Diatom (Thalassiosira pseudonana) cell information from experiments designed to study single-cell transcriptional profiling of nutrient acquisition heterogeneity in diatoms conducted in December of 2022

Website: https://www.bco-dmo.org/dataset/918860
Data Type: experimental
Version: 1
Version Date: 2024-01-30

Project
» EAGER: Diatom Programmed Cell Death at Single-Cell Resolution (Diatom Death)

Program
» Ocean Carbon and Biogeochemistry (OCB)

Contributors	Affiliation	Role
Orellana, Monica V.	University of Washington Friday Harbor Laboratories (FHL)	Principal Investigator
Lausted, Christopher	Institute for Systems Biology (ISB)	Co-Principal Investigator
Huang, Sui	Institute for Systems Biology (ISB)	Scientist
York, Amber D.	Woods Hole Oceanographic Institution (WHOI BCO-DMO)	BCO-DMO Data Manager

Abstract
This dataset includes cell information for diatom Thalassiosira pseudonana grown during experiments conducted as part of a study of "Single-Cell transcriptional profiling of nutrient acquisition heterogeneity in diatoms." See "Related Datasets" section for T. pseudonana physiological data and gene information collected as part of the same study and experiments. Study description: Diatoms (Bacillariophyceae) are unicellular photosynthetic algae, accounting for about 40% of total marine primary production (equivalent to terrestrial rainforests) and critical ecological players in the contemporary ocean. Diatoms can form enormous blooms in the ocean that can be seen from space and are the base of food webs in coastal and upwelling systems, support essential fisheries, and are central to the biogeochemical cycling of important nutrients such as carbon and silicon. Over geological time, diatoms have influenced the world's climate by changing the carbon flux into the oceans. Diatoms have traditionally been studied on a population level. Growth is often measured by the total increase in biomass, and gene expression is analyzed by isolating mRNA from thousands or millions of cells. These methods generate a valuable analysis on the population’s average functioning; however, they fail to show how each individual diatom cell contributes to the population phenotype. Bulk transcriptomes confound different stages and variability of cell states in heterogeneous populations. By contrast, single-cell transcriptomics measures gene expression in thousands of individual diatoms providing a quantitative and ultrahigh-resolution picture of transient cell states in population fractions enabling the reconstruction of the various phenotypic trajectories. Thus, the single-cell physiological and molecular parameters analysis allows an unsupervised assessment of cell heterogeneity within a population—a new dimension in diatoms and phytoplankton in general. In this dataset, we examine the model diatom Thalassiosira pseudonana clonal cells grown in different nitrogen conditions, at the single cell level when grown in a light: dark cycle (12:12 h). Nitrogen is the major limiting nutrient for primary production and growth in the ocean’s surface, specifically for diatoms and the food webs they support. We investigate nutrient limitation, starvation and recovery. We used droplet-based, single-cell transcriptomics to analyze ten samples in two stages. In the first stage ("starvation"), six samples were collected over four days of culture as nutrient levels decreased. In the second stage ("recovery"), four samples were collected over twelve hours after nutrients were replenished.

Table of Contents

- Coverage
- Dataset Description
  - Methods & Sampling
  - Data Processing Description
  - BCO-DMO Curation Notes
- Data Files

- Related Publications
- Related Datasets
- Parameters
- Instruments
- Project Information
- Program Information
- Funding

Coverage

Location: Baliga Laboratory Institute for Systems Biology
Temporal Extent: 2022-12-04 - 2022-12-09

