Metabarcoding zooplankton at station ALOHA: Operational taxonomic unit (OTU) tables and fasta files for representative sequences from each OTU (Plankton Population Genetics project)

Website: https://www.bco-dmo.org/dataset/700279
Data Type: Other Field Results
Version: 1
Version Date: 2017-05-25

Project
» Basin-scale genetics of marine zooplankton (Plankton Population Genetics)

Contributors	Affiliation	Role
Goetze, Erica	University of Hawaii (UH)	Principal Investigator, Contact
York, Amber D.	Woods Hole Oceanographic Institution (WHOI BCO-DMO)	BCO-DMO Data Manager

Abstract
Metabarcoding data for the zooplankton community in the epipelagic, mesopelagic and upper bathypelagic zones (0-1500m) of the North Pacific Subtropical Gyre. The goal of this study was to assess the hidden diversity present in zooplankton assemblages in midwaters, and detect vertical gradients in species richness, depth distributions, and community composition of the full zooplankton assemblage. Samples were collected in June 2014 from Station ALOHA (22.75, -158) using a 1-meter square Multiple Opening and Closing Nets and Environmental Sampling System (MOCNESS, 200um mesh), on R/V Falkor cruise FK140613. Next generation sequence data (Illumina MiSeq, V3 chemistry, 300-bp paired-end) of the zooplankton assemblage derive from amplicons of the V1-V2 region of 18S rRNA (primers described in Fonseca et al. 2010). The data includes sequences and read count abundance information for molecular OTUs from both holoplanktonic and meroplanktonic taxa. All results derive from analyses in mothur v1.36.1 (Schloss et al. 2009, Kozich et al. 2013).

Table of Contents

Coverage

Spatial Extent: **Lat**:22.75 **Lon**:-158
Temporal Extent: 2014-06-13 - 2014-06-19

Dataset Description

This data submission consists of metabarcoding data for the zooplankton community in the epipelagic, mesopelagic and upper bathypelagic zones (0-1500m) of the North Pacific Subtropical Gyre. The goal of this study was to assess the hidden diversity present in zooplankton assemblages in midwaters, and detect vertical gradients in species richness, depth distributions, and community composition of the full zooplankton assemblage. Samples were collected in June 2014 from Station ALOHA (22.75, -158) using a 1-meter square Multiple Opening and Closing Nets and Environmental Sampling System (MOCNESS, 200um mesh), on

R/V Falkor cruise FK140613. Next generation sequence data (Illumina MiSeq, V3 chemistry, 300-bp paired-end) of the zooplankton assemblage derive from amplicons of the V1-V2 region of 18S rRNA (primers described in Fonseca et al. 2010). The data includes sequences and read count abundance information for molecular OTUs from both holoplanktonic and meroplanktonic taxa. All results derive from analyses in mothur v1.36.1 (Schloss et al. 2009, Kozich et al. 2013).

Tables from four analyses are included in this submission:

1) 97_OTUtable: read counts for each OTU (clustered at 97% similarity) across 54 samples (depth, size fractionated), with NCBI BLAST results for the top representative sequence from each OTU.

2) 97_OTU_subsampled: read counts for each OTU (clustered at 97% similarity), subsampled for even sequencing coverage across 54 samples (depth, size fractionated, with NCBI BLAST results for the top representative sequence from each OTU.

3) 99_OTU: read counts for each OTU (clustered at 99% similarity) across 54 samples (depth, size fractionated), with NCBI BLAST results for the top representative sequence from each OTU.

4) 99_OTU_subsampled: read counts for each OTU (clustered at 99% similarity), subsampled for even sequencing coverage across 54 samples (depth, size fractionated), with NCBI BLAST results for the top representative sequence from each OTU.

Related dataset containing NCBI accession numbers for sequence data:
Metabarcoding zooplankton at station ALOHA: NCBI SRA accession numbers www.bco-dmo.org/dataset/700961

Methods & Sampling

Sample Codes:

Read count parameter names (e.g. FA3_N7_SF1) include the following three codes:

MOCNESS tow
FA3: Night sampling
FA4: Day sampling

Depth range:
N1: 1500-1000m
N2: 1000-700m
N3: 700-500m
N4: 500-300m
N5: 300-200m
N6: 200-150m
N7: 150-100m
N8: 100-50m
N9: 50m-0m

Wet-sieved zooplankton size fractions
SF1: 0.2-0.5 mm
SF2: 0.5-1.0 mm
SF3: 1.0-2.0 mm

Data parameters with the prefix (NCBI_) are from the National Center for Biotechnology Information (NCBI, https://www.ncbi.nlm.nih.gov/)

Data parameters with the prefix (silva_) are from the Silva database (https://www.arb-silva.de/).

