DYEatom Metatranscriptome metadata from RV/Point Sur cruise PS1312 in the Monterey Bay area, June-July 2013

Website: https://www.bco-dmo.org/dataset/768550
Data Type: Cruise Results
Version: 1
Version Date: 2019-05-29

Project
» Linking physiological and molecular aspects of diatom silicification in field populations (Diatom Silicification)

Contributors	Affiliation	Role
Thamatrakoln, Kimberlee	Rutgers University (Rutgers IMCS)	Principal Investigator
Allen, Andrew E.	J. Craig Venter Institute (JCVI)	Co-Principal Investigator
Copley, Nancy	Woods Hole Oceanographic Institution (WHOI BCO-DMO)	BCO-DMO Data Manager

Abstract

Metadata for assembled contigs and ORFS from metatranscriptome analysis from CTD casts in the Monterey Bay area on RV/Point Sur cruise PS1312, June-July 2013. Assembled contigs files are also available; see Supplemental Files.

Table of Contents

Coverage

Spatial Extent: **N**:38.265 **E**:-121.981 **S**:36.455 **W**:-123.969
Temporal Extent: 2013-06-28 - 2013-07-05

Dataset Description

Metadata for assembled contigs and ORFS from metatranscriptome analysis from CTD casts in the Monterey Bay area on RV/Point Sur cruise PS1312, June-July 2013.

Methods & Sampling

Water was collected using Niskin bottles mounted on a CTD rosette. Biomass for metatranscriptomic analysis was collected by filtration (after a 200 micron pre-filtration) onto 47 mm, 1.2-micron pore size polycarbonate filters at <5 psi for no longer than 15 min to minimize degradation. Filters were flash frozen in liquid nitrogen and stored at -80 degrees C. Upon analysis, filters were thawed and RNA was extracted using TRIzol reagent according to the manufacturer's protocol (Life Technologies). Metatranscriptome libraries were constructed

using 500 ng of total RNA and a TruSeq RNA Sample Preparation Kit (Illumina; San Diego, CA) following the Low-Throughput protocol. The mean size of the final libraries was confirmed to be between 359-420 base pairs (bp) using an Agilent Bioanalyzer 2100 (Santa Clara, CA). Libraries were paired-end sequenced (2x150 bp) on the Illumina HiSeq platform. ORFs were annotated via BLASTP alignment (e-value > 10-3) to a comprehensive protein database, phyloDB, as well as screened for function de-novo by assigning Pfams, TIGRfams and transmembrane tmHMMs with hmmer 3.0 (http://hmmer.org/). PhyloDB 1.076 consists of 24,509,327 peptides from 19,962 viral, 230 archaeal, 4,910 bacterial, and 894 eukaryotic taxa. It includes peptides at KEGG, GenBank, JGI, ENSEMBL, CAMERA, and various other repositories, as well as from the 410 taxa of the Marine Microbial Eukaryotic Transcriptome Sequencing Project. Taxonomic annotation of ORFs was also conducted via BLASTP to phyloDB.

All cruise related data are available publicly at the Biological and Chemical Oceanography Data Management Office under project number 550825 (https://www.bco-dmo.org/project/550825). The metatranscriptomic data have been deposited in the NCBI sequence read archive (BioProject accession no. PRJNA528986: BioSample accession nos. SAMN11263616 - SAMN11263639 and SAMN11258802-SAMN11258825). Assembled contigs used in this study can also be found at https://scripps.ucsd.edu/labs/aallen/data/.

Unassembled reads and rRNA data from this study: https://www.ncbi.nlm.nih.gov/bioproject/PRJNA528986
Assembled contigs: See Supplemental Files.

Data Processing Description

BCO-DMO Processing Notes:
- created flat table from info at GenBank and metadata submitted by PI
- added conventional header with dataset name, PI name, version date
- modified parameter names to conform with BCO-DMO naming conventions
- added cruise_id

[table of contents | back to top]

Data Files

File
DYEatom_Metatranscriptome.csv(Comma Separated Values (.csv), 11.13 KB) MD5:98e0291d6bd3f23b229286837252b328 Primary data file for dataset ID 768550

[table of contents | back to top]

Supplemental Files

File
Assembled contigs fasta file 1 filename: assembly.ctgs.fa (Octet Stream, 112.36 MB) MD5:493c34a721f4cd26221eb91bfbcdb1bd
Assembled contigs fasta file 2 filename: assembly.orf.faa (Octet Stream, 33.63 MB) MD5:a48bdbc28da74057c346a8db973e7518
Assembled contigs fasta file 3 filename: assembly.orf.ffn (Octet Stream, 90.77 MB) MD5:182af87c999986c62ddaeb8b2d5394e3

