Journal of Hazardous Materials 252–253 (2013) 123–131

Contents lists available at SciVerse ScienceDirect

Journal of Hazardous Materials

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

Molecular dissection of mercury-responsive transcriptome and sense/antisense genes in *Medicago truncatula*

Zhao Sheng Zhou [b], Sheng Ning Yang [c], Hua Li [a], Can Can Zhu [a], Zhao Pu Liu [b], Zhi Min Yang [a,*]

[a] Department of Biochemistry and Molecular Biology, College of Life Science, Nanjing Agricultural University, Nanjing, China
[b] Jiangsu Province Key Laboratory of Marine Biology, College of Resources and Environmental Science, Nanjing Agricultural University, Nanjing, China
[c] Department of Biological, Chemical and Physical Sciences, Illinois Institute of Technology, Chicago, IL, USA

HIGHLIGHTS

- Genome-wide sequencing of Hg-exposed *M. truncatula* was conducted to analyze gene expression and complexity.
- Millions of sequenced tags were obtained from Hg-exposed libraries.
- Most of the annotated genes were expressed differentially under Hg stress.
- GO and KEGG analysis showed modified molecular functions and metabolic pathways.

ARTICLE INFO

Article history:
Received 7 September 2012
Received in revised form 4 February 2013
Accepted 11 February 2013
Available online 18 February 2013

Keywords:
Medicago truncatula
Transcriptome
Heavy metal
Deep sequencing
Sense and antisense gene

ABSTRACT

We described a newly developed approach, namely next-generation tag sequencing, to identify global gene transcripts and complexity regulated by heavy metals in *Medicago truncatula*. Two cDNA libraries were generated from *M. truncatula* seedlings: treated and non-treated with the toxic heavy metal mercury Hg(II). With the large number of read-mapped genes generated, we observed that most of the genes were differentially expressed between the two libraries. In addition, several classes of new transcripts including transcription factors, antisense transcripts, and stress responsive genes were detected. The forty genes most altered in expression levels were associated with tolerance to environmental stress and secondary metabolism. Validation of genes by quantitative RT-PCR confirmed the results from deep-sequencing. Most of genes coding for metal transporters, sulfate metabolism, and cell wall solidification were significantly altered by Hg exposure. We also examined altered expression ratios of sense and antisense (S–AS) transcripts between the two libraries. By analyzing strand-specific information of read sequences, S–AS transcripts were found to be enriched with metal treatment. The transcriptome sequences were analyzed further with Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) and showed diverse biological functions and metabolic pathways under the metal stress.

1. Introduction

Whereas some essential metals such as Cu and Zn are required for plant growth and development, several other non-essential metals such as Cd and Hg are hazardous to plant growth. These non-essential metals are naturally occurring but also enriched in the environment due to anthropogenic activities [1–3]. Among these is mercury (Hg), one of the most toxic heavy metals detected in environments [4–6]. As a unique metal, mercury exists in various oxidative forms, but in well-oxygenated soils the soluble Hg^{2+} is dominant [4,7]. Growing studies have shown that Hg is readily accumulated by higher plants [8–10]. As a result of Hg overexposure, affected plants often display abnormalities in their physiological processes [11].

Plants have developed diverse strategies to cope with external and internal fluctuations of metal availability and tend to minimize adverse effects of metals [12–17]. Heavy metal-regulated morphological alterations of metabolic pathways are best presented at cellular and physiological levels. Sequestration of absorbed toxic metals in vacuoles and other organelles can be achieved by metal transporters such as IRT2 and MTP1 [18–21], both of which play crucial roles in cytoplasmic detoxification [22]. However,

Abbreviations: GO, Gene Ontology; KEGG, Kyoto Encyclopedia of Genes and Genomes; PCs, phytochelatins; IRT2, iron-regulated transporter; MTP1, metal tolerance protein 1; SOAP, Simple Object Access Protocol; TIGR, The Institute of Genomic Research Database; MTGI, *Medicago truncatula* gene index; DEGs, differentially expressed genes; FDR, false discovery rate; TPM, transcripts per million.

* Corresponding author at: No. 1 Weigang, Outside the Zhongshan Men, Building of Life Science, College of Life Science, Nanjing Agricultural University, Nanjing 210095, China. Tel.: +86 25 84395057; fax: +86 25 84073096.
E-mail address: zmyang@njau.edu.cn (Z.M. Yang).

0304-3894/$ – see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.jhazmat.2013.02.011

understanding of heavy metal-responsive genes and regulating networks is limited. Recent search for heavy metal responsive genes have benefited from genome-wide sequencing of distinct tissues or entire plants. This approach can be used to explore further the biological processes of metal uptake, translocation and compartmentalization. Using microarray many genes from metal-exposed *Arabidopsis* has been profiled [23–25]. Also, hundreds of Cu-responsive genes in rice leaves were detected using microarray and quantified by real-time PCR [26]. More importantly, a collection of metal-responsive microRNAs have been detected using microarray [27] and cloning method [28–30]. Recently, the next-generation sequencing technology has become a powerful tool to permit the concomitant sequencing of millions of signatures to the genome and identify specific and abundance of genes in a single tissue [31–33]. These outcomes highlight the applications of genome-wide analysis in obtaining a complete view of heavy metal-responsive gene expression in plants. However, transcriptome studies focusing on genes involved in plant response to heavy metals have not been described using the high-throughput sequencing approach. In this study, *Medicago truncatula* was used because it is a model legume species with completed genomic sequences. In addition, *M. truncatula* has contributed a great deal to our understanding of acclimation to various biotic and abiotic stresses [34]. Its unique traits make it an excellent species for studying the molecular and genetic complexity of metal tolerance [35]. As a result of the study, we annotated thousands of signatures matching predicted genes and quantified the transcript abundance in the species. Furthermore, we profiled the expression patterns of sense and antisense RNA and found changes between Hg-treated and control plants. This present study highlights the change in transcriptome and networks of gene expression in *M. truncatula* under the stress of heavy metals.

