NCBI accession numbers and associated metadata for raw low-coverage genomic sequence reads from 876 different Atlantic silverside individuals

Website: https://www.bco-dmo.org/dataset/754623
Data Type: Other Field Results, experimental
Version: 1
Version Date: 2019-02-04

Project
» High resolution genome changes during evolution in a classic fisheries experiment (Fishery Genome Changes)

Contributors	Affiliation	Role
Palumbi, Stephen R.	Stanford University - Hopkins (Stanford-HMS)	Principal Investigator
Therkildsen, Nina Overgaard	Cornell University (Cornell)	Contact
Rauch, Shannon	Woods Hole Oceanographic Institution (WHOI BCO-DMO)	BCO-DMO Data Manager

Abstract
NCBI accession numbers and associated metadata for raw low-coverage genomic sequence reads from 876 different Atlantic silverside individuals

Table of Contents

Coverage

Spatial Extent: **N**:47.4 **E**:-61.85 **S**:31.02 **W**:-81.43
Temporal Extent: 1998 - 2007

Dataset Description

Raw low-coverage genomic sequence reads: 1190 files with raw low-coverage genomic sequence reads from 876 different Atlantic silverside individuals (some individuals are represented by multiple sequence data files) are deposited in the NCBI Sequence Read Archive (SRA) with accession numbers SRS2001281- SRS2003099 associated with BioProject PRJNA376564 and BioSamples SAMN06436077 - SAMN06436952.

Methods & Sampling

Species: Atlantic silverside (*Menidia menidia)*
Source: From laboratory breeding experiment or wild caught at one of four locations along the east coast of

North America.
Library preparation method: Modified version of Illumina Nextera sample preparation kit.
Sequencing instrument: Illumina HiSeq 2000
Read type: Paired-end 125 bp
Sequencing strategy: Whole genome sequencing

The sequence data were prepared as described in Therkildsen & Palumbi (2017).

The tissue samples used for sequencing had been stored in minus 20 degrees C freezer for 8-17 years prior to DNA extraction. We used the Qiagen DNeasy Blood and Tissue kit to extract DNA from muscle tissue and evaluated the degradation level of each extract through 1.5% agarose gel electrophoresis. Only samples that showed clear high molecular weight bands and limited smearing were retained for library preparation. To ensure DNA integrity in the retained samples, we removed fragments shorter than ~1000 bp from each extract using Agencourt AMPure XP beads in a 0.4:1 AMPure to sample ratio and eluted the DNA in 10 mM Tris-Cl, pH 8.5.

We measured DNA concentrations with a Quant-iT high-sensitivity assay (Invitrogen) and prepared a separate barcoded library for each individual with Illumina's Nextera kit according to the protocol described in Therkildsen and Palumbi (2017). Briefly, the tagmentation reaction, which simultaneously fragments the DNA and incorporates partial adapters, was carried out in a 2.5 ul volume with 1.6–7.9 ng of input DNA for each library. We then used a two-step PCR procedure with a total of 12 cycles (8 + 4) to add the remaining Illumina adapter sequence with dual index barcodes and amplify the libraries. The PCR was conducted with the KAPA Library Amplification Kit and the Illumina Nextera index kit with primers N501– N508 + S511 and N701–N712 + N714. As a final step, we purified and size-selected the amplification products with Agencourt AMPure XP beads and quantified the concentration of the final libraries with the Invitrogen Quant-iT high-sensitivity assay. We also examined the fragment size distribution of multiple libraries from each plate on an Agilent Bioanalyzer instrument.

We combined equimolar amounts of 56–76 libraries into separate pools for sequencing across 13.5 lanes of paired-end 125-bp reads on an Illumina HiSeq 2000 (v4 chemistry) at the University of Utah's Bioinformatics Core Facility. To even out the data yield among samples, we repooled libraries that initially had obtained the lowest read output for supplementary sequencing in 4.5 additional HiSeq lanes.

