

# Sample information and genetic accession information for raw low-coverage genomic sequence reads from 248 different Atlantic silverside (*Menidia menidia*) collected along the east coast of North America between 2005 to 2007

**Website:** <https://www.bco-dmo.org/dataset/854878>

**Data Type:** Other Field Results

**Version:** 1

**Version Date:** 2021-06-30

## Project

» [Collaborative research: The genomic underpinnings of local adaptation despite gene flow along a coastal environmental cline](#) (GenomAdapt)

Contributors	Affiliation	Role
<a href="#">Therkildsen, Nina Overgaard</a>	Cornell University (Cornell)	Principal Investigator
<a href="#">Baumann, Hannes</a>	University of Connecticut (UConn)	Co-Principal Investigator
<a href="#">York, Amber D.</a>	Woods Hole Oceanographic Institution (WHOI BCO-DMO)	BCO-DMO Data Manager

## Abstract

Sample information and genetic accession information for raw low-coverage genomic sequence reads from 248 different Atlantic silverside individuals (some individuals are represented by multiple sequence data files). The raw low-coverage genomic sequence reads are deposited in the National Center for Biotechnology Information (NCBI) Sequence Read Archive (SRA) under BioProject PRJNA376564. These data were published in Wilder et al. (2020).

## Table of Contents

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

## Coverage

**Spatial Extent:** N:47.4 E:-61.85 S:31.02 W:-81.43

**Temporal Extent:** 2005 - 2007

## Methods & Sampling

The following is an excerpt of the methods used to generate these data. Full details can be found in the associated publication and supplemental material of Wilder et al. (2020).

## Data generation:

Atlantic silverside muscle samples collected from Georgia (GA), New York (NY), Gulf of Maine (GoM) and Gulf of St. Lawrence (GSL) were sequenced by Therkildsen & Palumbi (2017) to a final, average depth of 1.5x across the reference transcriptome. All the samples were collected in the spring. See Therkildsen & Palumbi (2017) and Therkildsen et al. (2019) for further detail. We added 47 additional samples from Oregon Inlet, North Carolina (NC) near Cape Hatteras.

Individually barcoded sample libraries were prepared at Cornell's Biotechnology Resource Center using similar methods to Therkildsen & Palumbi (2017), with a few minor modifications. Reagents from the Illumina Nextera kit (96 sample Nextera DNA Library Prep Kit) were used at 1/3 the recommended concentration in 1/10 the recommended volume (5ul instead of 50ul), with 2ng of input DNA. Individual libraries were pooled, and size-selected using a Pippin Prep to remove fragments <286 bp (150 bp insert plus 136 bp of Illumina adapters). Filtered mapped reads for the NC samples had a final average depth of 1.5x.

### Sampling and analytical procedures:

Samples were collected directly from the wild as described in Hice et al. (2012).

### Locations:

Five locations along the east coast of North America.

USA:Jekyll Island,Georgia 31.02 -81.43

USA:Oregon Inlet,NorthCarolina 35.77 -75.52

USA:Minas Basin,Gulf of Maine 45.2 -64.38

USA:Patchogue,NewYork 40.75 -73.00

Canada:Magdalen Island, Gulf of St. Lawrence 47.4 -61.85

### Problem report:

The following samples should be excluded from analysis because of possible errors during library preparation: Jekylls\_1223, Jekylls\_1224, MagdalenIs\_911, MagdalenIs\_912, MagdalenIs\_913, MagdalenIs\_914, MagdalenIs\_915, MagdalenIs\_916, MagdalenIs\_917, MagdalenIs\_918, Patchogue\_1030

### Data Processing Description

BCO-DMO Data Manager Processing Notes:

- \* Converted SampleLocationDetails.xlsx Sheet 1 to csv and imported into the BCO-DMO data system.
- \* Split latitude and longitude into respective columns from combined string column.
- \* replaced comma in geolocation name to semicolon for better support for csv exports.

