DNA and cDNA (RNA) V4 tag sequence genetic accessions from microbial eukaryotes sampled in the upper euphotic zone of the eastern North Pacific during R/V Yellowfin cruises from 2003-2018

Website: https://www.bco-dmo.org/dataset/770110
Data Type: Cruise Results
Version: 1
Version Date: 2019-06-07

Project
» Protistan, prokaryotic, and viral processes at the San Pedro Ocean Time-series (SPOT)

Contributors	Affiliation	Role
Caron, David	University of Southern California (USC)	Principal Investigator
Mesrop, Lisa	University of Southern California (USC)	Co-Principal Investigator
York, Amber D.	Woods Hole Oceanographic Institution (WHOI BCO-DMO)	BCO-DMO Data Manager

Abstract
DNA and cDNA (RNA) V4 tag sequences that characterize the spatiotemporal (monthly and seasonal) variability of eukaryotic microbial assemblages in the upper euphotic zone of the eastern North Pacific at 5m depth and Deep chlorophyll maximum (DCM) from 2003 - 2018. These genetic sequences under BioProject PRJNA311248 (https://www.ncbi.nlm.nih.gov/bioproject/PRJNA311248) will be made available at the National Center for Biotechnology Information (NCBI) on October 1st, 2021.

Table of Contents

Coverage

Spatial Extent: **Lat**:33.71 **Lon**:-118.48
Temporal Extent: 2003-09-08 - 2018-12-14

Dataset Description

DNA and cDNA (RNA) V4 tag sequences that characterize the spatiotemporal (monthly and seasonal) variability of eukaryotic microbial assemblages in the upper euphotic zone of the eastern North Pacific at 5m depth and Deep chlorophyll maximum (DCM) from 2003 - 2018.

These genetic sequences under BioProject PRJNA311248 will be made available at National Center for Biotechnology Information (NCBI) on June 1st, 2020.

Methods & Sampling

Methodology:

Seawater samples were collected from SPOT station (3333N, 11824W) from two depths. The SPOT station was sampled from 5m and DCM using 10 L niskin bottles mounted on a CTD rosette. The depth of the DCM was determined by in situ real-time fluorescence data from the CTD during each sample collection. Seawater samples were pre-screened through a 200 um nitex mesh and 80 um nitex using gravity filtration and and in-line filtration. Seawater collected at 5m (2L) and DCM (2L) were filtered onto a GF/F filter (nominal pore size 0.7 um; Whatman) and filters were stored in Caron Lab lysis buffer for DNA extraction and stored dry for RNA extraction. Filters were flash frozen in liquid nitrogen and stored in -80 C until further processing.
Total DNA extraction was accomplished by modifying the protocol described by DNeasy Plant Mini Kit (Qiagen, #69104). Filters were thawed for 5 min and 65C in a heat block and bead beaten by adding RNase-free silica beads and vortexed for 5 min with the addition of RNase A. Lysate was syringe-filtered and the appropriate proportions of neutralization buffer was added to the lysate. Total DNA was then extracted using DNeasy Plant Mini Kit, as per manufacturers instructions.
Total RNA was extracted using the RNeasy Mini Kit (Qiagen, #74104), as per manufacturers instructions. Filters were thawed on ice and 1.5 mls of RLT + (beta-Mercaptoethanol) was added to each filter and bead beaten by adding RNase-free silica beads and vortexed for 5 min. Genomic DNA was removed during RNA extraction (Qiagen, #79254) and checked for residual genomic DNA by performing PCR using (Stoeck et al., 2010) primers and observing an absence of amplification in an agarose gel. RNA was reverse transcribed into cDNA using iScript Select cDNA Synthesis Kit (BioRad, #1708896) with Random Primers. The resulting DNA and cDNA from each sample were PCR amplified using (Stoeck et al., 2010) primers with attached Illumina Adaptors (Forward- TCGTCGGCAGCGTCAGATGTGTATAAGAGACAG-CASCYGCGGTAATTCC, Reverse-GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAG-ACTTTCGTTCTTGATYRA). For each sample, we added 10 ng of genetic material to PCR tubes containing 0.5 um of each forward and reverse primer and 1X Q5 High Fidelity Master Mix for a final PCR volume of 25 ul and final concentration reaction mixture of .4 ng/ul. The PCR thermal profile had an initial activation step (Q5 specific) of 98C for 2 minutes, followed with 10 cycles of 98C for 10 seconds, 53C for 30 seconds, 72C for 30 seconds, and 15 cycles of 98C for 10 seconds, 48C for 30 seconds, and 72C for 30 seconds, and a final extension of 72C for 2 minutes. PCR products were purified using an AMPure bead clean up (Beckman Coulter #A63881). After purification, PCR products were normalized to one another and indexed using Illumina-specific P5 and P7 indices. Final indexes samples were pooled at equimolar concentration of 10 uM and sequenced using 250 bp PE sequences (Laragen).

