Pool-seq data from laboratory selection lines of copepods collected from Kiel Canal in Germany in 2017 and 2018

Website: https://www.bco-dmo.org/dataset/878335
Data Type: experimental
Version: 1
Version Date: 2022-08-29

Project
» Evolutionary Responses to Global Changes in Salinity and Temperature (Evolutionary genomics of a copepod)

Contributors	Affiliation	Role
Lee, Carol E.	University of Wisconsin (UW-Madison)	Principal Investigator
Stern, David B.	University of Wisconsin (UW-Madison)	Contact
Heyl, Taylor	Woods Hole Oceanographic Institution (WHOI BCO-DMO)	BCO-DMO Data Manager
Rauch, Shannon	Woods Hole Oceanographic Institution (WHOI BCO-DMO)	BCO-DMO Data Manager

Abstract
The dataset contains pool-seq data from a laboratory natural selection experiment conducted on Eurytemora affinis. The E. affinis copepods used in the experiment were collected from Kiel Canal in Kiel, Germany in 2017 (approximately 1000 copepods) and on May 30, 2018 (85 gravid females and 40 juveniles). Individual adult copepods (N = 50; 25 male and 25 female) were collected for sequencing from each laboratory selection line at generations six (after one generation at 0 PSU in the treatment lines) and ten (after five generations at 0 PSU in the treatment lines). Sampled copepods from each line were pooled and their DNA was extracted using the DNeasy Blood and Tissue Extraction kit (Qiagen, Inc.). Paired-end whole-genome sequencing libraries were prepared using the Nextera DNA kit (Illumina Inc.) and sequenced on four lanes of Illumina Hi-Seq 4000 and one lane of Illumina NovaSeq 6000 at the University of Chicago Genomics Facility, generating an average of approximately 117 million paired-end (100 bp) reads per pool. These data have been deposited in NCBI under BioProject number PRJNA844002.

Table of Contents

Coverage

Spatial Extent: **Lat**:54.3333 **Lon**:10.15
Temporal Extent: 2019 - 2019

Methods & Sampling

Sampling and analytical procedures:
The *Eurytemora affinis* copepods used in the laboratory natural selection experiment were collected from Kiel Canal in Kiel, Germany (latitude = 54° 19' 59.88"N, longitude = 10° 9' 0"W) in 2017 (approximately 1000 copepods) and on May 30, 2018 (85 gravid females and 40 juveniles). Wild *E. affinis* populations were collected from eight locations in the Baltic Sea using bongo and WP2 nets with 100 µm mesh in 2019 (see related

dataset 878322). The two collections of copepods were mixed and maintained at 15 PSU to increase population size and acclimate to laboratory conditions. Two samples of adult copepods (25 male and 25 female each) from the mixed culture were collected for pooled whole-genome sequencing (Pool-seq) to represent the starting population for the laboratory natural selection experiment and capture variance in starting SNP frequency. The culture was then split into 14 equally sized beakers. Control lines (N = 4) were maintained for the duration of the experiment in 15 PSU water made with Instant Ocean, along with Primaxin (20 milligrams per liter) to avoid bacterial infection. The control lines were fed the marine alga *Rhodomonas salina* every three to four days with the water changed weekly. The ten treatment lines were exposed to decreasing salinity over the first six generations until they reached 0 PSU (Lake Michigan water, ~300 µS/cm conductivity), and then maintained at 0 PSU for four additional generations.

Beginning at generation two, salinity declination proceeded at each generation as follows: 5 PSU, 1 PSU, 0.1 PSU, 0.01 PSU, 0 PSU. The generation number was monitored by assuming a generation time of approximately three weeks. Treatment lines were fed a 50:50 mixture of *R. salina* and the freshwater alga *R. minuta* at 5 PSU and only *R. minuta* at 1 PSU and below.

Individual adult copepods (N = 50; 25 male and 25 female) were collected for sequencing from each laboratory selection line at generations six (after one generation at 0 PSU in the treatment lines) and ten (after five generations at 0 PSU in the treatment lines). Sampled copepods from each line were pooled and their DNA was extracted using the DNeasy Blood and Tissue Extraction kit (Qiagen, Inc.). Paired-end whole-genome sequencing libraries were prepared using the Nextera DNA kit (Illumina Inc.) and sequenced on four lanes of Illumina Hi-Seq 4000 and one lane of Illumina NovaSeq 6000 at the University of Chicago Genomics Facility, generating an average of approximately 117 million paired-end (100 bp) reads per pool.

Data Processing Description

Raw sequence reads were mapped to a reference genome to call SNPs. We then detected SNPs and genomic regions under natural selection in response to salinity change.

The following software was used:
BLAST 2.7.1+, BWA-MEM v0.7.17, CD-HIT v4.7, PoPoolation2, SAMBLASTER v0.1.26, Samtools v1.3.1, Trinity v2.6.6, VarScan v2.4.3, BioPython v1.78, numpy v1.15.2, lme4 v1.1.21, poolfstat v1.1.1, qvalue v2.14.1, ACER v1.0.2, haplovalidate v0.1.4, BBTools v38, BEDOPS v2.4.39, Bowtie v2.3.5, Gowinda v1.12, HMMER v3.2.1, SLiM v3.7, RSEM v1.3.1, Transdecoder v5.5, Trimmomatic v0.39, TreeMix v1.13, https://github.com/jjberg2/PolygenicAdaptationCode, wtdbg v2.5, Racon v1.4.3, LiftOff v1.6.1, https://github.com/TheDBStern/Baltic_Lab_Wild (DOI:10.5281/zenodo.6615047)

These data have been deposited in NCBI under BioProject number PRJNA844002.

BCO-DMO Processing Description:
- Adjusted field/parameter names to comply with BCO-DMO naming conventions;
- Added a conventional header with dataset name, PI names, version date.

