

Genetic accession numbers and sampling information for metaviromes collected during the R/V Thomas G. Thompson cruise TN-280 along Line P in the northeast Pacific Ocean in May of 2012

Website: <https://www.bco-dmo.org/dataset/762912>

Data Type: Cruise Results

Version: 1

Version Date: 2019-03-25

Project

» [Ecology of diatom viruses: connecting physiology and field dynamics through host transcriptional responses](#)
(Diatom Viruses)

Contributors	Affiliation	Role
Rocap, Gabrielle	University of Washington (UW)	Principal Investigator
Carlson, Michael	University of Washington (UW)	Co-Principal Investigator
York, Amber D.	Woods Hole Oceanographic Institution (WHOI BCO-DMO)	BCO-DMO Data Manager

Table of Contents

- [Coverage](#)
- [Dataset Description](#)
 - [Methods & Sampling](#)
 - [Data Processing Description](#)
- [Related Publications](#)
- [Parameters](#)
- [Instruments](#)
- [Deployments](#)
- [Project Information](#)
- [Funding](#)

Coverage

Spatial Extent: N:48.82 E:-125.5 S:48.58 W:-128.66

Temporal Extent: 2012-05-17 - 2012-05-20

Dataset Description

This dataset contains genetic accession numbers and sampling information for metaviromes collected during the R/V Thomas G. Thompson cruise TN-280 along Line P in the northeast Pacific Ocean in May of 2012. The accessions identified in this dataset are housed at the National Center for Biotechnology Information (NCBI) under BioProject PRJNA437470. The status of this page will be "Data not available yet" until the genetic accessions are publicly available at NCBI.

Methods & Sampling

Samples were obtained from Stations P1 (48.58 N, 125.50 W), P4 (48.65 N, 126.67 W), P6 (48.74 N, 127.67 W) and P8 (48.82 N, 128.66 W) on cruise TN280 between May 17-20 2012.

Water from all samples was obtained from Niskin bottles on the CTD Rosette. Plankton were collected on a 2.0 µm pore-size filter (142 mm diameter polycarbonate) and a 0.2 µm pore-size filter (142 mm diameter Supor) following prefiltration through a 53 µm pore-size filter. 20L of the <0.2 µm filtrate was amended with iron chloride (1g L⁻¹) and incubated for at least 1 hour at room temperature to flocculate viruses (John et al. 2011). Viral flocculates were filtered onto 0.6 µm polycarbonate filters and stored at 4C. The GenBank BioSample

accession numbers are SAMN08663113- SAMN08663116.

Viruses were resuspended from filters by incubating them with 0.2M ascorbate-0.25M EDTA-Mg2-0.25M Tris-HCL for at least 24 hours with periodic shaking (John et al. 2011). Viral suspensions were incubated for 2 hours at room temperature with DNase 1 (100 U ml⁻¹, Ambion) to remove contaminating extracellular DNA. Viral DNA was extracted using a MoBio PowerSoil Total RNA isolation kit with the DNA elution column accessory, according to the manufacturers instructions. To convert single stranded DNA (ssDNA) viruses to double stranded DNA (dsDNA), complementary strands were synthesized using a modified random priming-mediated sequence-independent single-primer amplification (RP-SISPA) method designed to generate quantitative whole genome shotgun libraries (Djikeng et al. 2008, Culley et al. 2010). Triplicate 20 µl reactions for each sample and a no template control containing 10 µl template (or water), 0.2mM dNTPs, 0.5 mM DTT, 1 µM FR26RV-N primer (5-GCCGGAGCTCTGCAGATATCNNNNN -3), 1mM MgCl₂, and 1X Klenow buffer were heated to 94 C for 3 minutes and snap cooled on ice. 2.5 U of Klenow Fragment, 3-5 exo- (New England Biolands) was added to each reaction then incubated at 37 C for 60 minutes followed by 75 C for 10 minutes. Reactions were pooled, cleaned with ethanol precipitation, and resuspended in T low-E (10mM Tris-HCl, 0.1 mM EDTA pH 8.0).

To construct Illumina shotgun metagenome libraries, viral DNA was sheared to 800-1000 bp using a Covaris nebulizer and cleaned with AmPure XP beads. Viral metagenomes (viromes) were constructed using the Rubicon ThruPLEX-FD Kit, which uses qPCR to linearly amplify low concentration DNA templates. Libraries were sequenced on an Illumina HiSeq 2500 with 150 paired end reads at the Michigan State University Sequencing Center. FASTQ files of the raw sequence reads are deposited in the Sequence Read Archive under study SRP134205 and the individual file accession numbers SRR6819460-SRR6819463.

Virome reads were trimmed of adapter sequences and low quality regions and filtered to remove low quality reads with the Trimmomatic pipeline (Bolger et al., 2014). Sequences were de novo assembled into contigs by first using diginorm (Brown et al., 2012) followed by the velvet assembler (Zerbino and Birney, 2008) with k-mer length set at 25, 29, 33, 51 for P1, P4, P6, and P8 respectively. Assemblies have been deposited in the Whole Genome Shotgun archive under the accession numbers

GeoMICS_DNAvirome_P1 PYID00000000

GeoMICS_DNAvirome_P4 PYIE00000000

GeoMICS_DNAvirome_P6 PYIF00000000

GeoMICS_DNAvirome_P8 PYIG00000000

Data Processing Description

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

* Converted date to ISO 8601 format

* parsed column with lat and lon into separate latitude and longitude columns in decimal degrees.

"48.58 N 125.50 W" -> lat:"48.58" lon:"-125.50"

* "m" removed from Depth column values. Units are described in the metadata.