The CTD was operated per Sea-Birds suggested methods, and was powered up and allowed to stabilize at approximately 0-5 meters prior to profiling. The instrument was lowered at 40 m per minute (maximum rate) for the entire depth of the sampling site (0-890 m).

Sampling ranges:
SPOT DNA samples were extracted from 2003-2018 (omitting 2017). Ten DNA samples were extracted for each month for two depths (5 m and DCM) from 2003-2018 (omitting 2017).
SPOT RNA samples were extracted from SPOT cruises from 2013 - 2016. Three years of RNA samples were collected for each month for two depths (5m and DCM) from 2013 - 2016. RNA samples were extracted for each month consecutively from 2013 - 2016; full dataset from 2013-2016 for RNA samples.

Location:
The SPOT station (33˚33 N, 118˚24 W) is located approximately 10 miles offshore from the Port of Los Angeles in the southern California Bight.

Data Processing Description

Sequences have been analyzed using QIIME2 - DADA2 software (qiime2-2018.8).

The following programs were used to collect and process CTD data: SEASAVE V.7.2 (to display, record and playback data), SEACON (to enter calibration coefficients and configuration records) and Sea-Bird software Seasoft (data is recorded at a full scan rate at 24 Hz).

BCO-DMO Data Manager Processing Notes:
* added a conventional header with dataset name, PI name, version date
* modified parameter names to conform with BCO-DMO naming conventions. Removed spaces and replaced with underscores.
* added Latitude and Longitude columns with values in decimal degrees from a "Coordinate" column containing locations as strings with decimal degrees and directionals.
* Replaced comma in geographic locations column with a semicolon to better support export as csv format.
* Removed empty columns named: filename3, filename4, assembly, fasta_file

[table of contents | back to top]

Related Publications

Hu, S. K., Campbell, V., Connell, P., Gellene, A. G., Liu, Z., Terrado, R., & Caron, D. A. (2016). Protistan diversity and activity inferred from RNA and DNA at a coastal ocean site in the eastern North Pacific. FEMS Microbiology Ecology, fiw050. doi:10.1093/femsec/fiw050
Results

[table of contents | back to top]

Parameters

Parameter	Description	Units
Accession	BioSample accession number at NCBI	unitless
bioproject_accession	BioProject accession number at NCBI	unitless
sample_name	Sample name	unitless
library_ID	Library identifier	unitless
title	Title. Short name for experiment.	unitless
library_strategy	Sequencing technique for this library	unitless
library_source	Library source. Source material that is being sequenced	unitless
library_selection	Library selection (PCR)	unitless
library_layout	Library layout. Specifies whether to expect single, paired, or other configuration of reads. In the case of paired reads, information about the relative distance and rientation is specified.	unitless
platform	Sequencing platform	unitless
instrument_model	Instrument and model used for sequencing	unitless
design_description	Free-form description of the methods used to create the sequencing library; a brief materials and methods section.	unitless
filetype	Sequence Read Archives native format type (fastq is a combined nucleotide/qualities sequence file)	unitless
filename	Filename (see NCBI accession)	unitless
filename2	Filename2 (see NCBI accession)	unitless
Depth	Depth (nominal) of sample collection (in meters or "DCM")	unitless
Date_Collected	Date (UTC) of sample collection in ISO 8601 format YYYY-MM-DD	unitless
Geographic_Location	Sampling geolocation name	unitless
Latitude	Latitude	decimal degrees
Longitude	Longitude	decimal degrees

[table of contents | back to top]

Instruments

Dataset-specific Instrument Name	Illumina MiSeq
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.