Data Processing Description

The following samples should be excluded from analysis because of possible errors during library preparation:
D1Gen10_1496
D1Gen10_748
D1Gen10_749
D1Gen10_755
D1Gen10_767
D1Gen10_774
D1Gen10_776
D1Gen10_778
D1Gen10_780
D1Gen5_250
D1Gen5_332
D2Gen10_425
D2Gen10_445
JekyllIs_1223
JekyllIs_1224
MagdalenIs_911
MagdalenIs_912
MagdalenIs_913
MagdalenIs_914
MagdalenIs_915
MagdalenIs_916
MagdalenIs_917
MagdalenIs_918
Patchogue_1030
R1Gen5_1357

RGen0_701
U2Gen10_1199
U2Gen10_1200
U2Gen10_1201
U2Gen10_1479

[table of contents | back to top]

Data Files

File
raw_low_coverage_accessions.csv (Comma Separated Values (.csv), 95.60 KB) MD5:b8fada5d39a5c085106fe3674c5793ed Primary data file for dataset ID 754623

[table of contents | back to top]

Related Publications

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
Results
,
Methods

[table of contents | back to top]

Parameters

Parameter	Description	Units
BioSample_Accession	NCBI BioSample Accession number	unitless
BioProject_Accession	NCBI BioProject Accession number	unitless
SRA_Accession	NCBI SRA Accession number	unitless
Species	Species name	unitless
Sample_ID	Sample ID number	unitless
Breed	Laboratory breed or wild caught	unitless
Collection_Site	Collection site name	unitless
Latitude	Latitude of collection site; positive values = North.	decimal degrees
Longitude	Longitude of collection site; positive values = East.	decimal degrees
Collection_Year	Year of collection (YYYY)	unitless

[table of contents | back to top]

Instruments

Dataset-specific Instrument Name	Illumina HiSeq 2000
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.

[table of contents | back to top]

Deployments

Palumbi_silversides

Website	https://www.bco-dmo.org/deployment/687817
Platform	Poquott Beach
Description	Atlantic silverside (Menidia menidia) specimens collected from Poquott Beach (NY), Jekyll Island (GA), Patchogue (NY), Minas Basin (Nova Scotia), and Madalen Islands (Quebec).

[table of contents | back to top]

Project Information

High resolution genome changes during evolution in a classic fisheries experiment (Fishery Genome Changes)

Coverage: The East coast of North America

Description from NSF award abstract:
One of the strongest impacts there is on ocean species is reduction in population size due to fishing. Even for species that are not overfished, harvest takes a large fraction of the biggest and fastest growing individuals. As a result, fishing exerts strong natural selection on fish population, selecting for slow growing, small individuals. Experiments in artificial fishing have shown rapid evolution of growth rates, maturation size and other traits for lab populations. The classic Conover-Munsch experiments ten years ago showed the power of fishing to generate rapid evolution. However, no analysis of the genetic impact of fishing under such controlled conditions has been done, and no investigation of the way whole genomes respond to strong fisheries evolution has been attempted. Luckily, the Conover-Munsch samples - from the fish used in their classic experiment - have been preserved, and modern genomic techniques are now available that can analyze the way fisheries-induced evolution shaped the genetic diversity and genome architecture of these populations. Strong selection is known in other systems to leave a legacy of deleterious changes in the genome. The results of this study will be important components of understanding the long-term effects of fishing because they will for the first time allow a mechanistic understanding of how natural selection works on fish populations. The project will also compare the changes that occur after the relaxation of fishing pressure to estimate if there is a legacy of deleterious genome changes in fished species that impedes their recovery. The data from this study will show how fishing creates change at loci under selection, and also how this strong selection generates other, non-adaptive shifts because of genetic hitchhiking and inbreeding.

The investigators will use next generation DNA sequencing to sequence the protein coding regions of the

Conover-Munsch fish samples. They will discover, document and compare genetic variants across the genome in lines selected for large size, in lines selected for small size, and in the original populations in order to chart evolutionary changes at the genomic level imposed by fishing. They will use outlier analyses to pinpoint loci at which strong selection has acted to change allele frequencies. They will compare these changes to changes in fish populations after the relaxation of fishing in order to distinguish evolutionary changes that are easily reversible from those that are not. In addition, they will also compare genetic changes induced by fishing to those that occur naturally along an environmental gradient on the US east coast. Preliminary data indicate that many of the genetic variants selected for in the fisheries experiment are old variants - estimated by patterns of linkage disequilibrium - already present in the original population. Working on the genetics of the natural gradient will show which of these variants have been selected by evolutionary forces in the native environment, and subsequently been favored by the novel evolutionary pressure of fishing.

[table of contents | back to top]

Funding

Funding Source	Award
NSF Division of Ocean Sciences (NSF OCE)	OCE-1434325

[table of contents | back to top]