[[table of contents](#) | [back to top](#)]

Related Publications

Bolger, A. M., Lohse, M., & Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*, 30(15), 2114–2120. doi:[10.1093/bioinformatics/btu170](https://doi.org/10.1093/bioinformatics/btu170)
Methods

Brown, C. T., Howe, A., Zhang, Q., Pyrkosz, A. B., & Brom, T. H. (2012). A reference-free algorithm for computational normalization of shotgun sequencing data. arXiv preprint arXiv:[1203.4802](https://arxiv.org/abs/1203.4802).
Methods

John, S. G., Mendez, C. B., Deng, L., Poulos, B., Kauffman, A. K. M., Kern, S., ... Sullivan, M. B. (2010). A simple and efficient method for concentration of ocean viruses by chemical flocculation. *Environmental Microbiology Reports*, 3(2), 195–202. doi:[10.1111/j.1758-2229.2010.00208.x](https://doi.org/10.1111/j.1758-2229.2010.00208.x)
Methods

Zerbino, D. R., & Birney, E. (2008). Velvet: Algorithms for de novo short read assembly using de Bruijn graphs. *Genome Research*, 18(5), 821–829. doi:[10.1101/gr.074492.107](https://doi.org/10.1101/gr.074492.107)
Methods

[[table of contents](#) | [back to top](#)]

Parameters

Parameters for this dataset have not yet been identified

[[table of contents](#) | [back to top](#)]

Instruments

Dataset-specific Instrument Name	Illumina HiSeq 2500
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	
Generic Instrument Name	CTD - profiler
Generic Instrument Description	The Conductivity, Temperature, Depth (CTD) unit is an integrated instrument package designed to measure the conductivity, temperature, and pressure (depth) of the water column. The instrument is lowered via cable through the water column. It permits scientists to observe the physical properties in real-time via a conducting cable, which is typically connected to a CTD to a deck unit and computer on a ship. The CTD is often configured with additional optional sensors including fluorometers, transmissometers and/or radiometers. It is often combined with a Rosette of water sampling bottles (e.g. Niskin, GO-FLO) for collecting discrete water samples during the cast. This term applies to profiling CTDs. For fixed CTDs, see https://www.bco-dmo.org/instrument/869934 .

[[table of contents](#) | [back to top](#)]

Deployments

TN280

Website	https://www.bco-dmo.org/deployment/664928
Platform	R/V Thomas G. Thompson
Start Date	2012-05-16
End Date	2012-05-22

[[table of contents](#) | [back to top](#)]

Project Information

Ecology of diatom viruses: connecting physiology and field dynamics through host transcriptional responses (Diatom Viruses)

Coverage: Puget Sound, North Pacific

Extracted from the NSF award abstract:

Overview: Viruses play critical roles in aquatic ecosystems. Phages infecting marine bacteria are abundant members of the plankton that contribute to cell mortality, structure population diversity and drive genome evolution through horizontal gene transfer. Viruses infecting eukaryotic phytoplankton have been demonstrated to induce both life cycle switching and programmed cell death in coccolithophorids and be significant agents of mortality in blooms of pelagophytes, haptophytes and raphidophytes. However, much less is known about viruses infecting one of the largest, most diverse and most productive groups of algae, the diatoms. Only thirteen diatom infecting viruses have been reported, and little is known about their mechanisms of infection, effects on host metabolism or diversity and dynamics in the field. This is a remarkable knowledge gap considering the ecological importance of the diatoms. Infection with a clonal virus on *Pseudo-nitzschia multiseries* can result in complete host lysis within 12-16 hours. The *P. multiseries* virus (PmDNAV) is a single stranded DNA virus with an icosahedral capsid of 50 nm. The PmDNAV infects the widest host range of any marine eukaryote-infecting virus, lysing other strains of *P. multiseries*, other species of *Pseudo-nitzschia*, and other genera of diatoms including many centric diatoms. With the recent completion of the genome of the host, *P. multiseries*, we now have a model system to investigate the response of the host to viral infection and the potential impacts of viruses on diatom mortality in the field. The objectives of this project are to:

1. isolate and characterize additional diatom viruses utilizing established methods, using a variety of host strains and field viral concentrate combinations
2. use RNA-Seq to determine the transcriptional profiles of three diatoms (*P. multiseries*, *P. pungens* and *T. pseudonana*) during the course of viral infection
3. determine a surface water metavirome at four stations on a coastal to open ocean transect in diatom dominated waters in the Pacific Northwest (line P), with an emphasis on diversity and biogeography of ssDNA and ssRNA viruses.

Viral and host genes whose expression is diagnostic of viral infection, will be identified by observing genomic responses to infection in culture. These genes, along with viruses assembled in the metaviromes, will be combined with eukaryotic metatranscriptomes already available from the same waters to assess virus activity in the field.

Intellectual Merit: This project seeks to strengthen the model system initiated by the discovery of the *Pseudo-nitzschia multiseries* DNA virus. The host-virus transcriptomics will lay the groundwork for assessing the impact of viruses on diatom communities in the environment. In turn, the paired metaviromes and metatranscriptomes will reveal new questions about both diatom virus diversity and function that can then be further explored by controlled, culture-based experiments. This research will be the first extensive exploration of diatom virus ecology and function and will ultimately help further connect viruses and diatoms to global biogeochemical cycles, unravel complex organismal interactions, and inform ocean-related public health.

[[table of contents](#) | [back to top](#)]

Funding

Funding Source	Award
NSF Division of Ocean Sciences (NSF OCE)	OCE-1356779

[[table of contents](#) | [back to top](#)]