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

---

### Data Files

File
<b>854878_v1_raw_icwgs_sampleinfo.csv</b> (Comma Separated Values (.csv), 25.53 KB) MD5:9bb098ccb09b3113e0fc3dd587a536fd
Primary data file for dataset ID 854878, version 1

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

---

### Related Publications

Hice, L. A., Duffy, T. A., Munch, S. B., & Conover, D. O. (2012). Spatial scale and divergent patterns of variation in adapted traits in the ocean. *Ecology Letters*, 15(6), 568–575. doi:[10.1111/j.1461-0248.2012.01769.x](https://doi.org/10.1111/j.1461-0248.2012.01769.x)  
*Methods*

Therkildsen, N. O., & Palumbi, S. R. (2017). Practical low-coverage genomewide sequencing of hundreds of individually barcoded samples for population and evolutionary genomics in nonmodel species. *Molecular*

Ecology Resources, 17(2), 194–208. doi:[10.1111/1755-0998.12593](https://doi.org/10.1111/1755-0998.12593)  
Methods

Therkildsen, N. O., Wilder, A. P., Conover, D. O., Munch, S. B., Baumann, H., & Palumbi, S. R. (2019). Contrasting genomic shifts underlie parallel phenotypic evolution in response to fishing. *Science*, 365(6452), 487–490. doi:[10.1126/science.aaw7271](https://doi.org/10.1126/science.aaw7271)  
Methods

Wilder, A. P., Palumbi, S. R., Conover, D. O., & Therkildsen, N. O. (2020). Footprints of local adaptation span hundreds of linked genes in the Atlantic silverside genome. *Evolution Letters*, 4(5), 430–443. doi:[10.1002/evl3.189](https://doi.org/10.1002/evl3.189)  
Results

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

---

## Related Datasets

### IsRelatedTo

Cornell University (2017). *Menidia menidia* Raw sequence reads. 2017/02. NCBI:BioProject: PRJNA376564. Bethesda, MD: National Library of Medicine (US), National Center for Biotechnology Information; Available from: <http://www.ncbi.nlm.nih.gov/bioproject/PRJNA376564>.

Therkildsen, N. O., Baumann, H. (2021) **Methodology information and links to data access for allele frequencies and FST estimates for 1,904,119 SNPs analyzed in five population samples of Atlantic silverside (*Menidia menidia*) collected along the east coast of North America between 2005 to 2007**. Biological and Chemical Oceanography Data Management Office (BCO-DMO). (Version 1) Version Date 2021-06-29 <http://lod.bco-dmo.org/id/dataset/854895> [[view at BCO-DMO](#)]  
*Relationship Description: The raw low-coverage whole genome sequencing reads are the raw data used to generate the allele frequencies and FST estimates (for transcriptome regions only).*

Therkildsen, N. O., Baumann, H. (2024) **Sample and genetic accession information for RNA-seq data from whole Atlantic silverside (*Menidia menidia*) larvae from two populations and their F1 hybrids reared under different temperatures in 2017**. Biological and Chemical Oceanography Data Management Office (BCO-DMO). (Version 1) Version Date 2021-06-29 doi:10.26008/1912/bco-dmo.854887.1 [[view at BCO-DMO](#)]  
*Relationship Description: The "Raw low-coverage whole genome sequencing reads" are from population samples from five locations, including the two populations studied in the RNA-seq dataset, but the individuals are not related (the fish used for low-coverage whole genome sequencing were sampled ~10 years before the fish used in the RNA-seq study).*

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

---

## Parameters

Parameter	Description	Units
SampleID	Sample identifier	unitless
bioproject_accession	NCBI BioProject accession number	unitless
biosample_accession	NCBI BioSample accession number	unitless
Population	Description of the population sampled by region name	unitless
collection_date	Collection year in format yyyy	unitless
geo_loc_name	Geolocation name of the collection	unitless
lat	Collection latitude	decimal degrees
lon	Collection longitude	decimal degrees

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

---

## Instruments

<b>Dataset-specific Instrument Name</b>	Illumina HiSeq 2000
<b>Generic Instrument Name</b>	Automated DNA Sequencer
<b>Dataset-specific Description</b>	Illumina HiSeq 2000 with paired-end 125 bp (for samples from GA, NY, GoM, and GSL)
<b>Generic Instrument Description</b>	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.