[table of contents | back to top]

Data Files

File	
lab_samples-1.csv	(Comma Separated Values (.csv), 1.48 KB) MD5:5e0934373592dbc79af453914754dd13
Primary data file for dataset ID 878335	

[table of contents | back to top]

Related Publications

Juanitadiaz, & Stern, D. B. (2022). TheDBStern/Baltic_Lab_Wild: First release (Version v0.0.1) [Computer software]. Zenodo. https://doi.org/10.5281/ZENODO.6615047 https://doi.org/10.5281/zenodo.6615047

Software

Stern, D. B., Anderson, N. W., Diaz, J. A., & Lee, C. E. (2022). Genome-wide signatures of synergistic epistasis during parallel adaptation in a Baltic Sea copepod. Nature Communications, 13(1). https://doi.org/10.1038/s41467-022-31622-8
Results

[table of contents | back to top]

Related Datasets

IsRelatedTo

Lee, C. E., Diaz, J. A., Stern, D. B. (2023) **Pool-seq data from wild populations of copepods in the North Sea from May 2014 (Evolutionary genomics of a copepod project).** Biological and Chemical Oceanography Data Management Office (BCO-DMO). (Version 1) Version Date 2023-06-27 http://lod.bco-dmo.org/id/dataset/897977 [view at BCO-DMO]

Lee, C. E., Stern, D. B. (2022) **Pool-seq data from wild populations of copepods in the Baltic Sea from May 2018 through August 2019.** Biological and Chemical Oceanography Data Management Office (BCO-DMO). (Version 1) Version Date 2022-08-11 doi:10.26008/1912/bco-dmo.878322.1 [view at BCO-DMO]

University of Wisconsin - Madison. Genome-wide signatures of synergistic epistasis during parallel adaptation in a Baltic Sea copepod. 2022/05. In: BioProject [Internet]. Bethesda, MD: National Library of Medicine (US), National Center for Biotechnology Information; 2011-. Available from: http://www.ncbi.nlm.nih.gov/bioproject/PRJNA844002. NCBI:BioProject: PRJNA844002.

[table of contents | back to top]

Parameters

Parameter	Description	Units
Sample_Code	unique identifier for sample	unitless
Line_Beaker	Laboratory line	unitless
Collection_Year	Year of sample collection	unitless
Salinity	Salinity measured with handheld refractometer	PSU
Generation	Generation number after start of experiment	unitless
BioSample	NCBI BioSample	unitless
SRA_Run	NCBI SRA Run number	unitless

[table of contents | back to top]

Instruments

Dataset-specific Instrument Name	Illumina Hi-Seq 4000 and Illumina NovaSeq 6000
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	Hand-held refractometer
Generic Instrument Name	Refractometer
Generic Instrument Description	A refractometer is a laboratory or field device for the measurement of an index of refraction (refractometry). The index of refraction is calculated from Snell's law and can be calculated from the composition of the material using the Gladstone-Dale relation. In optics the refractive index (or index of refraction) n of a substance (optical medium) is a dimensionless number that describes how light, or any other radiation, propagates through that medium.

[table of contents | back to top]

Project Information

Evolutionary Responses to Global Changes in Salinity and Temperature (Evolutionary genomics of a copepod)

Coverage: St. Lawrence estuary, Gulf of Mexico, Great Lakes, Baltic Sea

NSF Award Abstract:

Drastic changes in the global water cycle and increases in ice melt are causing the freshening of Northern coastal seas. The combination of both reduced salinity and increased temperature will likely act in concert to reduce populations of estuarine and marine organisms. Data indicate that reduced salinity and high temperature would each increase the energy costs as well as reduce survival and reproduction of the common copepod Eurytemora affinis. This project will examine the joint effects of salinity reduction and temperature increase on the evolutionary responses of populations of E. affinis in the wild, as well as in selection experiments in the laboratory. This study will provide novel insights into responses of organisms to climate change, as no study has analyzed the joint impacts of salinity and temperature on evolutionary responses, and relatively few studies have examined the impacts of declining salinity. In general, how selection acts at the whole genome level is not well understood, particularly for non-model organisms. As a dominant estuarine copepod, E. affinis is among the most important species sustaining coastal food webs and fisheries in the Northern Hemisphere, such as salmon, herring, and anchovy. Thus, insights into its evolutionary responses with changing climate have important implications for sustainability of fisheries and food security. Two graduate students from historically underrepresented groups will be trained during this project. The project will have additional societal benefits, including development of educational modules for K-12 students and international collaboration.

This study will address the following questions: (1) To what extent could populations evolve in response to

salinity and temperature change, and what are the fitness and physiological costs? (2) How will populations respond to the impacts of salinity-temperature interactions? (3) Do wild populations show evidence of natural selection in response to salinity and temperature? To analyze the evolutionary responses of E. affinis populations to the coupled impacts of salinity and temperature, the investigator will perform laboratory selection experiments and population genomic surveys of wild populations. Selection experiments constitute powerful tools for determining the rate, trajectory, and limits of adaptation. During laboratory selection, evolutionary shifts in fitness-related traits and genomic expression will be examined, as well as genomic signatures of selection in response to low salinity and high temperature selection regimes. The investigator will also conduct population genomic sequencing of E. affinis populations that reside along salinity and temperature gradients in the St. Lawrence and Baltic Sea, and identify genes that show signatures of selection. The project will determine whether the loci that show signatures of selection in the wild populations are the same as those favored during laboratory selection. This reproducibility will provide greater confidence that the genes involved in adaptation to salinity and/or temperature have been captured.

[table of contents | back to top]

Funding

Funding Source	Award
NSF Division of Ocean Sciences (NSF OCE)	OCE-1658517

[table of contents | back to top]