Data Processing Description

BCO-DMO processing notes:

* commas in the data were replaced with semicolons to support export as csv format.

* fasta files added to dataset

Data Files

File
FASTA file - Representative sequences, 97% Clustering filename: RepSeqs97.fasta   (FASTA, 2.05 MB)   MD5:95c7b7787aaefcd69b78cf47a836bb3c FASTA file containing representative sequences for each OTU identified by abundance (top read) and with alignment gaps removed. Clustered at 97% similarity.
FASTA file - Representative sequences, 99% Clustering filename: RepSeqs99.fasta   (FASTA, 2.02 MB)   MD5:4a20d1ca10aef9d64f724918982b3c5d FASTA file containing representative sequences for each OTU identified by abundance (top read) and with alignment gaps removed. Clustered at 99% similarity.
OTU.csv   (Comma Separated Values (.csv), 11.59 MB)   MD5:63b6561afef4f0dea8f056305948289f Primary data file for dataset ID 700279

Parameters

Parameter	Description	Units
analysis_name	Descriptive name of analysis including percent clustering and subsampling or no subsampling	unitless
lat	Latitude of sample site (Station ALOHA)	decimal degrees
lon	Longitude of sample site (Station ALOHA)	decimal degrees
OTU_ID	Operational taxonomic unit (OTU) identifier	unitless
FA3_N1_SF1	Read counts for each OTU within each sample; Sample collected during the day with net 1 (depth 1500-1000m) for zooplankton size fraction 1 (0.2 - 0.5mm)	per OTU
FA3_N1_SF2	Read counts for each OTU within each sample; Sample collected during the day with net 1 (depth 1500-1000m) for zooplankton size fraction 2 (0.5 - 0.1mm)	per OTU
FA3_N1_SF3	Read counts for each OTU within each sample; Sample collected during the day with net 1 (depth 1500-1000m) for zooplankton size fraction 3 (1.0 - 2.0 mm)	per OTU
FA3_N2_SF1	Read counts for each OTU within each sample; Sample collected during the day with net 2 (depth 1000-700m) for zooplankton size fraction 1 (0.2 - 0.5mm)	per OTU
FA3_N2_SF2	Read counts for each OTU within each sample; Sample collected during the day with net 2 (depth 1000-700m) for zooplankton size fraction 2 (0.5 - 0.1mm)	per OTU
FA3_N2_SF3	Read counts for each OTU within each sample; Sample collected during the day with net 2 (depth 1000-700m) for zooplankton size fraction 3 (1.0 - 2.0 mm)	per OTU

FA3_N3_SF1	Read counts for each OTU within each sample; Sample collected during the day with net 3 (depth 700-500m) for zooplankton size fraction 1 (0.2 - 0.5mm)	per OTU
FA3_N3_SF2	Read counts for each OTU within each sample; Sample collected during the day with net 3 (depth 700-500m) for zooplankton size fraction 2 (0.5 - 0.1mm)	per OTU
FA3_N3_SF3	Read counts for each OTU within each sample; Sample collected during the day with net 3 (depth 700-500m) for zooplankton size fraction 3 (1.0 - 2.0 mm)	per OTU
FA3_N4_SF1	Read counts for each OTU within each sample; Sample collected during the day with net 4 (depth 500-300m) for zooplankton size fraction 1 (0.2 - 0.5mm)	per OTU
FA3_N4_SF2	Read counts for each OTU within each sample; Sample collected during the day with net 4 (depth 500-300m) for zooplankton size fraction 2 (0.5 - 0.1mm)	per OTU
FA3_N4_SF3	Read counts for each OTU within each sample; Sample collected during the day with net 4 (depth 500-300m) for zooplankton size fraction 3 (1.0 - 2.0 mm)	per OTU
FA3_N5_SF1	Read counts for each OTU within each sample; Sample collected during the day with net 5 (depth 300-200m) for zooplankton size fraction 1 (0.2 - 0.5mm)	per OTU
FA3_N5_SF2	Read counts for each OTU within each sample; Sample collected during the day with net 5 (depth 300-200m) for zooplankton size fraction 2 (0.5 - 0.1mm)	per OTU
FA3_N5_SF3	Read counts for each OTU within each sample; Sample collected during the day with net 5 (depth 300-200m) for zooplankton size fraction 3 (1.0 - 2.0 mm)	per OTU
FA3_N6_SF1	Read counts for each OTU within each sample; Sample collected during the day with net 6 (depth 200-150m) for zooplankton size fraction 1 (0.2 - 0.5mm)	per OTU
FA3_N6_SF2	Read counts for each OTU within each sample; Sample collected during the day with net 6 (depth 200-150m) for zooplankton size fraction 2 (0.5 - 0.1mm)	per OTU
FA3_N6_SF3	Read counts for each OTU within each sample; Sample collected during the day with net 6 (depth 200-150m) for zooplankton size fraction 3 (1.0 - 2.0 mm)	per OTU
FA3_N7_SF1	Read counts for each OTU within each sample; Sample collected during the day with net 7 (depth 150-100m) for zooplankton size fraction 1 (0.2 - 0.5mm)	per OTU
FA3_N7_SF2	Read counts for each OTU within each sample; Sample collected during the day with net 7 (depth 150-100m) for zooplankton size fraction 2 (0.5 - 0.1mm)	per OTU
FA3_N7_SF3	Read counts for each OTU within each sample; Sample collected during the day with net 7 (depth 150-100m) for zooplankton size fraction 3 (1.0 - 2.0 mm)	per OTU
FA3_N8_SF1	Read counts for each OTU within each sample; Sample collected during the day with net 8 (depth 100-50m) for zooplankton size fraction 1 (0.2 - 0.5mm)	per OTU
FA3_N8_SF2	Read counts for each OTU within each sample; Sample collected during the day with net 8 (depth 100-50m) for zooplankton size fraction 2 (0.5 - 0.1mm)	per OTU

