Metatranscriptome sequence NCBI accessions from samples collected during the SIMCO 1 & 2 incubation experiments, July & Oct 2014

Website: https://www.bco-dmo.org/dataset/735723
Data Type: experimental
Version: 1
Version Date: 2018-05-10

Project
» High Resolution Linkages Between DOC Turnover and Bacterioplankton in a Coastal Ocean (SIMCO)

Contributors	Affiliation	Role
Moran, Mary Ann	University of Georgia (UGA)	Principal Investigator
Copley, Nancy	Woods Hole Oceanographic Institution (WHOI BCO-DMO)	BCO-DMO Data Manager

Abstract
This dataset includes metatranscriptome sequence accessions from samples collected during the SIMCO 1 & 2 incubation experiments, July & Oct 2014, with links to the NCBI SRA pages.

Table of Contents

Coverage

Temporal Extent: 2014-07-23 - 2014-10-15

Dataset Description

This dataset includes metatranscriptome sequence accessions and links to the NCBI SRA pages.

Methods & Sampling

Experimental design, methods, data processing, and results are further described in (Vorobev et al., 2018).

Acquisition Description:

Near-surface water samples were collected from coastal waters near Sapelo Island, Georgia (31° 25' 4.08" N, 81° 17' 43.26" W), ~6 km from the mouth of Doboy Sound during July 2014 and October 2014. Samples were collected in daylight between 7:30 and 18:30. Six 20 L carboys were filled with water and wrapped in black plastic. Three were processed immediately, while the remaining three were returned to Doboy Sound for a 24 h dark incubation before processing by an identical protocol. This experimental scheme was carried out twice during each sampling event, once at high tide and once at low tide. To process the samples, 3 L were passed through a 3 µm pore-size filter to remove eukaryotic cells (Capsule Pleated 3 µm Versapor Membrane; Pall Life

Sciences, Ann Arbor, MI, USA) and then a 0.22 µm pore-size filter (Supor polyethersulfone; Pall Life Sciences). Subsamples for cell counts were collected prior to each filtration. The 0.22 µm filters were placed in Whirl-Pak® plastic bags (Nasco, Fort Artkinson, WI, USA) and immediately flash-frozen in liquid N2. The 0.22 µm filtrate was used for analysis of dissolved organic matter.

RNA was extracted from two replicate Time 0 h (T0) and Time 24 h (T24) filters (Table 1). The 50 ml lysis tubes contained 7.5 mL CTAB Extraction Solution (Teknova, Hollister, CA), 7.5 ml phenol:chloroform:isoamyl alcohol solution (25:24:1, pH 6.7), 750 µl 10% SDS, 200 µl 1% proteinase-K, 3 g RNA PowerSoil beads (Mo-Bio, Carlsbad, CA), and internal standards (described below). Frozen filters were broken into pieces inside the Whirl-Pak bags using a rubber mallet and transferred to the lysis tubes. Tubes were vortexed for 10 min, centrifuged, and the aqueous phase was transferred to a new tube. The extraction was performed again with chloroform, followed by centrifugation. RNA was precipitated from the aqueous phase with isopropanol, washed with cold 70% ethanol, and dissolved in 100 µl nuclease-free water. RNA was purified using an RNeasy Kit (Qiagen, Valencia, CA) followed by two successive treatments with the Turbo DNA-free kit (Invitrogen, Carlsbad, CA) to completely remove residual DNA.

Ribosomal RNA (rRNA) was selectively removed using biotinylated rRNA probes prepared for bacterial and archaeal 16S and 23S rRNA and eukaryotic 18S and 28S rRNA with DNA from a sample collected simultaneously (Stewart et al., 2010). Probe-bound rRNA was removed via hybridization to streptavidin-coated magnetic beads (New England Biolabs, Ipswich, MA), and successful removal of rRNA was confirmed using a 2200 TapeStation (Agilent Technologies, Santa Clara, CA).

rRNA-depleted samples were linearly amplified using the MessageAmp II-Bacteria Kit (Applied Biosystems, Austin, TX), and amplified mRNA was converted into cDNA using the Superscript III First Strand synthesis system (Invitrogen, Carlsbad, CA) with random primers, followed by the NEBnext mRNA second strand synthesis module (New England Biolabs,). Synthesized cDNA was purified using the PureLink PCR Micro Kit (Invitrogen) followed by ethanol precipitation and resuspension in 70 µL TE buffer, and stored at -80 degrees C until library preparation. The cDNA was sheared to ~300 bp with an E210 ultrasonicator (Covaris, Woburn, MA) and TruSeq libraries (Illumina Inc., San Diego, CA) were constructed. Libraries were sequenced on the Illumina HiSeq2500 platform to obtain 250 bp single end reads.

