Biochemical and Biophysical Research Communications xxx (2016) 1–8

Contents lists available at ScienceDirect

Biochemical and Biophysical Research Communications

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

Analysis on the transcriptome information of two different wheat mutants and identification of salt-induced differential genes

Xiaoli Ma [a, 1], Peihan Gu [a, 1], Wenji Liang [a, b], Yuxin Zhang [a], Xiaoli Jin [a], Shaoxiong Wang [a], Yinzhu Shen [a], Zhanjing Huang [a, *]

[a] College of Life Science, Hebei Normal University, Shijiazhuang 050024 Hebei, PR China
[b] College of Clinical Medicine, North China University of Science and Technology, Tangshan 063000 Hebei, PR China

ARTICLE INFO

Article history:
Received 7 April 2016
Accepted 8 April 2016
Available online xxx

Keywords:
Triticum aestivum
Salt tolerance
qRT-PCR
Transcriptome

ABSTRACT

This study established a wheat transcriptome library using RH8706-49 and RH8706-34. Salt-induced differential genes were screened by Illumina RNA sequencing (RNA-Seq). Five differential genes were chosen to study the functions by combining transcript sequencing result and gene chip. The expression changes of these five differential genes were analyzed using real-time quantitative PCR (qRT-PCR) technique to determine the reliability and accuracy of transcriptome sequencing and transplanted into *Arabidopsis thaliana* to obtain transgenic homozygote plants for the salt tolerance test. The salt tolerance test results show that the transgenic plants grew far better than the wild-type plant.

1. Introduction

Bread wheat (*Triticum aestivum*, AABBDD) is one of the most widely cultivated and consumed food crops in the world. However, the complex polyploid nature of its genome makes genetic and functional analyses extremely challenging [1]. Wheat, which is the most important food crop in the world, feeds approximately 40% of the global population. Accordingly, 20% of people worldwide acquire heat and proteins for human nutrition from wheat [2]. Wheat has since become a primary staple crop because of its enhanced adaptability to a wide range of climates and improved grain quality for the production of baker's flour [3].

Saline-alkali soil is extensively distributed on land. Incomplete statistics shows that the total area of saline-alkali soil is approximately 1 billion hm^2 [4], which accounts for approximately 10% of the total land area [5]. This size increases at an annual rate of $1.0–1.5 \times 10^6$ hm^2 [6]. Many countries and regions in the world face challenges as regards saline-alkali soil. Salinization influences 4.88% of the total available land area in China [7]. Salt stress influences plant yield, protein synthesis, photosynthesis, and energy metabolism [8] and is one of the main restrictions of agricultural production. Many experiments have shown that excessive salt in soil will hinder root systems in absorbing elements and influence the overall plant nutrition supply [9]. High salt stress will also cause osmotic stress to plants and disturb nutrition ion balance [10]. These effects will cause a series of secondary stresses to cells. Moreover, higher salt concentration causes a sharper reduction of the photosynthetic rate [11], thereby resulting in growth retardation and yield reduction of crops.

Wheat is sensitive to saline environments. Therefore, studying the salt tolerance mechanism of wheat is important. The whole genome shotgun sequencing of bread wheat [13], a drawing of the A genomic sequence draft of common wheat [1], and a drawing of the genomic sequence draft of *Aegilops tauschii*, which is a D genome donor of wheat [14], have been accomplished in recent years. These achievements have brought wheat study into a new development stage. Research on wheat breeding, germplasm resources, functional genomes, evolution, and comparative genomes has been greatly promoted.

Abbreviation: qPCR, quantitative polymerase chain reaction; SOM, self-organizing mapping; EST, expressed sequence tag; EMS, ethyl methyl sulfonate; RNA-Seq, RNA-Sequence; CDD, conserved domain database; Pfam, protein family database; NR, non-redundant database; KOG, clusters of orthologous groups of proteins for eukaryotic; KEGG, kyoto encyclopedia of genes and genomes; KAAS, KEGG automatic annotation server; GO, gene ontology; BWA, burrows-wheeler alignment tool; RPKM, reads per kilobases per million reads; FDR, false discovery rate; MS, Murashige/Skoog.

* Corresponding author.
E-mail addresses: maxiaoli04@163.com (X. Ma), jiajiajia343434@163.com (P. Gu), liangwenji_2005@163.com (W. Liang), 10303613@qq.com (Y. Zhang), 546884699@qq.com (X. Jin), 116575217@qq.com (S. Wang), 626869522@qq.com (Y. Shen), huangzhanjing@mail.hebtu.edu.cn, huangzhanjing@sohu.com (Z. Huang).

[1] These authors equally contributed to this work.

http://dx.doi.org/10.1016/j.bbrc.2016.04.039
0006-291X/

We have obtained a salt-tolerant mutant of wheat (RH8706-49; 49 salt-tolerant mutants) and salt-sensitive mutant of wheat (H8706-34; 34 salt-sensitive mutants) after years of breeding. Abundant EST related with salt tolerance have been acquired on the basis of the gene chip analysis of RH8706-49 collected at 5 moments (0, 1, 6, 12, and 72 h after 0.8% NaCl stress) [15].

Imperfect wheat genome information is available. Therefore, reading the transcription assembly from the beginning by using the short RNA sequence of Illumina is the most economical and effective environment-friendly technique to analyze generated and collected transcriptomes. This method can quickly and comprehensively collect almost all of the transcription sequence information of a tissue or an organ of a special species. The transcriptome sequence has been used in research on a transcriptional level difference between the early embryo and endosperm of corn seeds [16]. The transcriptome sequence has also been utilized in the response genes of Litchi to different light conditions and reactive oxygen [17,18], gene control network of corn seed [19], and new abundant transcripts with low expressive abundance in rice [20].