FA3_N8_SF3	Read counts for each OTU within each sample; Sample collected during the day with net 8 (depth 100-50m) for zooplankton size fraction 3 (1.0 - 2.0 mm)	per OTU
FA3_N9_SF1	Read counts for each OTU within each sample; Sample collected during the day with net 9 (depth 50-0m) for zooplankton size fraction 1 (0.2 - 0.5mm)	per OTU
FA3_N9_SF2	Read counts for each OTU within each sample; Sample collected during the day with net 9 (depth 50-0m) for zooplankton size fraction 2 (0.5 - 0.1mm)	per OTU
FA3_N9_SF3	Read counts for each OTU within each sample; Sample collected during the day with net 9 (depth 50-0m) for zooplankton size fraction 3 (1.0 - 2.0 mm)	per OTU
FA4_N1_SF1	Read counts for each OTU within each sample; Sample collected during the night with net 1 (depth 1500-1000m) for zooplankton size fraction 1 (0.2 - 0.5mm)	per OTU
FA4_N1_SF2	Read counts for each OTU within each sample; Sample collected during the night with net 1 (depth 1500-1000m) for zooplankton size fraction 2 (0.5 - 0.1mm)	per OTU
FA4_N1_SF3	Read counts for each OTU within each sample; Sample collected during the night with net 1 (depth 1500-1000m) for zooplankton size fraction 3 (1.0 - 2.0 mm)	per OTU
FA4_N2_SF1	Read counts for each OTU within each sample; Sample collected during the night with net 2 (depth 1000-700m) for zooplankton size fraction 1 (0.2 - 0.5mm)	per OTU
FA4_N2_SF2	Read counts for each OTU within each sample; Sample collected during the night with net 2 (depth 1000-700m) for zooplankton size fraction 2 (0.5 - 0.1mm)	per OTU
FA4_N2_SF3	Read counts for each OTU within each sample; Sample collected during the night with net 2 (depth 1000-700m) for zooplankton size fraction 3 (1.0 - 2.0 mm)	per OTU
FA4_N3_SF1	Read counts for each OTU within each sample; Sample collected during the night with net 3 (depth 700-500m) for zooplankton size fraction 1 (0.2 - 0.5mm)	per OTU
FA4_N3_SF2	Read counts for each OTU within each sample; Sample collected during the night with net 3 (depth 700-500m) for zooplankton size fraction 2 (0.5 - 0.1mm)	per OTU
FA4_N3_SF3	Read counts for each OTU within each sample; Sample collected during the night with net 3 (depth 700-500m) for zooplankton size fraction 3 (1.0 - 2.0 mm)	per OTU
FA4_N4_SF1	Read counts for each OTU within each sample; Sample collected during the night with net 4 (depth 500-300m) for zooplankton size fraction 1 (0.2 - 0.5mm)	per OTU
FA4_N4_SF2	Read counts for each OTU within each sample; Sample collected during the night with net 4 (depth 500-300m) for zooplankton size fraction 2 (0.5 - 0.1mm)	per OTU
FA4_N4_SF3	Read counts for each OTU within each sample; Sample collected during the night with net 4 (depth 500-300m) for zooplankton size fraction 3 (1.0 - 2.0 mm)	per OTU
FA4_N5_SF1	Read counts for each OTU within each sample; Sample collected during the night with net 5 (depth 300-200m) for zooplankton size fraction 1 (0.2 - 0.5mm)	per OTU

