South African Journal of Botany 116 (2018) 200–206

Contents lists available at ScienceDirect

South African Journal of Botany

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

The transcriptome properties of reeds under cadmium stress in Liaohe Estuary wetland

Q. Sun [1], F. Su [1], T. Wang *, Z. Ding

College of Bioscience and Biotechnology, Shenyang Agricultural University, Shenyang 110866, People's Republic of China
Liaoning Shuangtai Estuary Wetland Ecosystem Research Station, Panjin 124000, People's Republic of China

ARTICLE INFO

Article history:
Received 17 October 2017
Received in revised form 8 February 2018
Accepted 20 March 2018
Available online xxxx

Edited by M Vaculík

Keywords:
Reed
Cadmium stress
Transcriptome
Beta-glucosidase
Phenylpyruvate tautomerase
Oxidoreductase
Photosynthesis

ABSTRACT

Reed contains a naturally-occurring gene resource base to resist cadmium stress. Transcriptome profiling is conducive to cloning the cadmium-tolerance genes in reed. Reed roots under cadmium stress were sampled, and the Illumina HiSeq™ 2500 transcriptome sequencing was performed. A total of 286,439 unigenes were obtained after *de novo* assembly. There were 22,304 up-regulating differentially expressed unigenes and 15,711 down-regulation. Two-level clustering formed an accurate heatmap to compare between expression fold-change classes and between treatment classes. The gene ontology terms were mainly enriched into biological regulation, metabolic process, response to stimulus, cell component, catalytic activity, proton-transporting two-sector ATPase and binding. Noticeably, the catalytic activity was mainly forwarded to oxidoreductase activity and antioxidant activity. There were obviously-enriched Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways, phenylpropanoid biosynthesis, phenylalanine metabolism, ribosome and photosynthesis. Beta-glucosidase EC: 3.2.1.21 was up-regulation in the phenylpropanoid biosynthesis pathway. Up-regulation of phenylpyruvate tautomerase EC: 5.3.2.1 catalyzes the production of 2-hydroxy-3-phenylpropanoate in the phenylalanine metabolism pathway. In the photosynthesis pathway, there was up-regulation of petN, petH, photosystems I/II, LHCA4 and LHCB3. The protein-protein interaction network showed that beta-glucosidase EC: 3.2.1.21 interacted with c129889_g1 and c115092_g3; up-regulation of photosystems II c111052 interacted with c93463_g1 and c121697_g3. Two up-regulating enzymes EC: 3.2.1.21 and EC: 5.3.2.1 may be the candidates to improve cadmium tolerance in reed. The photosynthesis with light-harvesting and electron transport plays important roles in reed cadmium tolerance. The interacting proteins c129889_g1, c115092_g3, c93463_g1 and c121697_g3 are putative to involve in cadmium tolerance. The current study is promising to provide the profile data for cloning these reed genes.

1. Introduction

Heavy metal cadmium (Cd) is an extremely toxic environmental pollutant classified as a human carcinogen [Group 1-according to International Agency for Research on Cancer] (Cancer, 1994); Group 2a-according to Environmental Protection Agency (EPA) (Cadmium, 1992); and 1B carcinogen classified by European Chemical Agency (Faroon et al., 2012). In plants, cadmium is a non-essential and toxic metal, rapidly taken up by roots and accumulated in various plant tissues which hamper the crop growth and productivity (Gill and Tuteja, 2011). Inside cells, cadmium ions act as a hydrogen peroxide generator. Hydrogen peroxide can also convert thiol groups on proteins into nonfunctional sulfones and is also capable of directly attacking nuclear DNA (Nordberg, 2010).

Cadmium-resistant (or-tolerant) plants are found to develop several mechanisms in response to cadmium stress, including metallothionein (Sigel et al., 2009), antioxidative enzymes (Wang et al., 2008; Cuypers et al., 2010; Lin and Aarts, 2012), cadmium ion transporter P1B-ATPase, natural resistance-associated macrophage proteins, cation diffusion facilitators, and ATP-binding Cassettes (EZG et al., 2013). Therefrom, phytoremediation is a low-cost effective management method for removing the cadmium pollution in soil and/or water source. To recycle the plant with accumulation of heavy metal cadmium can reduce the heavy metal content in the soil and water, thus improve the ecological environment.

Reed [*Phragmites australis* (Cav.) Trin. ex Steud.] is the major cadmium cleanser in wetland. It contains a naturally-occurring gene resource base in response to cadmium stress. Studying and cloning the reed genes that are involved in cadmium stress can help us to understand the cadmium-resistant mechanism in reed, and subsequently endow other plants with the cadmium-resistant features by transgenic manipulation.

* Corresponding author.
E-mail address: sufangli8@163.com (T. Wang).
[1] These authors contributed equally.

https://doi.org/10.1016/j.sajb.2018.03.012
0254-6299/

The current study was designed to perform Illumina the Illumina HiSeq™ 2500 transcriptome sequencing of reed under cadmium stress. Through bioinformatics methods, we analyzed the reed transcriptome data, and mined important reed genes and/or transcription factors that are associated with the cadmium-resistant (or-tolerant) metabolic pathways. The current study was promising to provide the profile data for cloning these genes and study their function in future.

2. Materials and methods

2.1. Materials

Reed (*Phragmites australis* (Cav) Trin. Ex Steud.) was sampled on July 28, 2015, in the Comprehensive Experimental Base of Shenyang Agricultural University. Reed was exposed to cadmium at a concentration of 5 mM for 30 days. Reeds in the cadmium-stress group and the non-cadmium tress (CK) group were sampled. Briefly, 5 g reed roots for each group were sheared, with three biological repeats for each sampling. Samples were rinsed with distilled water, coated with fin foil and frozen immediately in liquid nitrogen.

