ARISA relative abundances (bacterial community structure), San Pedro Channel, 5 depths, 2014 (Bacterial, Archaeal, and Protistan Biodiversity project, Marine Viral Dynamics project)

Website: https://www.bco-dmo.org/dataset/535915
Version: 2014-11-05
Version Date: 2014-11-03

Project
» Pattern and Process in Marine Bacterial, Archaeal, and Protistan Biodiversity, and Effects of Human Impacts (Bacterial, Archaeal, and Protistan Biodiversity)
» Marine viral dynamics and incorporation into microbial association networks (Marine Viral Dynamics)

Program
» Dimensions of Biodiversity (Dimensions of Biodiversity)

Contributors	Affiliation	Role
Fuhrman, Jed A.	University of Southern California (USC)	Principal Investigator
Cram, Jacob A.	University of Southern California (USC)	Student, Contact
Copley, Nancy	Woods Hole Oceanographic Institution (WHOI BCO-DMO)	BCO-DMO Data Manager

Table of Contents

- Dataset Description
  - Methods & Sampling
  - Data Processing Description
- Data Files
- Parameters
- Instruments
- Deployments
- Project Information
- Program Information
- Funding

Dataset Description

ARISA relative abundances (bacterial community structure) and environmental measurements. All measurements were taken from August 2000 – January 2011 at five depths from the San Pedro Channel 5m, the Deep Chlorophyll Maximum (DCM or CMAX), 150m, 500m, 890m. Additional data will be added periodically.

Related Dataset:
SPOT environmental data
ARISA Bin Taxonomy
SPOT cruises

Methods & Sampling

Water samples were collected on monthly trips to the SPOT station in the San Pedro Basin.

Cruise logs are available at: http://dornsife.usc.edu/spot/document-access/

Data Processing Description

Detailed information on the **methodology** (pdf), including:
- Satellite measurements
- Assigning Taxonomic Identities to ARISA peaks
- Environmental parameter variability
- Seasonal variability of microbial community structure
- Mantel test approach
- Interannual variability of microbial community structure
- Alpha diversity:
  - Variability between depths
  - Relation to season
  - Relation to community similarity between depths
  - Relation to community change
  - Environmental parameters and community structure: Mantel tests
- Temporal dynamics of microbial taxa over time
  - Transformations
  - Taxonomic Groups
  - OTUs

Relevent References:

Beman JM, Steele JA, Fuhrman JA. (2011). Co-occurrence patterns for abundant marine archaeal and bacterial lineages in the deep chlorophyll maximum of coastal California. ISME J 5: 1077-1085.

* Cram JA, Chow C-ET, Sachdeva R, et al. (2014) Seasonal and interannual variability of the marine bacterioplankton community throughout the water column over ten years. The ISME Journal. doi: 10.1038/ismej.2014.153.

Frouin R, Franz BA, Werdell PJ (2003) The SeaWiFS PAR product. Algorithm updates for the fourth SeaWiFS data reprocessing 46-50.

Fuhrman J, Azam F (1982) Thymidine incorporation as a measure of heterotrophic bacterioplankton production in marine surface waters - evaluation and field results. marine biology 66:109-120.

Kirchman D, K'nees E, Hodson R (1985) Leucine incorporation and its potential as a measure of protein synthesis by bacteria in natural aquatic systems. Appl Environ Microbiol 49:599-607.

Morel A, Gentili B (2009) A simple band ratio technique to quantify the colored dissolved and detrital organic material from ocean color remotely sensed data. Remote Sensing of Environment 113:998-1011. doi: 10.1016/j.rse.2009.01.008

Noble RT, Fuhrman JA. (1998). Use of SYBR Green I for rapid epifluorescence counts of marine viruses and bacteria. Aquat. Microb. Ecol 14: 113-118.

Parsons TR (1984) A manual of chemical and biological methods for seawater analysis, 1st ed. Pergamon Press, Oxford [Oxfordshire]; New York

Patel A, Noble RT, Steele JA, Schwalbach MS, Hewson I, Fuhrman JA. (2007). Virus and prokaryote enumeration from planktonic aquatic environments by
epifluorescence microscopy with SYBR Green I. Nat Protoc 2: 269-276.