Methods & Sampling

Cultures: Axenic *Thalassiosira pseudonana* (CCMP 1335, National Center for Marine Algae and Microbiota, Maine, USA; LSID urn:lsid:marinespecies.org:taxname:148934) batch cultures were grown in enriched artificial seawater (ESAW) medium modified with reduced levels of nitrate (170 µM) to characterize starvation. Before the experiment, the diatoms were acclimated to a constant 20 °C temperature and under a 12: 12 h dark: light diurnal cycle at 300 µmol photons·m_{-2}·s_{-1} using cool fluorescent lights. The cultures were continuously equilibrated at ambient (420 ppm CO_2 by bubbling mixed gasses (air and CO_2 at 0.4 L/min) regulated using mass-flow controllers (GFC17, Aalborg) and monitored with a CO_2 analyzer (model Q-S151; Qubit Systems) into a 1.5-L glass bioreactor system. The bioreactors were inoculated with 150,000 cells/ml of acclimated, axenic *T. pseudonana* and grown for 5 days on a 12: 12 h dark: light cycle. The pH was monitored spectrophotometrically (Dickson et al. 2007); the photochemical yield of photosystem II (variable fluorescence/maximum fluorescence Fv/Fm) was measured with an AquaPen AP100; and total dissolved nitrogen was measure using the Nitrite/Nitrate, Colorimetric Test Roche # 11746081001) kit after syringe-filtering (0.2µm) the samples. Samples were taken twice a day, in the middle of dark-time, and the middle of the light-time representing different growth conditions regulating the metabolism based on light and nitrogen on days one and two (Ashworth et al. 2013). On days three and four, samples were only taken during the light time for experiment b. This experiment was repeated to evaluate the recovery of the cells from starvation by amending NO_3 (170 uM+/- 5uM) on day 5, after 70 hrs of starvation for *T. pseudonana* and sampled at T1: 0h, T2: 1h and T3: 3 hr, and T4: 6 hr after adding the supplemental nitrogen. Samples T1, T2, T3, were set on ice in the dark before analysis at when T4 was sampled.

Single cell transcriptomics: we used 10x Genomics Chromium single-cell 3' gene expression protocols to profile samples of 1000-10,000 diatoms. We used the standard throughput kits (v3.0, cat. #1000094) for exploration, targeting 10K cells/sample for measurements in a and c, and low throughput (LT) kits (v3.1, cat. #1000325) for a time series targeting 1000 cells/sample in experiment b. Fresh cells were harvested, washed with 1 × Phosphate buffered Saline (PBS) in RNase free water, at pH 7.4 and resuspended at 1 × 106 cells per ml in 1x PBS and 0.10% bovine serum albumin (x). Cellular suspensions were loaded on a Chromium instrument to generate single-cell Gelbead-In-EMulsion (GEM) droplets. Reverse transcription (RT) was performed in a C1000 Touch thermocycler (Biorad, Hercules, CA). After RT, GEMs were harvested and the cDNAs were amplified and cleaned with SPRIselect Reagent Kit (Beckman Coulter, Brea, CA). Indexed sequencing libraries were constructed using the Chromium Single-Cell 3’ Library Kit for enzymatic fragmentation, end-repair, A-tailing, adapter ligation, ligation cleanup, sample index PCR, and PCR cleanup. The barcoded sequencing libraries were quantified by quantitative PCR using the KAPA Library Quantification Kit (KAPA Biosystems, Wilmington, MA). Sequencing libraries were loaded on a NextSeq500 (Illumina, San Diego, CA) and run 150 cycles (26 bp for Read 1 and 124 bp for Read 2).

Experiment Metadata (physiology, gene and cell information):

Experiment
Start_Date 12-4
Light:Dark 12:12
Lights_On 1:00 PM
Lights_(umols/m.s) 1000 umoles m-2s-1
CO2_ppm ~420
scRNAseq_Timepoints 18 hours, 30 hours, 42 hours, 54 hours, 68 hours, 80 hours, 92 hours
scRNAseq_Timepoint_Labels
Notes scRNA, Full ESAW nitrate limited (170uM). Starting innoculate = 150k/m

Data Processing Description

Sequencing reads were aligned to the Diatom Consortium T. pseudonana CCMP1335 genome (GCA_000149405.2) and quantified using the Cell Ranger (v.3.0). Filtering, normalization and clustering was performed using Scanpy (Wolf et al., 2018). Cells were filtered to include only droplets containing at least 50 genes. Genes were then filtered to include only those appearing in at least 20 cells. Unique transcripts were identified by 10x unique molecular indexes (UMIs). UMI counts were normalized to 10,000 per cell. Gene expression was log transformed and scaled to unit variance and zero mean. Cells were clustered using the Leiden community detection clustering algorithm with the default settings.

BCO-DMO Curation Notes

* Tables within files "starve_obs_cell_info.csv" and "recover_obs_cell_info.csv" were imported into the BCO-DMO data system.
** Missing data values are displayed differently based on the file format you download. They are blank in csv files, "NaN" in MatLab files, etc.