This study establishes a wheat transcriptome library using RH8706-49 and RH8706-34. Salt-induced differential genes are screened by Illumina RNA sequencing (RNA-Seq). The functions of some differential genes are verified. This sequencing acquires the salt-induced transcripts of wheat as comprehensively as possible, which provides valuable information and lays certain theoretical basis for improving and culturing salt-tolerant crop varieties. This study aims to provide references for the genetic improvement of wheat salt tolerance, which is significant to agricultural researchers.

2. Materials and methods

2.1. Plant materials

The single-seed generation of wheat Punong 365/Bainong 3039F1 was gained after another culture, EMS mutagenesis, and repeated salt tolerance screening. This generation was stably inherited for 15 generations. Consequently, RH8706-49(SR) and RH8706-34(SS) were obtained. These 2 mutants had similar genetic backgrounds. The salt tolerance index of RH8706-49 is $\geq$1.3, whereas that of RH8706-34 is $\leq$0.5. This result shows significant salt tolerance differences. Wild *Arabidopsis thaliana* (Columbia) seeds were provided by our laboratory.

2.2. RNA-Seq

Pre-processing of the read quality: a large quantity of sample data was obtained by the paired-end sequencing of Solexa RNA. The quality of the original data was pre-processed by considering the effect of the error rate of the Solexa data on results.

Pre-processing steps: the low-quality segments were eliminated using the sliding window method. The quality threshold was 20 (error rate = 1%). The window size was 5 bp, and the length threshold was 35 bp. Sequences with N in reads were cut, and the length threshold was set to 35 bp.

Pollution detection: 100,000 sequences were randomly chosen from the QC data of samples for comparison with the nt database. The best sequence ($e < 1e{-}10$ and coverage >80%) was chosen. Most sequences were compared with wheat genomes. No evident pollution were found in all samples.

2.3. Annotation

The sample unigenes were compared with the public data genes. The gene similarity comparison was mainly based on the BLAST algorithm. The sample gene sequence was compared with the Swiss-Prot, TrEMBL, CDD(Conserved Domain Database), Pfam(Protein family database), NR(non-redundant database), and KOG(Clusters of Orthologous Groups of proteins for Eukaryotic) libraries. All information on annotation with a similarity of >30% and $e < 1e{-}5$ were collected by gene combination. We analyzed the KEGG(Kyoto Encyclopedia of Genes and Genomes) pathway of the gained genes and predicted the corresponding ko number through KAAS(KEGG Automatic Annotation Server). The ko number corresponded to the KEGG pathway to analyze the relational file between genes and the enzyme annotation in KEGG as well as the information mapped onto the pathway. A statistical analysis on the GO(Gene Ontology) term of the genes under the biological process, cellular component, and molecular function classifications was conducted on the basis of BLAST uniprot (a combination result of Swiss-Prot and TrEMBL) with the GO classification of gained genes. The gained uniprot number was used to compare the GO terms.

2.4. Differential gene transcript analysis

Transcript expression abundance: the transcript expression abundance in different samples was calculated by a sequence similarity comparison by using the spliced transcript ($\geq$200 bp) as the library. One read is allowed to blast onto several transcripts by the single-end mapping method of bwa(BWA-Burrows-Wheeler Alignment tool) [21]. At this moment, every read was counted 1/n on every transcript. The reads on all transcript blast under each unigene were counted under unigene. Abundance RPKM(Reads Per Kilobases per Million reads) is defined as the reads per kilobase of exon model per million mapped reads.

Screening of differential genes: FDR(false discovery rate) $\leq$ 0.001 and fold change $\geq$ 2. The differential gene screening method used the sequencing-based method published by Audic et al. [22] on *Genome Research*.

2.5. Differential gene function annotation

We analyzed GO and KEGG enrichment by using differential genes as the foreground and all genes as the background for the gained differential transcripts. The P value of a specific branch in the same GO/pathway classification of the foreground transcripts was calculated through hypergeometric distribution (phyper function) for the FDR calibration.

2.6. Expression pattern analysis of differential genes

The probes with expression increase in SOM cluster analysis in RH8706-49 after salt stress were chosen. The EST was searched by NCBI according to the corresponding serial probe number. Electronic cloning was conducted by these EST. Blastn and Blastx were then compared to obtain the full-length cDNA sequence and salt-stress related genes.

The total RNA in the root and the leaf was extracted from RH8706-34 and RH8706-49 with two leaves and one heart at 0, 1, 6, 12, and 72 h of 0.8% NaCl. This RNA was inversely translated into cDNA for PCR analysis on Rotor-Gene 3000 qualitative PCR instrument (Gene Company Limited). The experimental results were analyzed by using RG3000 6.0 (Gene Company Limited). The wheat β-actin gene (GenBank accession no. AB181991) was used as an internal reference, whereas SYBR Green was used as dye. The Delta Delta CT quantitative method was subsequently implemented. The process was repeated thrice for each sample. Table 1 shows a list of the quantitative primers.

Please cite this article in press as: X. Ma, et al., Analysis on the transcriptome information of two different wheat mutants and identification of salt-induced differential genes, Biochemical and Biophysical Research Communications (2016), http://dx.doi.org/10.1016/j.bbrc.2016.04.039