International Journal of Genomics and Data Mining

Li JJ, et al. Int J Genom Data Min 02: 121.

Research Article

DOI: 10.29011/2577-0616.000121

Transcriptomic Data of Utilization Processes for Nitrogen and Phosphorus in *Prorocentrum donghaiense*

Jing-Jing Li[1], Wei Xia[1], Hong-Po Dong[2*], Lin-Jian Ou[1*]

[1]Key Laboratory of Eutrophication and Red Tide Prevention of Guangdong Higher Education Institutes, College of Life Science, Jinan University, China

[2]School of Ocean and Meteorology, Guangdong Ocean University, China

***Corresponding authors:** Hong-Po Dong, Key Laboratory of Eutrophication and Red Tide Prevention of Guangdong Higher Education Institutes, College of Life Science, Jinan University, Guangzhou 510632, China. Tel: +862085220010; Email: donghongpo2001@hotmail.com

Lin-Jian Ou, School of Ocean and Meteorology, Guangdong Ocean University, Zhanjiang, Guangdong Province, China. Tel: +8607592396055; Email: torangeou@jnu.edu.cn

Citation: Li JJ, Xia W, Dong HP, Ou LJ (2018) Transcriptomic Data of Utilization Processes for Nitrogen and Phosphorus in *Prorocentrum donghaiense*. Int J Genom Data Min 02: 121. DOI: 10.29011/2577-0616.000121

Received Date: 15 December, 2017; **Accepted Date:** 30 December, 2017; **Published Date:** 08 January, 2018

Abstract

Prorocentrum donghaiense is one of the most frequently occurred harmful algal blooms in the East China sea. In order to reveal the response mechanisms of *P. donghaiense* to different nutrient status, *de novo* transcriptome sequencing was used to examine transcriptomic differences in *P. donghaiense* that was grown under replete, nitrogen-limited or phosphorus-limited conditions. We noted that transcripts down-regulated by phosphate limitation included those encoding proteins involved in RNA transport, oxidative phosphorylation, photosynthesis, endocytosis, pyrimidine metabolism, glycolysis/gluconeogenesis, biosynthesis of amino acids, vitamin digestion and absorption, protein processing in endoplasmic reticulum, while the expression of genes involved in ribosomal protein metabolism were significantly up-regulated. The abundance of 896 transcripts were elevated or inhibited by nitrogen limitation and they were involved in metabolic processes similar to P depletion. Here, we presented the experimental procedures and analytical processes in detail.

Keywords: Nitrogen Limitation; Phosphate Limitation; *Prorocentrum donghaiense*; Transcriptome

Specification	
Organism/cell line /tissue	*Prorocentrum. donghaiense*
Sex	N/A
Sequencer or array type	Illumina HiSeq4000
Data format	Raw data: FASTA file, analyzed data: text file
Experimental factor	*Prorocentrum. donghaiense* grown on nutrient replete, N-free or P-free medium
Experimental features	Comparative transcriptomic analyses among *P. donghaiense* cells grown on nutrient replete, N-free or P-free medium
Consent	N/A
Sample source location	Zhu Jiang, China

Direct Link to Deposited Data

https://www.ncbi.nlm.nih.gov/Traces/study/?acc=SRP125985

Experimental Design, Materials and Methods

Algal strain and culture conditions

Prorocentrum donghaiense (MEL203) was originally isolated from Zhu Jiang, China on 2009, and preserved in research center of harmful algae and marine biology in Jinan University. Before formal experiments, P. *donghaiense* was revived in natural seawater media supplemented with f/2 nutrients. Subsequently, the cultures were inoculated into f/2 media made from artificial seawater and incubated in a light-dark cycle of 12 h: 12 h with an intensity of 100 μmol/(m^2·s) provided by cool fluorescent tubes and 20 ± 1°C. The two antibiotics were added into the media to inhibit bacterial growth (final concentration: penicillin G, 30 mg L^{-1}; streptomycin sulphate, 50 mg L^{-1}).

Experimental Design

For the nitrogen and phosphate limitation experiment, cultures during exponential growth phase were centrifuged at 1400 g for 5 min, and the resultant pellets were transferred into nutrient-replete (882 μmol L^{-1} NO_3^- and 36 μmol L^{-1} PO_4^{3-}), N-free (0 μmol L^{-1} NO_3^- and 36 μmol L^{-1} PO_4^{3-}), and P-free media (882 μmol L^{-1} NO_3^- and 0 μmol L^{-1} PO_4^{3-}), respectively. The cultures were maintained at 20 ± 1°C in a light-dark cycle of 12 h: 12 h. Cultures were harvested at 12 h after inoculation, and pellets were covered with RNA later solution(Sigma) and stored at -80°C for further analysis.