FA4_N5_SF2	Read counts for each OTU within each sample; Sample collected during the night with net 5 (depth 300-200m) for zooplankton size fraction 2 (0.5 - 0.1mm)	per OTU
FA4_N5_SF3	Read counts for each OTU within each sample; Sample collected during the night with net 5 (depth 300-200m) for zooplankton size fraction 3 (1.0 - 2.0 mm)	per OTU
FA4_N6_SF1	Read counts for each OTU within each sample; Sample collected during the night with net 6 (depth 200-150m) for zooplankton size fraction 1 (0.2 - 0.5mm)	per OTU
FA4_N6_SF2	Read counts for each OTU within each sample; Sample collected during the night with net 6 (depth 200-150m) for zooplankton size fraction 2 (0.5 - 0.1mm)	per OTU
FA4_N6_SF3	Read counts for each OTU within each sample; Sample collected during the night with net 6 (depth 200-150m) for zooplankton size fraction 3 (1.0 - 2.0 mm)	per OTU
FA4_N7_SF1	Read counts for each OTU within each sample; Sample collected during the night with net 7 (depth 150-100m) for zooplankton size fraction 1 (0.2 - 0.5mm)	per OTU
FA4_N7_SF2	Read counts for each OTU within each sample; Sample collected during the night with net 7 (depth 150-100m) for zooplankton size fraction 2 (0.5 - 0.1mm)	per OTU
FA4_N7_SF3	Read counts for each OTU within each sample; Sample collected during the night with net 7 (depth 150-100m) for zooplankton size fraction 3 (1.0 - 2.0 mm)	per OTU
FA4_N8_SF1	Read counts for each OTU within each sample; Sample collected during the night with net 8 (depth 100-50m) for zooplankton size fraction 1 (0.2 - 0.5mm)	per OTU
FA4_N8_SF2	Read counts for each OTU within each sample; Sample collected during the night with net 8 (depth 100-50m) for zooplankton size fraction 2 (0.5 - 0.1mm)	per OTU
FA4_N8_SF3	Read counts for each OTU within each sample; Sample collected during the night with net 8 (depth 100-50m) for zooplankton size fraction 3 (1.0 - 2.0 mm)	per OTU
FA4_N9_SF1	Read counts for each OTU within each sample; Sample collected during the night with net 9 (depth 50-0m) for zooplankton size fraction 1 (0.2 - 0.5mm)	per OTU
FA4_N9_SF2	Read counts for each OTU within each sample; Sample collected during the night with net 9 (depth 50-0m) for zooplankton size fraction 2 (0.5 - 0.1mm)	per OTU
FA4_N9_SF3	Read counts for each OTU within each sample; Sample collected during the night with net 9 (depth 50-0m) for zooplankton size fraction 3 (1.0 - 2.0 mm)	per OTU
silva_tax_level1	Taxonomic classification level 1 for each OTU; assigned by mothur based on the Silva database	unitless
silva_tax_level2	Taxonomic classification level 2 for each OTU; assigned by mothur based on the Silva database	unitless
silva_tax_level3	Taxonomic classification level 3 for each OTU; assigned by mothur based on the Silva database	unitless
silva_tax_level4	Taxonomic classification level 4 for each OTU; assigned by mothur based on the Silva database	unitless
silva_tax_level5	Taxonomic classification level 5 for each OTU; assigned by mothur based on the Silva database	unitless

