold stress. The study provided some valuable information for crop improvement to deal with low temperature.

2. Materials and methods

2.1. Plant growth and cold treatments

In our previous study, a new method was developed to characterize the cold tolerance of different tobacco genotypes, and the cold-tolerant tobacco (*Nicotiana tobacum* L.) cultivar NC567 and cold-sensitive Taiyan8 were identified [8]. Both NC567 and Taiyan8 were cultured in plastic pots in a ventilated greenhouse. The seedlings with 6–8 leaves were moved into a growth chamber (Binder, Tuttlingen, Germany) and cultured at about 6 °C with humidity of 80% (16/8 h, light/dark). The third leaf from the top from three plants were collected at 0 h (as control), 4 h, 24 h, and 48 h after cold treatment, washed carefully, and separately frozen in liquid nitrogen and stored at −80 °C prior to RNA extraction. In total, 8 RNA pools were achieved, and each of the RNA pool was made by the mixture.

2.2. RNA isolation, cDNA library construction and sequencing

Total RNA was isolated from each pooled sample using Trizol reagent (Invitrogen, Carlsbad, CA) according to the manufacturer's protocol. RNA quality and quantity were determined with Agilent 2100 Bioanalyzer (Agilent, MN). Illumina sequencing libraries were constructed following a modified strand-specific RNA-seq protocol. The samples were amplified with TruSeq PCR primers and sequenced on the Illumina HiSeq2500 platform at Shanghai Biochip Corporation. RNA-seq raw data were deposited to the NCBI Sequence Read Archive (NCBI SRA) under accession number SRP059707.

2.3. De novo assembly and functional annotation

In order to remove the low quality nucleotides, adapter, and PCR primer sequences, raw RNA-seq reads were processed with the ShortRead package. Reads longer than 35 nt and with no more than 2 N (ambiguous nucleotides) were retained. The resulting cleaned reads of each sample was then assembled as unigenes using the Trinity package with optimized k-mer length of 25 [9], respectively. The assembled unigenes of all the samples were cleaned by removing redundancies and further assembled as all-unigenes. The cleaned reads of each sample were aligned to the assembled all-unigenes using FANSe2 allowing 7 nt mismatch, respectively. The all-unigene with at least 10 mapped reads were considered as reliably assembled unigenes.

The gene annotations for the entire all-unigenes were performed via blastx search against Nr (NCBI non-redundant protein sequences), UniProt (The Universal Protein Resource), and Cluster of Orthologous Groups (COG) database with an E-value of 1e-05.

2.4. Quantification of gene expression levels and differential expression analysis

Clean data were mapped back onto the assembled transcriptome for the quantification of gene expression. Readcount for each gene was obtained from the mapping results and normalized to reads per kilobase of exon model per million mapped reads (RPKM). The unigenes annotated by PGSC (Potato Genome Sequencing Consortium, http://www.potatogenome.net) in all 8 samples with a cutoff E-value of 1e-05 were selected, and the readcounts of the selected unigenes were normalized by edgeR package [10] after deleting the unigenes whose count is below 10 in all 8 samples. The differentially expressed genes (DEGs) between the cold-treated samples and the control were determined with FDR $\leq$ 0.01 and |log2 (fold change)|>1. The time course DEGs were referred as the ones that were differentially expressed in at least one time point during cold treatment in NC567 and Taiyan8, respectively. The fold change (FC) value was log2 transformed in the further analysis. The pathway analysis was performed with MapMan [11] and KEGG database.

2.5. Quantitative real-time PCR (qRT-PCR) analysis

To validate RNA-seq results, the expression of 10 genes related to transcription factors, UPS, and circadian clock was verified by qRT-PCR. The primer sequences were designed in the laboratory and synthesized by Generay Biotech (Generay, PRC) based on the assembled unigene sequences (Table S9). The RNA preparation, cDNA synthesis, qRT-PCR and data process were accomplished as described [12] using L25, a ribosome gene, as the internal control.

3. Results

3.1. High-throughput sequencing and de novo assembly

More than 50 million reads for each sample were generated from eight tobacco RNA libraries. The low-quality and rRNA reads were removed, resulting in the least clean ratio of 95% for eight samples (Table S1). After the first assembly for each sample using Trinity software, the N50 length is from 1357 to 1475 bp (Table 1), and the unigenes in the range of 200–800 bp account for 63.5% of total unigenes (Fig. S1). The final unigenes for two tobacco genotypes were generated based on the first assembly sequences, resulting in 152,688 unigenes in NC567, and 144,160 unigenes in Taiyan8, N50 length of 1760 and 1766 bp, and the average length of 1213 and 1218 bp for NC567 and Taiyan8, respectively (Table 1).

3.2. Annotation and expression quantification of unigenes

As the tobacco genome is still not fully available, the final unigenes from NC567 and Taiyan8 were pooled together, and aligned against the UniProt databases [13]. The alignment analysis showed that the highest hits to species is potato (*Solanum tuberosum*) with a cutoff E-value of 1e-05 (Fig. 1). The number of unigenes commonly annotated by PGSC is 56,181 (Table S2), and the potato gene IDs were adopted for the study.

To verify the expression level of the unigenes identified by RNA-seq, 10 genes related to transcription factor, the E3 ubiquitin ligase, and circadian clock were selected, and quantitative real-time PCR (qRT-PCR) was performed. The majority of these candidate genes exhibited up-regulated expression pattern during cold stress in both tobacco genotypes according to qRT-PCR data. The results showed the excellent consistence between RNA-seq and qRT-PCR (Table 2).

Please cite this article in press as: R. Hu, et al., Comparative transcriptome analysis revealed the genotype specific cold response mechanism in tobacco, Biochemical and Biophysical Research Communications (2015), http://dx.doi.org/10.1016/j.bbrc.2015.12.040