Reference sequences, genes, and K0 numbers for sampled diatoms on the Laurence Gould (LMG1411) in the Western Antarctica Peninsula during 2014. (Polar Transcriptomes project)

Website: https://www.bco-dmo.org/dataset/661492
Data Type: Cruise Results
Version: 1.1
Version Date: 2025-02-25

Project
» Iron and Light Limitation in Ecologically Important Polar Diatoms: Comparative Transcriptomics and Development of Molecular Indicators (Polar_Transcriptomes)

Contributors	Affiliation	Role
Marchetti, Adrian	University of North Carolina at Chapel Hill (UNC-Chapel Hill)	Principal Investigator, Contact
Ake, Hannah	Woods Hole Oceanographic Institution (WHOI BCO-DMO)	BCO-DMO Data Manager

Abstract
Reference sequences, genes, and K0 numbers for sampled diatoms on the Laurence Gould (LMG1411) in the Western Antarctica Peninsula during 2014. (Polar Transcriptomes project)

Table of Contents

Coverage

Temporal Extent: 2014 - 2014

Dataset Description

Reference sequences, genes, and K0 numbers for sampled diatoms on the Laurence Gould (LMG1411) in the Western Antarctica Peninsula during 2014.

Diatom isolates were obtained from the Western Antarctic Peninsula surface waters.

Methods & Sampling

Nine species of diatoms were isolated from the Western Antarctic Peninsula along the PalmerLTER sampling grid in 2013 and 2014. Isolations were performed using an Olympus CKX41 inverted microscope by single cell isolation with a micropipette (Anderson 2005). Diatom species were identified by morphological characterization and 18S rRNA gene (rDNA) sequencing. DNA was extracted with the DNeasy Plant Mini Kit according to the manufacturer's protocols (Qiagen). Amplification of the nuclear 18S rDNA region was achieved with standard PCR protocols using eukaryotic-specific, universal 18S forward and reverse primers.

Primer sequences were obtained from Medlin et al. (1982). The length of the region amplified is approximately 1800 base pairs (bp). Pseudo-nitzschia species are often difficult to identify by their 18S rDNA sequence, therefore, additional support of the taxonomic identification of P. subcurvata was provided through sequencing of the 18S-ITS1-5.8S regions. Amplification of this region was performed with the 18SF-euk and 5.8SR_euk primers of Hubbard et al. (2008). PCR products were purified using either QIAquick PCR Purification Kit (Qiagen) or ExoSAP-IT (Affymetrix) and sequenced by Sanger DNA sequencing (Genewiz). Sequences were edited using Geneious Pro software (http://www.geneious.com, Kearse et al., 2012) and BLASTn sequence homology searches were performed against the NCBI nucleotide non-redundant (nr) database to determine species with a cutoff identity of 98%.

BUSCO (Benchmarking Universal Single-Copy Orthologs) was used to assess the completeness of genomes and transcriptomes based on sets of single copy orthologous groups derived from OrthoDB that are highly conserved within multiple lineages (Felipe et al. 2015). Completed, duplicated and fragmented orthologs were determined by meeting an 'expected score' and having aligned sequences within two standard deviations of the BUSCO gene's length. A second metric of completeness was performed by evaluating conserved pathways, such as the ribosome and spliceosome, using the single-directional best-hit method in the KEGG Automatic Annotation Server (KAAS) (Moriya et al. 2007). Finally contiguity, was calculated at the 0.75 level as according to Martin and Wang (2011) with custom scripts.

For each transcriptome, unassembled sequence reads were aligned to the final Trinity assembly using Bowtie 2 (Langmead 2012). Mapped reads were normalized by the Reads per Kilobase per Million reads method (RPKM) (Mortazavi et al. 2008).

Gene biogeographical distributions - 20 genes of interest were selected in the study to investigate the molecular basis of iron and light limitation in polar diatoms. Reference sequences for each of these genes were obtained from the F. cylindrus and P. tricornutum JGI genome portals and T. pseudonana and T. oceanica NCBI and GenBank repositories. Reference sequences were identified in the transcriptomes by translated nucleotide homology searches (tBLASTn) with an e-value cutoff of <10-5. A reciprocal tBLASTn homology search was performed for each transcriptome against the KEGG GENES database, using the single-directional best-hit method in the KAAS online tool to ensure consistent gene annotations (Moriya et al. 2007).

Subsequently, reference sequences were identified in the MMETSP protein database by BLASTp (e-value <10-5) homology searches among the diatom transcriptomes. The transcriptomes and their associated latitude and longitude were obtained from iMicrobe Data Commons (Project Code CAM_P_0001000) and the National Center for Marine Algae and Microbiota (NCMA). Custom Matlab scripts allowed global biogeographical distribution of key genes of interest to be mapped.