Stramski D, Reynolds RA, Babin M, et al. (2008) Relationships between the surface concentration of particulate organic carbon and optical properties in the eastern South Pacific and eastern Atlantic Oceans. Biogeosciences 5:171-201.

BCO-DMO Processing:

- split environmental and biological data into separate datasets to reduce overall size of each dataset
- added conventional header with dataset name, PI name, version date
- renamed parameters to BCO-DMO standard or more descriptive terms
- moved columns amoA through bcsim89 before the ARISA_# columns
- transformed ARISA_####.### columns to rows with new column 'arisa_frag' for the arisa name and 'rel_abund' for the values
- replaced NA with nd
- reformated data from m/d/yyyy to yyyy-mm-dd
- changed year from 1 or 2 digits to 4 digits

- added lat, lon, cruise_id, cruise_id2
- missing depth data added from CTD logs where available

[table of contents | back to top]

Data Files

File
arisa_latlon_sort_vess_bio.csv (Comma Separated Values (.csv), 48.04 MB) MD5:9f283db4482f7694c4a12983eb986c48 Primary data file for dataset ID 535915

[table of contents | back to top]

Parameters

Parameter	Description	Units
year	year of sampling	YYYY
cruise_id2	cruise identification	unitless
cruise_id	secondary deployment identification: same for all lab work (lab_Fuhrman_2014)	unitless
lat	latitude; north is positive	decimal degrees
lon	longitude; east is positive	decimal degrees
month	month of sampling	MM
date_local	sampling date	YYYY-MM-DD
Dayl	Day of the month (Imputed with 15 if day not recorded)	unitless
depth_mixlayr	Mixed layer depth	meters
Cmax_Depth	Chlorophyll maximum depth	meters
depth	Actual depth	meters
depth_n	sample depth; CMAX = depth of chlorophyl maximum	meters

bact_abund	Bacterial abundance	cells/ml
viral_abund	Viral abundance	VLP/ml
virus_bact_ratio	Virus to bacteria ratio	unitless
amoA	Arch amoA gene from qPCR	copies/ng
Archaea_16S	Archaea 16S from qPCR	copies/ng
Group1_16S	Group 1 16S from qPCR	copies/ng
Leucine_uptake	[bacteria]/bacterial productivity (Leucine)	Days
TurnoverLeu	[bacteria]/bacterial productivity (Leucine)	Days
Thymedine_uptake	[bacteria]/bacterial productivity (Thymedine)	Days
TurnoverThy	[bacteria]/bacterial productivity (Thymedine)	Days
ShannonH	Shannon Biodiversity Index	unitless
InvSimpson	Inverse Simpson Index	unitless
Richness10000	Richness 0.01%	# operational taxonomic units
Richness1000	Richness 0.1%	# operational taxonomic units
Richness100	Richness 1%	# operational taxonomic units
PeilouJ	Pelou's Evenness Index	unitless
BrayCurtis_shift	Bray-Curtis Community Shift	unitless
BrayCurtissim_5m	Bray-Curtis similarity to 5m sample	unitless
BrayCurtissim_DCM	Bray-Curtis similarity to deep chlophyll maximum sample	unitless

BrayCurtissim_150m	Bray-Curtis similarity to 150m sample	unitless
BrayCurtissim_500m	Bray-Curtis similarity to 500m sample	unitless
BrayCurtissim_890m	Bray-Curtis similarity to 890m sample	unitless
arisa_frag	ARISA fragment identification	unitless
rel_abund	ARISA fragment relative abundance	proportion

[table of contents | back to top]

Instruments

Dataset-specific Instrument Name	CTD Sea-Bird
Generic Instrument Name	CTD Sea-Bird
Generic Instrument Description	A Conductivity, Temperature, Depth (CTD) sensor package from SeaBird Electronics. This instrument designation is used when specific make and model are not known or when a more specific term is not available in the BCO-DMO vocabulary. Refer to the dataset-specific metadata for more information about the specific CTD used. More information from: http://www.seabird.com/

Dataset-specific Instrument Name	Epifluorescence Microscope
Generic Instrument Name	Fluorescence Microscope
Generic Instrument Description	Instruments that generate enlarged images of samples using the phenomena of fluorescence and phosphorescence instead of, or in addition to, reflection and absorption of visible light. Includes conventional and inverted instruments.