* tables were combined with an additional "stage" column added with values "starve" or "recover"

[table of contents | back to top]

Data Files

File	
918860_v1_cell-information.csv	(Comma Separated Values (.csv), 405.81 KB) MD5:6d993b9df7aa578af1b7643d6ce43c8e
Primary data file for dataset ID 918860, version 1	

[table of contents | back to top]

Related Publications

Dickson, A.G.; Sabine, C.L. and Christian, J.R. (eds) (2007) Guide to best practices for ocean CO2 measurement. Sidney, British Columbia, North Pacific Marine Science Organization, 191pp. (PICES Special Publication 3; IOCCP Report 8). DOI: https://doi.org/10.25607/OBP-1342
Methods

Wolf, F. A., Angerer, P., & Theis, F. J. (2018). SCANPY: large-scale single-cell gene expression data analysis. Genome Biology, 19(1). https://doi.org/10.1186/s13059-017-1382-0
Methods

[table of contents | back to top]

Related Datasets

IsRelatedTo

Orellana, M. V., Lausted, C., Huang, S. (2024) **Diatom (Thalassiosira pseudonana) gene information from experiments designed to study single-cell transcriptional profiling of nutrient acquisition heterogeneity in diatoms conducted in December of 2022.** Biological and Chemical Oceanography Data Management Office (BCO-DMO). (Version 1) Version Date 2024-01-30 doi:10.26008/1912/bco-dmo.918852.1 [view at BCO-DMO]
Relationship Description: Datasets were part of the same experiments conducted as part of a study of "Single-Cell transcriptional profiling of nutrient acquisition heterogeneity in diatoms."

Orellana, M. V., Lausted, C., Huang, S. (2024) **Diatom (Thalassiosira pseudonana) physiological data from experiments designed to study single-cell transcriptional profiling of nutrient acquisition heterogeneity in diatoms conducted in December of 2022.** Biological and Chemical Oceanography Data Management Office (BCO-DMO). (Version 1) Version Date 2024-01-30 doi:10.26008/1912/bco-dmo.918841.1 [view at BCO-DMO]
Relationship Description: Datasets were part of the same experiments conducted as part of a study of "Single-Cell transcriptional profiling of nutrient acquisition heterogeneity in diatoms."

[table of contents | back to top]

Parameters

Parameter	Description	Units
stage	Experimental stage ("starve" or "recover").	unitless
cell_barcode	Unique DNA sequence identifying each cell.	unitless
sample	Sample name indicating the time of collection.	unitless
n_genes	Number of genes detected in each cell.	unitless
n_counts	Number of unique transcripts counted in each cell.	unitless
cluster	Leiden cluster name.	unitless

[table of contents | back to top]

Instruments

Dataset-specific Instrument Name	Nextseq 500 DNA sequencer, Illumina (San Diego, CA)
Generic Instrument Name	Automated DNA Sequencer
Generic Instrument Description	A DNA sequencer is an instrument that determines the order of deoxynucleotides in deoxyribonucleic acid sequences.

Dataset-specific Instrument Name	Chromium Controller droplet generator, 10x Genomics (Pleasanton, CA)
Generic Instrument Name	Chromium Controller
Generic Instrument Description	The Chromium Controller by 10x Genomics uses advanced microfluidics to perform single-cell partitioning and barcoding. Powered by Next GEM technology, the Chromium Controller enables integrated analysis of single cells at massive scale. Chromium Single Cell products can capture molecular readouts of cell activity in multiple dimensions, including gene expression, cell surface proteins, immune clonotype, antigen specificity, and chromatin accessibility. (https://www.10xgenomics.com/instruments/chromium-controller).

Dataset-specific Instrument Name	CO2 analyzer (model Q-S151; Qubit Systems)
Generic Instrument Name	CO2 Analyzer
Generic Instrument Description	Measures atmospheric carbon dioxide (CO2) concentration.

Dataset-specific Instrument Name	C1000 Touch PCR Thermocycler, Biorad (Hercules, CA)
Generic Instrument Name	Thermal Cycler
Generic Instrument Description	A thermal cycler or "thermocycler" is a general term for a type of laboratory apparatus, commonly used for performing polymerase chain reaction (PCR), that is capable of repeatedly altering and maintaining specific temperatures for defined periods of time. The device has a thermal block with holes where tubes with the PCR reaction mixtures can be inserted. The cycler then raises and lowers the temperature of the block in discrete, pre-programmed steps. They can also be used to facilitate other temperature-sensitive reactions, including restriction enzyme digestion or rapid diagnostics. (adapted from http://serc.carleton.edu/microbelife/research_methods/genomics/pcr.html)

