Data in Brief

Transcriptomic analyses of nitrogen assimilation processes in a Chinese strain of *Aureococcus anophagefferens*

Li-Na Chen, Hong-Po Dong *, Kai-Xuan Huang, Hua-Long Wang, Song-Hui Lu *, Jing-Yi Cen, Yue-Lei Dong

Research Center for Harmful Algae and Marine Biology, Key Laboratory of Eutrophication and Red Tide Prevention of Guangdong Higher Education Institutes, Jinan University, Guangzhou 510632, China

A B S T R A C T

*Aureococcus anophagefferens* is a harmful alga that dominates plankton communities during brown tides in North America, Africa, and Asia. In order to figure out the processes of nitrogen assimilation in a Chinese strain of *A. anophagefferens*, RNA-seq technology was used to examine transcriptomic differences in *A. anophagefferens* that was grown on urea, nitrate, or a mixture of urea and nitrate, and that was under N-replete, limited and recovery conditions. We noted that transcripts upregulated by nitrate and N-limitation included those encoding proteins involved in amino acid, nucleotide and aminosugar transport, degradation of amides and cyanates, and nitrate assimilation pathway. The data suggest that *A. anophagefferens* possesses an ability to utilize a variety of dissolved organic nitrogen. Moreover, transcripts for synthesis of proteins, glutamate-derived amino acids, spermines and sterols were upregulated by urea. Transcripts encoding key enzymes that are involved in the ornithine–urea cycle (OUC) and TCA cycle were differentially regulated by urea and nitrogen concentration, which suggests that the OUC may be linked to the TCA cycle and involved in reallocation of intracellular carbon and nitrogen. These genes regulated by urea may be crucial for the rapid proliferation of *A. anophagefferens* when urea is provided as the N source. Here, we provide the experimental procedures and analytical processes in detail. The data set is deposited in GEO with the accession number GSE60576.

© 2015 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

1. Direct link to deposited data

http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE60576

2. Experimental design, materials and methods

2.1. Algal strain and culture conditions

*A. anophagefferens* was collected from the coastal water of Qinhuangdao in the Bohai Sea, China on June 20, 2012 at station X01 (119°37.911’ E, 39°54.111’ N). Station X01 was located in a region that was experiencing brown tide on that date. *A. anophagefferens* cells were isolated using capillary pipettes under an inverted microscope and subsequently cultures from a single cell were established.

The cultures were grown in artificial seawater medium, which was enriched with f/2 nutrients, at 18 °C on a 12:12 h light:dark cycle under cool-white fluorescence lights (100 μmol photon m⁻² s⁻¹).
2.2. Experimental design

For the N source experiment, cultures were harvested during the late exponential growth phase and then inoculated into 1/2 media with three different sources of N, including nitrate (882 μmol L\(^{-1}\) final concentration), urea (441 μmol L\(^{-1}\) final concentration), and nitrate + urea (Mixture N, 441 μmol L\(^{-1}\) nitrate, 220 μmol L\(^{-1}\) urea, final concentration). The cells were harvested at the onset of the stationary phase by centrifugation (8000 \(\times\) g for 10 min), covered with RNAlater solution (Sigma) and stored at \(-80^\circ\)C.

For nitrogen-limited and recovery experiments, cells from an exponential culture grown in 1/2 media with urea as the N source (882 μmol L\(^{-1}\) final concentration) were collected by centrifugation (6000 \(\times\) g for 5 min), washed once with nitrogen-free media and then inoculated in N-replete (441 μmol L\(^{-1}\) urea) and N-limited (20 μmol L\(^{-1}\) urea) media, respectively. The cells were harvested by centrifugation (3000 \(\times\) g for 10 min) in the stationary phase when nitrogen was depleted in the N-limited media. The remaining nitrogen-starved cultures were divided into two parts and then received addition of either 882 μmol L\(^{-1}\) NO\(_3\) or 441 μmol L\(^{-1}\) urea. RNA samples were collected at 24 h.