Dataset-specific Instrument Name	Dissolved Oxygen Sensor
Generic Instrument Name	Oxygen Sensor
Dataset-specific Description	Sea-bird, model 13
Generic Instrument Description	An electronic device that measures the proportion of oxygen (O2) in the gas or liquid being analyzed

Dataset-specific Instrument Name	Thermal Cycler
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]

Deployments

lab_Fuhrman_2014

Website	https://www.bco-dmo.org/deployment/535519
Platform	USC
Start Date	2014-10-17
End Date	2014-10-17
Description	Microbial diversity laboratory studies. Monthlly cruises to collect water samples in Los Angeles, California area.

[table of contents | back to top]

Project Information

Pattern and Process in Marine Bacterial, Archaeal, and Protistan Biodiversity, and Effects of Human Impacts (Bacterial, Archaeal, and Protistan Biodiversity)

Website: http://dornsife.usc.edu/labs/usc-microbial-observatory/

Coverage: San Pedro Ocean Time Series; approx. 33N, 118W

Description from NSF award abstract:
Bacteria, Archaea, and Protists dominate global elemental cycling and are immensely diverse genetically, taxonomically, and functionally. Yet the extent of marine microbial diversity, its patterns, and relationships among genetic, taxonomic, and functional diversity are very poorly characterized, even though the ocean covers 70% of the planet's surface. Among the least well known variables is the effect of human impacts on native marine microbial systems, although it is recognized that impacted systems are more prone to events like harmful algal blooms. Knowledge of these relationships and impacts are necessary to anticipate the responses of biota to global changes and feedback mechanisms that may alter the extents, rates, and even pathways of such changes. This project will expand upon an existing NSF-funded 10+-year monthly ocean time series (Microbial Observatory) that has focused on a single site midway between Los Angeles and Santa Catalina Island, to also include quarterly sampling adjacent to the impacted LA Harbor region to the barely-impacted Catalina coast. USC already runs facilities in LA Harbor and Catalina, with daily boats between (no cost). Measurements include (1) Genetic diversity: high throughput DNA sequences of "housekeeping" and functional genes. (2) Taxonomic diversity: high throughput tag sequences of small subunit ribosomal RNA genes, flow

cytometry, automated image analysis (3) Functional Diversity: (a) Functional measurements (carbon fixation and respiration rates, microbial growth and grazing rates, cell size, morphology, and biomass variations), (b) distribution and expression of particular target functional genes involved with processes central to the cycles of carbon, nitrogen, and sulfur, (c) exploratory metatranscriptomics to explore functionalities that were not anticipated. (4) Integrating these: Multivariate statistical and network approaches including newly developed techniques (e.g. Bayesian networks to examine cause-effect relationships), and high speed computational approaches to assess the relationships among the genetic, taxonomic, and functional aspects of biodiversity observed. The PIs will also examine the collected data for signatures and specific effects (on organism identity and functions) associated with human impacted harbor site vs. the relatively pristine one.

The PIs will use network and time series analysis, along with other statistical tools to integrate "classical" microbial and oceanographic rate process measurements, flow cytometric and microscopic characterizations of communities, along with targeted as well as untargeted metagenomics and metatranscriptomics to relate genetic and taxonomic diversity with specific functions (at organismal, food web, and system levels). For example, they should be able to determine how different variants of particular taxa (e.g. at resolution levels ranging from what might be considered near the subspecies to genus levels) would differ in their association with particular measured functions, functional genes, or particular other taxa - or they might see how particular clusters of related organisms behave similarly or differently in their associations. This project offers an unprecedented and potentially transformative opportunity to combine and integrate measurements of genetic, taxonomic, and functional diversity along with direct measurements of system function in a well studied marine system that includes a gradient from one of the world's busiest harbors to a largely pristine ocean habitat. Far beyond just describing the distributions of organisms and functions (itself a necessary first step), they will specifically link spatial and temporal variations in a variety of functions with variations in genetic and taxonomic community composition.