2.2. RNA extraction, sequencing, reads filtering, and de novo assembly

Total RNA was extracted using RNeasy Plant Mini Kit (Qiagen, Shanghai, China) under the guidance of manufacturer's protocols. RNA quality was assessed using a NanoDrop 2000 UV–Vis Spectrophotometer (Thermo scientific). The mRNA with poly (A) tail was separated from 20 μg total RNA using the Dynabeads® Oligo $(dT)_{25}$ according to the protocol. And then, sequencing libraries were constructed using NEBNext® Ultra™ RNA Library Prep Kit for Illumina® (NEB, USA). The dichromate-treated group and untreated group were sequenced using the Illumina HiSeq™ 2500 platform.

After trimming the adaptor sequences, the quality of raw reads was filtering by FastQC, and low-quality bases (Q-value < 20) at both ends were abandoned by a custom Perl script. And then, all the clean reads were de novel assembled by Trinity program (version: v2014-04-13) (Grabherr et al., 2011; Haas et al., 2013). The unigenes shorter than 300 bp were removed to avoid possible assembly errors and reduce redundancy. Finally, the clean high-quality reads were collected for the further study.

2.3. Functional annotation and gene ontology (GO) analysis

The assembled unigenes were blast against the NR protein database from NCBI using BLASTX with a cut off of E-value $\leq 1 \times 10^{-5}$. According the blast result, the unigenes were annotated to GO (Gene Ontology) with a cut off of E-value $\leq 1 \times 10^{-6}$ to characterize the unigenes to three generalized ontological vocabularies which also categorized into different functional subcategories. The unigenes were also mapped to KO (Kyoto Encyclopedia of Genes and Genomes (KEGG) Ortholog database) using the BLASTX algorithm with an E-value threshold of 10^{-5} to assign the unigenes to Enzyme commission (EC) number (Mao et al., 2005).

2.4. Differential expression analysis

The clean reads of each sample were mapping to the assembled unigenes by bowtie2 (Langmead and Salzberg, 2012). After that, the RPKM (Reads Per Kilobases per Millionreads) of unigenes in each sample were calculated to present the unigenes' expression levels by a custom Perl script. Differential expression analysis of two group was performed using edgeR (Robinson et al., 2010). Unigenes adjusted P-value <0.05 and log (fold change) >2 were assigned as differentially expressed genes (DEGs). DEGs were clustered at two levels, fold change clustering and hierarchical clustering, forming a clustering heatmap. DEGs GO enrichment was based on the standard hypergeometric distribution. As for KEGG enrichment analysis of the DEGs, the KOBAS software was recruited to test the statistical enrichment of DEGs in KEGG pathways (Xie et al., 2011). Rich factor was used to represents enrichment intensiveness, which means that the ratio of the DEGs number and the number of genes have been annotated in this pathway.

2.5. DEGs' protein-protein interaction networks

The DEGs were mapped to STRING database (version: 9.1) using BLAST program. STRING database provide the easy and intuitive interfaces for searching and browsing the protein interaction data, as well as for inspecting the underlying evidence for study protein-protein interaction (Franceschini et al., 2013). Based on the sequence alignment, the Known and predicted associations from over 1100 organisms are scored and integrated to visualize comprehensive protein networks. The visualization of the DEGs protein network was visualized using Cytoscape software (Smoot et al., 2011).

3. Results and discussion

3.1. Transcriptome sequencing and assembly

To our knowledge, the current study first reported the transcriptome-wide profiling of reed response to cadmium stress. To obtain a comprehensive overview of underlying molecular mechanisms cadmium ion tolerance influence on reed, two group of reed root cDNA libraries, CK group and MiT group (cadmium ion tolerance group), were constructed with three biologic repetition. The transcriptome sequencing was carried out using the Illumina HiSeq™ 2500 high-throughput sequencing platform. A total of 286, 217, 184 reads were generated from the sequencing. After trimming the adapter sequence and filtering low quality reads, 40, 840, 648, 37, 479, 934 and 43, 152, 744 clean reads were obtained from CK1 to CK3 respectively, and 33, 130, 804, 32, 338, 160 and 36, 435, 834 clean reads were acquired from MiT1 to MiT3 respectively (Table 1). A high-quality assembly was implemented using clean reads and Trinity software. A total of 286, 439 unigenes, in the range of 201–13,318 bp with a N50 of 1809 bp, were obtained after *de novo* assembly (Table 2, Fig. 1).

Table 1
Summary of transcriptome sequencing.

Sample	CK group			MiT group		
	CK1	CK2	CK3	MiT1	MiT2	MiT3
Raw reads	50,367,450	49,417,448	50,807,998	45,329,464	42,779,058	47,515,766
Clean reads	40,840,648	37,479,934	43,152,744	33,130,804	32,338,160.00	36,435,834
Clean bases	3,426,097,200	5,621,990,100	6,472,911,600	4,969,620,600	4,850,724,000	5,465,375,100
GC content (%)	53.72%	54.35%	52.31%	54.54%	54.72%	54.46%
≥Q20 (%)	98.84%	98.04%	98.63%	98.97%	98.99%	98.88%
≥Q30 (%)	88.95%	84.86%	87.40%	90.00%	90.10%	89.41%