Data Processing Description

BCO-DMO Data Processing Notes:

-reformatted column names to comply with BCO-DMO standards
-added a column to accession number links
-removed all special characters and commas
-replaced spaces in gene and diatom columns with underscores

[table of contents | back to top]

Data Files

File
reference_sequences.csv(Comma Separated Values (.csv), 11.18 KB) MD5:cadb079b403f5e8f2f8e562269754a3c

[table of contents | back to top]

Parameters

Parameter	Description	Units
diatom	Reference diatom species	unitless
gene	Gene name	unitless
gene_shortName	Ancronym for gene name	unitless
source	Database source	unitless
accession_number	Molecular level functions in KEGG; NCBI accession number	unitless
accession_link_url	Link to NCBI	unitless
accession_link_text	NCBI Acession ID	unitless
accession_link	NBII Acession link	unitless
K0_number	Molecular level functions in KEGG; NCBI accession number	unitless
K0_accession_link_url	Link to K0	unitless
K0_accession_link_text	K0 Accession ID	unitless
K0_accession_link	Accession link for K0 number	unitless
K0_description	Molecular function	unitless

[table of contents | back to top]

Instruments

Dataset-specific Instrument Name	Agilent Bioanalyzer 2100
Generic Instrument Name	Bioanalyzer
Dataset-specific Description	Used to determine RNA integrity
Generic Instrument Description	A Bioanalyzer is a laboratory instrument that provides the sizing and quantification of DNA, RNA, and proteins. One example is the Agilent Bioanalyzer 2100.

Dataset-specific Instrument Name	Olympus CKX41
Generic Instrument Name	Inverted Microscope
Dataset-specific Description	Used to perform isolations
Generic Instrument Description	An inverted microscope is a microscope with its light source and condenser on the top, above the stage pointing down, while the objectives and turret are below the stage pointing up. It was invented in 1850 by J. Lawrence Smith, a faculty member of Tulane University (then named the Medical College of Louisiana). Inverted microscopes are useful for observing living cells or organisms at the bottom of a large container (e.g. a tissue culture flask) under more natural conditions than on a glass slide, as is the case with a conventional microscope. Inverted microscopes are also used in micromanipulation applications where space above the specimen is required for manipulator mechanisms and the microtools they hold, and in metallurgical applications where polished samples can be placed on top of the stage and viewed from underneath using reflecting objectives. The stage on an inverted microscope is usually fixed, and focus is adjusted by moving the objective lens along a vertical axis to bring it closer to or further from the specimen. The focus mechanism typically has a dual concentric knob for coarse and fine adjustment. Depending on the size of the microscope, four to six objective lenses of different magnifications may be fitted to a rotating turret known as a nosepiece. These microscopes may also be fitted with accessories for fitting still and video cameras, fluorescence illumination, confocal scanning and many other applications.

[table of contents | back to top]

Deployments

LMG1401

Website	https://www.bco-dmo.org/deployment/675566
Platform	ARSV Laurence M. Gould
Start Date	2014-11-27
End Date	2014-12-21

[table of contents | back to top]

Project Information

Iron and Light Limitation in Ecologically Important Polar Diatoms: Comparative Transcriptomics and Development of Molecular Indicators (Polar_Transcriptomes)

Website: http://www.nsf.gov/awardsearch/showAward?AWD_ID=1341479

Coverage: Antarctica

The Southern Ocean surrounding Antarctica is changing rapidly in response to Earth's warming climate. These changes will undoubtedly influence communities of primary producers (the organisms at the base of the food chain, particularly plant-like organisms using sunlight for energy) by altering conditions that influence their growth and composition. Because primary producers such as phytoplankton play an important role in global biogeochemical cycling, it is essential to understand how they will respond to changes in their environment. The growth of phytoplankton in certain regions of the Southern Ocean is constrained by steep gradients in chemical and physical properties that vary in both space and time. Light and iron have been identified as key variables influencing phytoplankton abundance and distribution within Antarctic waters. Microscopic algae known as diatoms are dominant members of the phytoplankton and sea ice communities, accounting for significant proportions of primary production. The overall objective of this project is to identify the molecular bases for the physiological responses of polar diatoms to varying light and iron conditions. The project should provide a means of evaluating the extent these factors regulate diatom growth and influence net community productivity in Antarctic waters. The project will also further the NSF goals of making scientific discoveries available to the general public and of training new generations of scientists. It will facilitate the teaching and learning of polar-related topics by translating the research objectives into readily accessible educational materials for middle-school students. This project will also provide funding to enable a graduate student and several undergraduate students to be trained in the techniques and perspectives of modern biology.

Although numerous studies have investigated how polar diatoms are affected by varying light and iron, the cellular mechanisms leading to their distinct physiological responses remain unknown. Using comparative transcriptomics, the expression patterns of key genes and metabolic pathways in several ecologically important polar diatoms recently isolated from Antarctic waters and grown under varying iron and irradiance conditions will be examined. In addition, molecular indicators for iron and light limitation will be developed within these polar diatoms through the identification of iron- and light-responsive genes -- the expression patterns of which can be used to determine their physiological status. Upon verification in laboratory cultures, these indicators will be utilized by way of metatranscriptomic sequencing to examine iron and light limitation in natural diatom assemblages collected along environmental gradients in Western Antarctic Peninsula waters. In order to fully understand the role phytoplankton play in Southern Ocean biogeochemical cycles, dependable methods that provide a means of elucidating the physiological status of phytoplankton at any given time and location are essential.

[table of contents | back to top]

Funding

Funding Source	Award
NSF Office of Polar Programs (formerly NSF PLR) (NSF OPP)	PLR-1341479

[table of contents | back to top]