<b>Dataset-specific Instrument Name</b>	Illumina NextSeq500
<b>Generic Instrument Name</b>	Automated DNA Sequencer
<b>Dataset-specific Description</b>	Illumina NextSeq500 paired-end 75 bp (for samples from NC)
<b>Generic Instrument Description</b>	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.

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

---

## Project Information

**Collaborative research: The genomic underpinnings of local adaptation despite gene flow along a coastal environmental cline (GenomAdapt)**

**Website:** <https://befel.marinesciences.uconn.edu/2018/03/07/research-news-new-nsf-grant-to-study-silverside-genes/>

**Coverage:** Eastern coastline of North America

NSF Abstract:

Oceans are large, open habitats, and it was previously believed that their lack of obvious barriers to dispersal would result in extensive mixing, preventing organisms from adapting genetically to particular habitats. It has recently become clear, however, that many marine species are subdivided into multiple populations that have

evolved to thrive best under contrasting local environmental conditions. Nevertheless, we still know very little about the genomic mechanisms that enable divergent adaptations in the face of ongoing intermixing. This project focuses on the Atlantic silverside (*Menidia menidia*), a small estuarine fish that exhibits a remarkable degree of local adaptation in growth rates and a suite of other traits tightly associated with a climatic gradient across latitudes. Decades of prior lab and field studies have made Atlantic silverside one of the marine species for which we have the best understanding of evolutionary tradeoffs among traits and drivers of selection causing adaptive divergence. Yet, the underlying genomic basis is so far completely unknown. The investigators will integrate whole genome sequencing data from wild fish sampled across the distribution range with breeding experiments in the laboratory to decipher these genomic underpinnings. This will provide one of the most comprehensive assessments of the genomic basis for local adaptation in the oceans to date, thereby generating insights that are urgently needed for better predictions about how species can respond to rapid environmental change. The project will provide interdisciplinary training for a postdoc as well as two graduate and several undergraduate students from underrepresented minorities. The findings will also be leveraged to develop engaging teaching and outreach materials (e.g. a video documentary and popular science articles) to promote a better understanding of ecology, evolution, and local adaptation among science students and the general public.

The goal of the project is to characterize the genomic basis and architecture underlying local adaptation in *M. menidia* and examine how the adaptive divergence is shaped by varying levels of gene flow and maintained over ecological time scales. The project is organized into four interconnected components. Part 1 examines fine-scale spatial patterns of genomic differentiation along the adaptive cline to a) characterize the connectivity landscape, b) identify genomic regions under divergent selection, and c) deduce potential drivers and targets of selection by examining how allele frequencies vary in relation to environmental factors and biogeographic features. Part 2 maps key locally adapted traits to the genome to dissect their underlying genomic basis. Part 3 integrates patterns of variation in the wild (part 1) and the mapping of traits under controlled conditions (part 2) to a) examine how genomic architectures underlying local adaptation vary across gene flow regimes and b) elucidating the potential role of chromosomal rearrangements and other tight linkage among adaptive alleles in facilitating adaptation. Finally, part 4 examines dispersal - selection dynamics over seasonal time scales to a) infer how selection against migrants and their offspring maintains local adaptation despite homogenizing connectivity and b) validate candidate loci for local adaptation. Varying levels of gene flow across the species range create a natural experiment for testing general predictions about the genomic mechanisms that enable adaptive divergence in the face of gene flow. The findings will therefore have broad implications and will significantly advance our understanding of the role genomic architecture plays in modifying the gene flow - selection balance within coastal environments.

This award reflects NSF's statutory mission and has been deemed worthy of support through evaluation using the Foundation's intellectual merit and broader impacts review criteria.

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

---

## Funding

Funding Source	Award
<a href="#">NSF Division of Ocean Sciences (NSF OCE)</a>	<a href="#">OCE-1756316</a>
<a href="#">NSF Division of Ocean Sciences (NSF OCE)</a>	<a href="#">OCE-1756751</a>

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