Total RNA Extraction and Library Preparation

Total RNA was prepared from frozen cells using the total RNA extraction kit (Magen, shanghai, China). The extracted RNA was eluted in RNA-free water, and the concentration was determined using a spectrophotometer (Aglient technologies, CA, USA). The mRNA was purified from total RNA using poly-T oligo-attached magnetic beads. The purified mRNA was cut into fragments using divalent cations under high temperature. These RNA fragments were generated into first strand cDNA using random hexamer primer and RNase H. After that, the second strand of cDNA was subsequently synthesized using the first strand buffer, dNTPs, DNA polymerase I and RNase H. The cDNA fragments were purified with QiaQuick PCR kits and washed with EB buffer. And then, these fragments were terminally repaired, and poly(A)-tails and adapters were added. The aimed products were separated by agarose gel electrophoresis, and the fragments were PCR amplificated to create a cDNA library. The clustering of the index-coded samples was performed on a cBot cluster generation system using HiSeq PE Cluster Kit v4-cBot-HS (Illumina) and then the library preparations were sequenced using an illummina HiSeq 4000 sequencer and 150 bp paried-end reads were generated Raw data files have been deposited in the NCBI's Gene Expression Omnibus(GEO) (SRX3437735).

Quality Control, Assembly of Reads, Coding Region Prediction and Annotation

In order to get high-quality reads, raw data was processed with Perl scripts to get rid of reads with adaptor sequence, low-quality reads and reads with number of N accounting for more than 5%. High-quality reads were assembled by Trinity software (version 20140710) [1]. The clean data were mapped to the assembled transcript by Bowtie [2] to post assembly evalution [4]. TransDecoder (version 20140710) was used to recognize candidate coding regions of the assemblied reads. The functional annotation of unigenes and Open Reading Frames (ORFs) were achieved using Trinotate (version 20140717). Trinotate is a tool which is widely used for annotating *de novo* assembled transcriptomes. The software can carry out multiple functions, including homology search, the structure identification of the protein domains (HMMER/PFAM), protein signal prediction (SignalP/TmHMM), etc.

Analysis of Differentially Expressed Genes

Reads count for each gene in each sample was counted by HTSeq v0.6.0, RPKM (Reads Per Kilobase Millon Mapped Reads) was then used to quantitatively estimate gene expression values in each sample. The final set of the genes were used for differential gene expression analysis [3]. DEGseq was used to compare genes that were upregulated and downregulated between two samples using a model based on the negative binomial distribution. The P-value could be assigned to each gene and adjusted following the Benjamini and Hochberg's correction for controlling the false discovery rate. Genes with $q \leq 0.05$ and $|\log2\_ratio| \geq 1$ are identified as Differentially Expressed Genes (DEGs) [4].

Discussion

In summary, most of differently expressed genes were downregulated in N-or P-depleted conditions relative to N-replete conditions. Various metabolic process in cells were regulated by N- or P- stress. These differently expressed genes will lay the foundation for in-depth study on mechanisms of nutrient utilization of *P. donghaiense*.

Conflict of interest

The authors have no conflict of interest

Acknowledgment

This study was funded by the Guangdong Natural Science Foundation (2015A030313326, S2013010013406), the National Natural Science Foundation of China (41576123, U1301235, 41206126), the Strategic Priority Research Program of the Chinese Academy of Sciences (XDA11020304), and the International Science and Technology cooperation project (GASI-IPOVI-04).

References

1. Manfred G Grabherr (2011) Full-Length Transcriptome assembly from RNA-Seq Data without a reference genome. Nature Biotechnology 29: 644-652.
2. Langmead B, Salzberg SL (2012) Fast gapped-read alignment with Bowtie 2. Nature Methods 9: 357-359.
3. Mortazavi A, Williams BA, McCue K, Schaeffer L, Wold B (2008) Mapping and quantifying mammalian Transcriptomes by RNA-seq. Nat Methods 5: 621-628.
4. M I Love, W Huber, S Anders (2014) Moderated estimation off old change and dispersion for RNA-seq data with DESeq2. Genome Biology 15: 550.