2. Materials and methods

2.1. Library construction and sequencing

Seeds of *M. truncatula* were surface-sterilized, rinsed and germinated. To eliminate the fungal colonization that attenuated plant responses to heavy metals, the germinating seedlings were hydroponically grown under the condition described previously [36]. Seedlings were treated with 0 μM (Hg-free, control) or 10 μM $HgCl_2$ (Hg-treated) and harvested after 6, 12, 24 and 48 h, respectively. Total RNA was isolated at each time-point for their respective treatments with the TRIzol purification system (Invitrogen). The 3′-tag digital gene expression libraries were prepared using the Illumina Gene Expression Sample Prep Kit based on the method by Zhang et al. [37]. Briefly, total RNA (6 μg) was used for mRNA capture with magnetic oligo (dT) beads. The first and second strand cDNA was synthesized and stranded bead-bound cDNA was subsequently digested with NlaIII. The 3′-cDNA fragments attached to the oligo (dT) beads were ligated to the Illumina GEX NlaIII Adapter 1, which contained a recognition site for the Endonuclease MmeI for cutting 17 bp downstream of the recognition site (CATG) to produce reads with adapter 1. After removing 3′ fragments with magnetic beads precipitation, an Illumina GEX adapter 2 was introduced at the site of MmeI cleavage. The resulting adapter-ligated cDNA reads were amplified using PCR-primers that were annealed to the adaptor ends for 15 cycles. The 85 base pair fragments were purified and recovered by 6% polyacrylamide Tris–borate–EDTA gel.

2.2. Data processing and digital read profiling

The raw sequence reads, on average 35 nucleotides, were filtered by the Illumina pipeline. The 3′ adaptor sequence was removed from raw sequences. Low quality reads including reads smaller than 21 nt, empty reads, and singletons were removed. The high quality sequences (21 nt) were first mapped to MTGI reference sequences (ftp://occams.dfci.harvard.edu/pub/bio/tgi/data/Medicago_truncatula/), and the remaining reads to the TIGR reference sequences (http://www.tigr.org) by SOAP. When the mapping events on both strands were monitored, the sense and antisense sequences were included in the data collection. We mapped all expressed reads onto a preprocessed database of 17 bases-long sequences of MTGI and TIGR and only one mismatch was allowed. Reads mapped to more than one transcript were excluded from our analysis.

2.3. Analysis of GO and KEGG pathway

All read-mapped genes were identified by Blastx searching against the GO protein database (http://www.geneontology.org/) and KEGG databases (http://www.genome.jp/kegg/pathway.html). The GO enrichment analysis of functional significance was subjected to the ultra-geometric test with Benjamini–Hochberg correction [38]. GO terms with corrected P-value < 0.05 were regarded as significant enrichment for the differentially expressed genes (DEGs) compared to the genome background. Pathways with Q value < 0.05 indicate significantly enriched in DEGs.

2.4. Quantitative RT-PCR analysis

The first strand cDNA was synthesized from 1.0 μg total RNA isolated above by Moloney Murine Leukemia Virus Reverse Transcriptase (Promega) using oligo (dT) primers [39]. The quantitative RT-PCR was performed on a MyiQ Single Color Real-time PCR system (Bio-Rad) in a final volume of 20 μL containing 2 μL of a 1/10 dilution of cDNA in water, 10 μL of the 2× SYBR Premix Ex Taq (TaKaRa) and 200 nM of forward and reverse primers (Table S1). The thermal cycling conditions were 35 cycles of 95 °C for 5 s for denaturation and 60 °C for 30 s for annealing and extension. PCR efficiency was determined by a series of 2-fold dilutions of cDNAs. The calculated efficiency of all primers was 0.9–1.0. Relative expression levels were normalized with the constitutive control *MtEF1α* [40] and presented as $2^{-\Delta CT}$ to simplify the presentation of data [41]. All experiments were repeated in triplicate.

2.5. Statistical evaluation

Statistical analysis was performed to identify differentially expressed genes between the libraries using a rigorous algorithm described previously [42]. The gene expression was normalized to transcripts per million clean reads. The statistical t-test was used to identify genes expressed between libraries. P-values were adjusted by the multiple testing procedures described by Benjamini and Yekutieli [38], by controlling false discovery rate (FDR). In this study, we used stringent value FDR < 0.001 and the absolute value of $\log_2$ ratio ≤ 1 as the threshold to judge the significant difference of gene expression. The correlation of the detected count numbers between parallel libraries was statistically assessed by calculation of Pearson correlation coefficient.

3. Results

3.1. Data generation and analysis

The total number of reads per library ranged from 3.5 to 3.6 million, and the number of reads entitled with distinct sequences ranged from 0.34 to 0.36 million (Table S2). To ensure that high quality data were analyzed in the Tag-seq libraries, we removed potentially erroneous reads using the Solexa automated