Marine viral dynamics and incorporation into microbial association networks (Marine Viral Dynamics)

Website: http://dornsife.usc.edu/labs/fuhrmanlab/research/

Coverage: Southern California between Los Angeles and Santa Catalina Island; Approx. 33.5N, 118.5 W

Description from NSF award abstract:
Marine microbes are tremendously abundant and are major players and driving forces in global biogeochemical cycles of carbon, nitrogen, phosphorus, and iron. We learned over the past two decades that viruses are pervasive elements in marine systems, with significant ecological, biogeochemical, genetic, and evolutionary effects on cellular marine organisms, but we have remarkably little information about the dynamics of marine viral community structure and how it relates to the community structure of their hosts (largely bacteria and phytoplankton). Such information is critical for developing proper conceptual and practical models of the roles of viruses and how these change over time and space. The goals of this project are:
(1) primarily, to characterize a significant subset of the natural virus community and its dynamics, along with bacterial host communities, as they change over daily to monthly time scales at the USC well-studied marine Microbial Observatory site (midway between Los Angeles and Santa Catalina Island), testing hypotheses regarding repeating patterns, host range effects, and taxa-time relationships, and
(2) secondarily, to incorporate these viruses into microbial association networks by statistically connecting particular types of viruses to specific potential hosts.

Approaches for this study include:
(a) nested daily, weekly, and monthly collection of bacteria and viruses for nucleic acid samples,
(b) amplification of conserved genes, as proxy phylogenetic markers, from a few moderately-well-characterized broad viral groups previously readily found in seawater (i.e. the T4-like myoviruses, T7-like podoviruses), as well as bacterial rRNA genes,
(c) extensive sequencing, after screening by community fingerprinting, from the mixed amplified products,
(d) binning of the sequences or fingerprint fragments into operational taxonomic units (OTUs) at different levels of resolution,
(e) evaluation of the results with statistical approaches to examine temporal patterns, relationships (including time-lagged ones) with other viral OTUs, bacteria, protists (monthly only), and environmental parameters,
(f) incorporating the viral OTUs mathematically into microbial association networks.

Data on environmental parameters, bacteria, and protists are already being collected monthly for an existing Microbial Observatory, so the viral work is complementary to this project, providing a major value-added component. Similarly, this project will add selected daily and weekly microbial data to the Microbial Observatory. Data from the literature and from the PI's preliminary results show they have the technology and capability to meet the first goal, and to our knowledge this would be the first such data set of its scope and kind. The investigators have already published in 2006 that the bacterial communities at the 5m depth of this site show a predictable repeating annual cycle in bacterial community composition, so the expectation of a predictable repeating viral community is not unreasonable. They also have some preliminary data showing some repeated viral occurrences. The second goal requires that there are indeed significant statistical relationships between the viruses and other measured parameters, which the PI anticipates to be the case, but of course cannot predict; if they cannot be demonstrated, this result itself would be informative and would constrain the possible modes of microbial/viral interactions.

[table of contents | back to top]

Program Information

Dimensions of Biodiversity (Dimensions of Biodiversity)

Website: http://www.nsf.gov/funding/pgm_summ.jsp?pims_id=503446

Coverage: global

(adapted from the NSF Synopsis of Program)
Dimensions of Biodiversity is a program solicitation from the NSF Directorate for Biological Sciences. FY 2010 was year one of the program. [MORE from NSF]

The NSF Dimensions of Biodiversity program seeks to characterize biodiversity on Earth by using integrative, innovative approaches to fill rapidly the most substantial gaps in our understanding. The program will take a broad view of biodiversity, and in its initial phase will focus on the integration of genetic, taxonomic, and functional dimensions of biodiversity. Project investigators are encouraged to integrate these three dimensions to understand the interactions and feedbacks among them. While this focus complements several core NSF programs, it differs by requiring that multiple dimensions of biodiversity be addressed simultaneously, to understand the roles of biodiversity in critical ecological and evolutionary processes.

[table of contents | back to top]

Funding

Funding Source	Award
NSF Division of Ocean Sciences (NSF OCE)	OCE-1136818
NSF Division of Molecular and Cellular Biosciences (NSF MCB)	MCB-0703159

[table of contents | back to top]