[table of contents | back to top]

Related Publications

Kranzler, C. F., Krause, J. W., Brzezinski, M. A., Edwards, B. R., Biggs, W. P., Maniscalco, M., … Thamatrakoln, K. (2019). Silicon limitation facilitates virus infection and mortality of marine diatoms. Nature Microbiology. doi:10.1038/s41564-019-0502-x
Results

[| back to top]

Parameters

Parameter	Description	Units
BioProject_type	NCBI BioProject type	unitless
BioProject_id	NCBI BioProject identifier	unitless
BioSample	NCBI BioSample identifier	unitless
Sample_name	NCBI Sample identifier	unitless
SRA_id	NCBI SRA identifier	unitless
Package_type	NCBI Package type	unitless
version	NCBI version	unitless
Accession	NCBI Accession	unitless
ID	NCBI ID	unitless
cruise_id	cruise identifier	unitless
CTD_cast	CTD cast number	unitless
Lat	latitude; north is positive	decimal degrees
Long	longitude; east is positive	decimal degrees
Date_collection	date of collection; formatted as yyyy-mm-dd	unitless
station	station number	unitless
depth_m	depth of sample	meters

[table of contents | back to top]

Instruments

Dataset-specific Instrument Name	Illumina HiSeq platform
Generic Instrument Name	Automated DNA Sequencer
Generic Instrument Description	General term for a laboratory instrument used for deciphering the order of bases in a strand of DNA. Sanger sequencers detect fluorescence from different dyes that are used to identify the A, C, G, and T extension reactions. Contemporary or Pyrosequencer methods are based on detecting the activity of DNA polymerase (a DNA synthesizing enzyme) with another chemoluminescent enzyme. Essentially, the method allows sequencing of a single strand of DNA by synthesizing the complementary strand along it, one base pair at a time, and detecting which base was actually added at each step.

[table of contents | back to top]

Deployments

PS1312

Website	https://www.bco-dmo.org/deployment/701341
Platform	R/V Point Sur
Start Date	2013-06-27
End Date	2013-07-06
Description	Cruise DOI: 10.7284/903425

Project Information

Linking physiological and molecular aspects of diatom silicification in field populations (Diatom Silicification)

Coverage: Oregon/California Coastal Upwelling Zone, between 34-44N and 120-124W

Description from NSF award abstract:
Diatoms, unicellular, eukaryotic photoautotrophs, are among the most ecologically successful and functionally diverse organisms in the ocean. In addition to contributing one-fifth of total global primary productivity, diatoms are also the largest group of silicifying organisms in the ocean. Thus, diatoms form a critical link between the carbon and silicon (Si) cycles. The goal of this project is to understand the molecular regulation of silicification processes in natural diatom populations to better understand the processes controlling diatom productivity in the sea. Through culture studies and two research cruises, this research will couple classical measurements of silicon uptake and silica production with molecular and biochemical analyses of Silicification-Related Gene (SiRG) and protein expression. The proposed cruise track off the West Coast of the US will target gradients in Si and iron (Fe) concentrations with the following goals: 1) Characterize the expression pattern of SiRGs, 2) Correlate SiRG expression patterns to Si concentrations, silicon uptake kinetics, and silica production rates, 3) Develop a method to normalize uptake kinetics and silica production to SiRG expression levels as a more accurate measure of diatom activity and growth, 4) Characterize the diel periodicity of silica production and SiRG expression.

It is estimated that diatoms process 240 Teramoles of biogenic silica each year and that each molecule of silicon is cycled through a diatom 39 times before being exported to the deep ocean. Decades of oceanographic and field research have provided detailed insight into the dynamics of silicon uptake and silica production in natural populations, but a molecular understanding of the factors that influence silicification processes is required for further understanding the regulation of silicon and carbon fluxes in the ocean. Characterizing the genetic potential for silicification will provide new information on the factors that regulate the distribution of diatoms and influence in situ rates of silicon uptake and silica production. This research is expected to provide significant information about the molecular regulation of silicification in natural populations and the physiological basis of Si limitation in the sea.

Funding

Funding Source	Award
NSF Division of Ocean Sciences (NSF OCE)	OCE-1333929