Dataset-specific Instrument Name	CTD SBE 911plus
Generic Instrument Name	CTD Sea-Bird SBE 911plus
Dataset-specific Description	The CTD is operated per Sea-Birds suggested methods, and is powered up and allowed to stabilize at approximately 0-5 meters prior to profiling.
Generic Instrument Description	The Sea-Bird SBE 911 plus is a type of CTD instrument package for continuous measurement of conductivity, temperature and pressure. The SBE 911 plus includes the SBE 9plus Underwater Unit and the SBE 11plus Deck Unit (for real-time readout using conductive wire) for deployment from a vessel. The combination of the SBE 9 plus and SBE 11 plus is called a SBE 911 plus. The SBE 9 plus uses Sea-Bird's standard modular temperature and conductivity sensors (SBE 3 plus and SBE 4). The SBE 9 plus CTD can be configured with up to eight auxiliary sensors to measure other parameters including dissolved oxygen, pH, turbidity, fluorescence, light (PAR), light transmission, etc.). more information from Sea-Bird Electronics

[table of contents | back to top]

Deployments

UpRISEE_SPOT_13-14

Website	https://www.bco-dmo.org/deployment/684011
Platform	R/V Yellowfin
Start Date	2013-01-16
End Date	2014-06-19
Description	A series of cruises were conducted from January 2013 to June 2014 to the San Pedro Ocean Time-Series (SPOT) station. These cruises were part of a study aimed at characterizing the biological response to upwelling at SPOT: the Upwelling Regime In-Situ Ecosystem Efficiency (Up.R.I.S.E.E.) study.

[table of contents | back to top]

Project Information

Protistan, prokaryotic, and viral processes at the San Pedro Ocean Time-series (SPOT)

Coverage: San Pedro Channel off the coast of Los Angeles

Planktonic marine microbial communities consist of a diverse collection of bacteria, archaea, viruses, protists (phytoplankton and protozoa) and small animals (metazoan). Collectively, these species are responsible for virtually all marine pelagic primary production where they form the basis of food webs and carry out a large fraction of respiratory processes. Microbial interactions include the traditional role of predation, but recent research recognizes the importance of parasitism, symbiosis and viral infection. Characterizing the response of pelagic microbial communities and processes to environmental influences is fundamental to understanding and modeling carbon flow and energy utilization in the ocean, but very few studies have attempted to study all of these assemblages in the same study. This project is comprised of long-term (monthly) and short-term (daily) sampling at the San Pedro Ocean Time-series (SPOT) site. Analysis of the resulting datasets investigates co-occurrence patterns of microbial taxa (e.g. protist-virus and protist-prokaryote interactions, both positive and negative) indicating which species consistently co-occur and potentially interact, followed by examination gene expression to help define the underlying mechanisms. This study augments 20 years of baseline studies of microbial abundance, diversity, rates at the site, and will enable detection of low-frequency changes in composition and potential ecological interactions among microbes, and their responses to changing environmental forcing factors. These responses have important consequences for higher trophic levels and ocean-atmosphere feedbacks. The broader impacts of this project include training graduate and undergraduate students, providing local high school student with summer lab experiences, and PI presentations at local K-12 schools, museums, aquaria and informal learning centers in the region. Additionally, the PIs advise at the local, county and state level regarding coastal marine water quality.

This research project is unique in that it is a holistic study (including all microbes from viruses to small metazoa) of microbial species diversity and ecological activities, carried out at the SPOT site off the coast of southern California. In studying all microbes simultaneously, this work aims to identify important ecological interactions among microbial species, and identify the basis(es) for those interactions. This research involves (1) extensive analyses of prokaryote (archaean and bacterial) and eukaryote (protistan and micro-metazoan) diversity via the sequencing of marker genes, (2) studies of whole-community gene expression by eukaryotes and prokaryotes in order to identify key functional characteristics of microorganismal groups and the detection of active viral infections, and (3) metagenomic analysis of viruses and bacteria to aid interpretation of transcriptomic analyses using genome-encoded information. The project includes exploratory metatranscriptomic analysis of poorly-understood aphotic and hypoxic-zone protists, to examine their stratification, functions and hypothesized prokaryotic symbioses.

[table of contents | back to top]