Internal Standard Addition - During sample processing, known copy numbers of two artificial internal mRNA standards ~1000 nt in length were added to each sample prior to cell lysis (Satinsky et al., 2014); and dx.doi.org/10.17504/protocols.io.ffwbjpe). Standards were synthesized using custom templates that were transcribed in vitro to RNA (Satinsky et al., 2013). The number of internal standards recovered in the sequence library was quantified via BLAST homology searches.

Data Processing Description

Bioinformatic Processing - Reads from the libraries were compared to a custom database containing small and large subunit rRNAs (derived from the SILVA database; www.arb-silva.de) and the internal standard sequences (Gifford et al., 2011; Gifford et al., 2013) using BLASTn. Reads with a bit score >50 to either database were removed from further analysis. Hits to the internal standard sequences were tallied. The remaining potential protein encoding reads were annotated by a homology search using RAPSearch2 against MarineRef II (http://ssharma.marsci.uga.edu/Lab/MarineRef2/), an in-house database consisting of marine bacterial, archaeal, viral, and eukaryotic genome sequences, and eukaryotic cDNAs from the MMESTP project (Keeling et al., 2014) that have been annotated based on SFams (Sharpton et al., 2012), Pfam (https://pfam.xfam.org), NCBI Protein Clusters (https://www.ncbi.nlm.nih.gov/proteinclusters), TIGRFAMS (http://www.jcvi.org/cgi-bin/tigrfams/Listing.cgi), and PhyloDB (https://drive.google.com/drive/folders/0B-BsLZUMHrDQfldGeDRIUHNZMEREY0g3ekpEZFhrTDlQSjQtbm5heC1QX2V6TUxBeFlOejQ). One replicate T0 sample from July high tide collection had contaminating sequences in the final library and was eliminated from subsequent analyses.

BCO-DMO Processing:
- added conventional header with dataset name, PI name, version date
- column names reformatted to comply with BCO-DMO standards
- added 'SRA_link' column with links to NCBI pages

[table of contents | back to top]

Data Files

File
SIMCO_transcript_accessions.csv (Comma Separated Values (.csv), 1.86 KB) MD5:24263c3a1d18d3b9d5f8f8fb927f2478 Primary data file for dataset ID 735723

[table of contents | back to top]

Related Publications

Corilo, Y. (2015) PetroOrg Software; Florida State University: Tallahassee, FL, 2014.
Methods

Dittmar, T., Koch, B., Hertkorn, N., & Kattner, G. (2008). A simple and efficient method for the solid-phase extraction of dissolved organic matter (SPE-DOM) from seawater. Limnology and Oceanography: Methods, 6(6), 230–235. doi:10.4319/lom.2008.6.230
Methods

Gifford, S. M., Sharma, S., Booth, M., & Moran, M. A. (2013). Expression patterns reveal niche diversification in a marine microbial assemblage. The ISME Journal, 7(2), 281–298. doi:10.1038/ismej.2012.96
Methods

Gifford, S. M., Sharma, S., Rinta-Kanto, J. M., & Moran, M. A. (2011). Quantitative analysis of a deeply sequenced marine microbial metatranscriptome. The ISME Journal, 5(3), 461–472. doi:10.1038/ismej.2010.141
Methods

Keeling, P. J., Burki, F., Wilcox, H. M., Allam, B., Allen, E. E., Amaral-Zettler, L. A., … Bell, C. J. (2014). The Marine Microbial Eukaryote Transcriptome Sequencing Project (MMETSP): Illuminating the Functional Diversity of Eukaryotic Life in the Oceans through Transcriptome Sequencing. PLoS Biology, 12(6), e1001889. doi:10.1371/journal.pbio.1001889
Methods

M., B., M., S., C., B., Smith, C., & Ann, M. (2016). Preparation of Custom Synthesized RNA Transcript Standard v2. Protocols.io. doi:10.17504/protocols.io.ffwbjpe
Methods

Rossel, P. E., Vähätalo, A. V., Witt, M., & Dittmar, T. (2013). Molecular composition of dissolved organic matter from a wetland plant (Juncus effusus) after photochemical and microbial decomposition (1.25 yr): Common features with deep sea dissolved organic matter. Organic Geochemistry, 60, 62–71. doi:10.1016/j.orggeochem.2013.04.013
Methods

Satinsky, B. M., Crump, B. C., Smith, C. B., Sharma, S., Zielinski, B. L., Doherty, M., … Moran, M. A. (2014). Microspatial gene expression patterns in the Amazon River Plume. Proceedings of the National Academy of Sciences, 111(30), 11085–11090. doi:10.1073/pnas.1402782111
Methods