[table of contents | back to top]

Project Information

EAGER: Diatom Programmed Cell Death at Single-Cell Resolution (Diatom Death)

Coverage: Laboratory study

NSF Award Abstract:
Diatoms are important primary producers in sunlit oceans and lakes across the globe. They are key players in spring phytoplankton blooms, which in turn support food webs that include productive fisheries. In addition, these photosynthetic diatom cells are known to capture enormous amounts of carbon, which are exported to depth as blooms die off. In oceanic systems, this process effectively removes carbon from the atmosphere for thousands of years. However, the biological mechanisms that lead to carbon sequestration, the transfer of carbon from the atmosphere into the deep ocean, are not well understood. As diatom populations reach the end of the bloom cycle, individual cells start to deteriorate and undergo cell death. The collapse of diatom populations is regarded as a critical point (or tipping point), which can be predicted by understanding the genetic activity of the population. This project seeks to elucidate how environmental change influences diatom cell-death processes. A mechanistic understanding of these cellular processes will elucidate how climate change could alter carbon-removal from the atmosphere and affect ocean productivity. The knowledge and methods developed during this study are applicable across organisms from all domains of life. The broader

impacts of this project are focused on high school education and new ideas and approaches for three-dimensional learning opportunities in support of Next Generation Science Standards (NGSS). Specifically, researchers are working with high school educators and partners to teach concepts of systems approaches, tipping points, and carbon sequestration in the marine environment through educational modules.

The goal is to determine the structure of diatom populations during the transition of actively growing cells towards population collapse by identifying the point of commitment (tipping point) at the level of individual cells. The model system is the diatom, Thalassiosira pseudonana, which is widespread and a common bloom-forming species. The random decision process of individual diatom cells is being characterized by establishing a genome-wide gene expression space through the physical and biochemical characterization of the cells. Implementation of the project is focused on measuring the phenotypic responses in the diatom to environmental factors including severe stress. Transcriptomic analyses during the transition of cell proliferation towards culture collapse include bulk (RNA-Seq) and single-cell (scRNA-Seq) high-throughput sequencing. Using a systems biology approach, the genetic information obtained from the samples is incorporated into predictive models to identify genetic transitions that occur prior to population collapse. The systems approach can detect changes in transcriptomic state that precede a critical point in the cell death process leading to predictions of how diatoms respond to environmental change. This work opens new vistas for the elucidation of mechanistic pathways of diatom cell populations.

This award reflects NSF's statutory mission and has been deemed worthy of support through evaluation using the Foundation's intellectual merit and broader impacts review criteria.

[table of contents | back to top]

Program Information

Ocean Carbon and Biogeochemistry (OCB)

Website: http://us-ocb.org/

Coverage: Global

The Ocean Carbon and Biogeochemistry (OCB) program focuses on the ocean's role as a component of the global Earth system, bringing together research in geochemistry, ocean physics, and ecology that inform on and advance our understanding of ocean biogeochemistry. The overall program goals are to promote, plan, and coordinate collaborative, multidisciplinary research opportunities within the U.S. research community and with international partners. Important OCB-related activities currently include: the Ocean Carbon and Climate Change (OCCC) and the North American Carbon Program (NACP); U.S. contributions to IMBER, SOLAS, CARBOOCEAN; and numerous U.S. single-investigator and medium-size research projects funded by U.S. federal agencies including NASA, NOAA, and NSF.

The scientific mission of OCB is to study the evolving role of the ocean in the global carbon cycle, in the face of environmental variability and change through studies of marine biogeochemical cycles and associated ecosystems.

The overarching OCB science themes include improved understanding and prediction of: 1) oceanic uptake and release of atmospheric CO2 and other greenhouse gases and 2) environmental sensitivities of biogeochemical cycles, marine ecosystems, and interactions between the two.

The OCB Research Priorities (updated January 2012) include: ocean acidification; terrestrial/coastal carbon fluxes and exchanges; climate sensitivities of and change in ecosystem structure and associated impacts on biogeochemical cycles; mesopelagic ecological and biogeochemical interactions; benthic-pelagic feedbacks on biogeochemical cycles; ocean carbon uptake and storage; and expanding low-oxygen conditions in the coastal and open oceans.

[table of contents | back to top]

Funding

Funding Source	Award
NSF Division of Ocean Sciences (NSF OCE)	OCE-2029738

[table of contents | back to top]