2.3. Total RNA extraction and Illumina sequencing

Total RNA was extracted from frozen cell pellets using TRIzol reagent (Invitrogen, CA, USA) according to the manufacturer’s instructions. RNA purity and concentration were determined by spectrophotometry. RNA integrity was assessed using Agilent 2100 Bioanalyzer. RNA-seq libraries were constructed following an Illumina gene expression sample preparation kit. Briefly, total RNA (5–10 mg) was treated with RNase-free DNase I. Poly(A) mRNA was isolated using oligo(dT) magnetic beads and then fragmented into short fragments (approximately 200 bp). The first-strand synthesis of cDNA was performed using random hexamer-primed reverse transcription. The second-strand synthesis was performed by adding the first strand cDNA synthesis reaction to a second strand reaction mix consisting of first strand buffer, second strand buffer, a dNTP mix, RNase H (Invitrogen) and DNA polymerase I (Invitrogen). The double stranded cDNA was subsequently purified using magnetic beads. End reparation and 39-end single nucleotide A addition was performed. Then, the cDNA fragments were connected with sequencing adaptors and were enriched by PCR amplification. Finally, the library was sequenced in BGI-tech (Shenzhen) using an Illumina HiSeq 2000 sequencer. RNA-seq raw data have been deposited in the NCBI Gene Expression Omnibus (GEO) database with experiment series accession number GSE60576.

2.4. Analysis of differentially expressed genes

To obtain high-quality reads, the raw reads were filtered to remove reads with adaptor sequence, low-quality reads and reads with high percentage of unknown bases using BGI-tech’s in-house software SOAPnuk. All processed clean reads were mapped to the reference genome and transcript of A. anophagefferens using the program SOAPaligner/SOAP2 (version 2.21) [1], respectively, which were downloaded from http://genome.jgi.doe.gov/Auran1/Auran1. This alignment allowed no more than two mismatches. In addition, considering that the A. anophagefferens genome just contains 1185 scaffolds and lacks prediction of gene models, the reads that mapped to the genome were not used for quantification analysis. In contrast, the reference transcript includes a total of 11,501 gene models built by homology to known proteins from other model organisms and ab initio gene predictions as well as from available A. anophagefferens EST and cDNA data, so the differential gene expression analysis was performed on the reads that mapped to the transcript. The number of clean reads for each gene was calculated and then normalized to RPKM (number of transcripts per million clean reads), which is related to the read number with gene expression levels [2]. Fold changes in the differential gene expression between conditions were calculated using the log2 ratio of RPKM. The significance of differentially expressed genes between two experimental groups (p-value) were performed following a published method [3]. A false discovery rate (FDR) ≤0.001 and an absolute value of log2 ratio ≥ 1 were used as cutoffs to judge the significance of gene expression differences.

No function annotation in the transcript which was used as reference genes is provided. In order to annotate these mapped genes, we performed a BLAST search against the non-redundant (NR) database in NCBI with an e-value cut-off of 1e\(^{-5}\). Those best hits with a specific function whose score is the highest and e-value > 1e\(^{-5}\) were chosen. The Blast2GO program was used to obtain the Gene ontology (GO) annotation of the genes.

3. Discussion

In this briefings, we show the experimental procedures and the transcriptomic analytical processes. A number of differentially expressed genes were found in A. anophagefferens grown on the different N resources, and under N-replete, limited and recovery conditions. This study helps us to better understand the strategies of nitrate and urea assimilation in a Chinese strain of A. anophagefferens. Our conclusions may provide a partial explanation for blooms of A. anophagefferens in estuaries with elevated levels of organic matter [4].

Conflict of interest

The authors have no conflicts of interest.

Acknowledgments

We thank BGI-tech (Shenzhen) for the sequencing and Y. Zhang for the bioinformatic analysis.

References

[1] R. Li, C. Yu, Y. Li, T.W. Lam, S.M. Yiu, et al., SOAP2: an improved ultrafast tool for short read alignment. Bioinformatics 25 (2009) 1966–1967.

[2] A.S. Morrissy, R.D. Morin, A. Delaney, T. Zeng, H. McDonald, et al., Next-generation tag sequencing for cancer gene expression profiling. Genome Res. 19 (2009) 1823–1835.

[3] S. Chen, P. Yang, F. Jiang, Y. Wei, Z. Ma, et al., De novo analysis of transcriptome dynamics in the migratory locust during the development of phase traits. PLoS One 5 (2010) e15633.

[4] H.P. Dong, K.X. Huang, H.L. Wang, S.H. Lu, J.Y. Cen, Y.L. Dong, Understanding strategy of nitrate and urea assimilation in a Chinese strain of Aureococcus anophagefferens through RNA-Seq analysis. PLoS One 9 (2014) 1–15.