Satinsky, B. M., Gifford, S. M., Crump, B. C., & Moran, M. A. (2013). Use of Internal Standards for Quantitative Metatranscriptome and Metagenome Analysis. Microbial Metagenomics, Metatranscriptomics, and Metaproteomics, 237–250. doi:10.1016/b978-0-12-407863-5.00012-5 https://doi.org/10.1016/B978-0-12-407863-5.00012-5
Methods

Sharpton, T. J., Jospin, G., Wu, D., Langille, M. G., Pollard, K. S., & Eisen, J. A. (2012). Sifting through genomes with iterative-sequence clustering produces a large, phylogenetically diverse protein-family resource. BMC Bioinformatics, 13(1), 264. doi:10.1186/1471-2105-13-264
Methods

Stewart, F. J., Ottesen, E. A., & DeLong, E. F. (2010). Development and quantitative analyses of a universal rRNA-subtraction protocol for microbial metatranscriptomics. The ISME Journal, 4(7), 896–907. doi:10.1038/ismej.2010.18
Methods

Vorobev, A., Sharma, S., Yu, M., Lee, J., Washington, B. J., Whitman, W. B., … Moran, M. A. (2018). Identifying

labile DOM components in a coastal ocean through depleted bacterial transcripts and chemical signals. Environmental Microbiology, 20(8), 3012–3030. doi:10.1111/1462-2920.14344
Results
,
Methods

Wu, Z., Rodgers, R. P., & Marshall, A. G. (2004). Two- and Three-Dimensional van Krevelen Diagrams: A Graphical Analysis Complementary to the Kendrick Mass Plot for Sorting Elemental Compositions of Complex Organic Mixtures Based on Ultrahigh-Resolution Broadband Fourier Transform Ion Cyclotron Resonance Mass Measurements. Analytical Chemistry, 76(9), 2511–2516. doi:10.1021/ac0355449
Methods

[table of contents | back to top]

Parameters

Parameter	Description	Units
Sample_Name	sample identifier	unitless
SRA_Experiment	NCBI SRA expreriment accession number	unitless
SRA_link	link to NCBI SRA accession page	unitless
Date	incubation end date	unitless
Tidal_Stage	tidal stage: LT = low tide; HT = high tide	unitless
Time_Point	Hours of dark incubation	unitless

[table of contents | back to top]

Instruments

Dataset-specific Instrument Name	Illumina HiSeq2500
Generic Instrument Name	Automated DNA Sequencer
Dataset-specific Description	mRNA-enriched material was sequenced
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

Moran_Sapelo_2012-14

Website	https://www.bco-dmo.org/deployment/661864
Platform	Univ_Georgia
Start Date	2012-09-01
End Date	2014-10-31
Description	Microbial 'omics studies

[table of contents | back to top]

Project Information

High Resolution Linkages Between DOC Turnover and Bacterioplankton in a Coastal Ocean (SIMCO)

Coverage: Southeastern U.S. coastal ocean, 31.4° N Lat, 81.3° W Lon

Description from NSF award abstract:
Long-standing questions regarding the fate of dissolved organic carbon (DOC) in coastal oceans require a better understanding of the network that links bacterioplankton metabolism with carbon transformation. These questions address uncertainties about the composition of the bioreactive DOC components transformed in ocean margins, and the role of bacterial taxonomic and genetic composition in determining the fate of DOC.

This project will infuse a new type of data into coastal carbon cycle research based on high-resolution chemical analysis coupled with bacterial gene expression measures. It will extend DOC process studies down to the single-compound level and bacterial activity studies down to the single-gene level, and integrate this information into existing bioinformatic resources for biogeochemical and modeling applications.

The specific goals of this project are:
1) To reconstruct major components of the network linking DOC composition, DOC turnover, and bacterial heterotrophy in the coastal ocean (the composition of the DOC pool, the major bioreactive components, the bacterioplankton taxa mediating transformations, and the bacterial genes and pathways responsible).

2) To test hypothesized network links for selected DOC compounds using a simplified system that queries individual DOC compounds against a complex natural microbial community.

3) To test hypothesized network links for marine bacteria using a simplified system that queries a single generalist heterotrophic bacteria against a complex natural DOC pool.

4) To verify predicted DOC-gene linkages that are most informative about heterotrophic activities of bacterioplankton.

This research addresses fundamental questions on bacterial mediation of organic carbon fate in the ocean and atmosphere. As such, these investigations linking the chemical changes in dissolved organic carbon with patterns of gene expression in coastal bacterioplankton communities will be of interest to scientists across several disciplines.

Note: The project acronym, SIMCO, means "Sapelo Island Microbial Carbon Observatory".

[table of contents | back to top]

Funding

Funding Source	Award
NSF Division of Ocean Sciences (NSF OCE)	OCE-1356010

[table of contents | back to top]