silva_tax_level6	Taxonomic classification level 6 for each OTU; assigned by mothur based on the Silva database	unitless
silva_tax_level7	Taxonomic classification level 7 for each OTU; assigned by mothur based on the Silva database	unitless
silva_tax_level8	Taxonomic classification level 8 for each OTU; assigned by mothur based on the Silva database	unitless
silva_tax_level9	Taxonomic classification level 9 for each OTU; assigned by mothur based on the Silva database	unitless
silva_tax_level10	Taxonomic classification level 10 for each OTU; assigned by mothur based on the Silva database	unitless
silva_tax_level11	Taxonomic classification level 11 for each OTU; assigned by mothur based on the Silva database	unitless
silva_tax_level12	Taxonomic classification level 12 for each OTU; assigned by mothur based on the Silva database	unitless
silva_tax_level13	Taxonomic classification level 13 for each OTU; assigned by mothur based on the Silva database	unitless
silva_tax_level14	Taxonomic classification level 14 for each OTU; assigned by mothur based on the Silva database	unitless
assigned_habitat	Assigned community/habitat, based on both mothur classification and NCBI top hit	unitless
NCBI_query_sequence_id	Query sequence (BLAST qseqid); header description of representative sequence from fasta file	unitless
NCBI_top_hit	Subject sequence id (BLAST sseqid): sequence identifier for subject sequence from NCBI; Contains two NCBI accession numbers	unitless
NCBI_percent_identity	Percent identity (BLAST pident): percentage of identical matches in the alignment	unitless
NCBI_percent_coverage	Percent coverage (BLAST qcovs): query coverage per sequence for all high-scoring segment pairs (HSPs)	unitless
NCBI_sequence_length	Alignment length (BLAST length)	unitless
NCBI_mismatch	Number of mismatches (BLAST mismatch)	unitless
NCBI_evalue	Expectation value of match (BLAST evalue)	unitless
NCBI_bitscore	Normalized alignment score for comparison across searches (BLAST bitscore)	unitless
NCBI_percent_positive	Percentage of positive-scoring matches (BLAST ppos)	unitless
NCBI_top_hit_taxID	Taxonomic identifiers (BLAST staxids): unique subject taxonomy ID(s); multiple IDs are separated by a semicolon	unitless
NCBI_top_hit_title	Subject title from NCBI entry (BLAST stitle)	unitless
analysis_description	Analysis description	unitless

Instruments

Dataset-specific Instrument Name	Illumina MiSeq
Generic Instrument Name	Automated DNA Sequencer
Dataset-specific Description	Illumina MiSeq using V3 chemistry (300-bp, paired-end)
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	Agilent 2100 Bioanalyzer
Generic Instrument Name	Bioanalyzer
Generic Instrument Description	A Bioanalyzer is a laboratory instrument that provides the sizing and quantification of DNA, RNA, and proteins. One example is the Agilent Bioanalyzer 2100.

Dataset-specific Instrument Name	quantitative PCR by the Evolutionary Genetics Core Facility (Hawaii Institute of Marine Biology)
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

FK140613

Website	https://www.bco-dmo.org/deployment/700613
Platform	R/V Falkor
Start Date	2014-06-13
End Date	2014-06-19
Description	Student Cruise #3 More about this cruise from the Schmidt Ocean Institute page:https://schmidtocean.org/cruise/net-gains-at-station-aloha/

Project Information

Basin-scale genetics of marine zooplankton (Plankton Population Genetics)

Coverage: Atlantic Ocean, 46 N - 46 S

Description from NSF award abstract:
Marine zooplankton show strong ecological responses to climate change, but little is known about their capacity for evolutionary response. Many authors have assumed that the evolutionary potential of zooplankton is limited. However, recent studies provide circumstantial evidence for the idea that selection is a dominant evolutionary force acting on these species, and that genetic isolation can be achieved at regional spatial scales in pelagic habitats. This RAPID project will take advantage of a unique opportunity for basin-scale transect sampling through participation in the Atlantic Meridional Transect (AMT) cruise in 2014. The cruise will traverse more than 90 degrees of latitude in the Atlantic Ocean and include boreal-temperate, subtropical and tropical waters. Zooplankton samples will be collected along the transect, and mitochondrial and microsatellite markers will be used to identify the geographic location of strong genetic breaks within three copepod species. Bayesian and coalescent analytical techniques will test if these regions act as dispersal barriers. The physiological condition of animals collected in distinct ocean habitats will be assessed by measurements of egg production (at sea) as well as body size (condition index), dry weight, and carbon and nitrogen content. The PI will test the prediction that ocean regions that serve as dispersal barriers for marine holoplankton are areas of poor-quality habitat for the target species, and that this is a dominant mechanism driving population genetic structure in oceanic zooplankton.

Note: This project is funded by an NSF RAPID award. This RAPID grant supported the shiptime costs, and all the sampling reported in the AMT24 zooplankton ecology cruise report (PDF).

Online science outreach blog at: https://atlanticplankton.wordpress.com

Funding

Funding Source	Award
NSF Division of Ocean Sciences (NSF OCE)	OCE-1255697
NSF Division of Ocean Sciences (NSF OCE)	OCE-1338959
NSF Division of Ocean Sciences (NSF OCE)	OCE-1029478