DRAFT: Annotated genome of the Atlantic dog whelk, Nucella lapillus

Website: https://www.bco-dmo.org/dataset/991343
Version: 1
Version Date: 2026-01-13

Project
» Local adaptation and the evolution of plasticity under predator invasion and warming seas: consequences for individuals, populations and communities (evolution of plasticity)

Contributors	Affiliation	Role
Trussell, Geoffrey C.	Northeastern University	Principal Investigator
Vollmer, Steven V.	Northeastern University	Scientist
Ford, Meghan	Northeastern University	Student
Gerlach, Dana Stuart	Woods Hole Oceanographic Institution (WHOI BCO-DMO)	BCO-DMO Data Manager

Abstract
Nucella lapillus has been a focal organism of ecological and evolutionary studies on rocky shores for decades. As a direct developer, this species has limited dispersal but is broad geographic range makes it ideal for studies of population structure, isolation by distance, and selection across different environmental gradients. The fully annotated genome generated ~37X coverage and the genome assembly is 2.32 Gbp. Genomic resources for mollusks are relatively limited compared to other phyla and this new genome will enhance our understanding of genomic variation in mollusks and studies seeking to link genomic variation with organismal performance and community-level processes under various dimensions of environmental change.

Table of Contents

Coverage

Location: Laboratory extraction of whelk specimen from Nahant, MA (42.419732, −70.902171)
Spatial Extent: **Lat**:42.419732 **Lon**:-70.902171
Temporal Extent: 2024-05 - 2024-05

Dataset Description

Nucella lapillus is an important player in rocky shore food chains and has been a focal organism of ecological and evolutionary studies for decades. Despite poor dispersal, they have a broad geographic range, which makes them an ideal species to examine isolation by distance and selection across environmental gradients. A fully annotated genome of *N. lapillus* generated with Oxford Nanopore Techonology (ONT) sequencing at ~37× coverage was described in detail in the following publication:

Ford, M. R., Vollmer, S. V., & Trussell, G. C. (2025). Annotated genome of the Atlantic dog whelk, Nucella lapillus. G3: Genes, Genomes, Genetics, 15(10). https://doi.org/10.1093/g3journal/jkaf182

This dataset provides information on the study and includes links to the genetic dataset at the National Center for Biotechnology Information (NCBI).

Methods & Sampling

Sample collection and DNA extraction
An adult Nucella lapillus individual was collected in May 2024 from Nahant, MA (42.419732, −70.902171). The foot tissue was used for DNA extraction and isolation.

High molecular weight (HMW) DNA was extracted from the foot via the CTAB method (1.4 M NaCL and 2% CTAB), followed by 3 chloroform reactions. The full extraction protocol is available on the project GitHub repository (https:// github.com/meghanclownfish/Nucella-lapillus-genome/tree/main/ 1_extraction). The HMW DNA was precipitated in ethanol (EtOH) and resuspended in Tris-EDTA buffer. The sample was further purified using a Genomic DNA Clean and Concentrator kit (gDCC-10, ZYMO Research, Irvine, CA, USA) per manufacturer's instructions. Sample quality and concentration were assessed by running 2 µL of the sample on Nanodrop (Thermo Fisher Scientific, Singapore). A sample was deemed ready for sequencing if the 260/280 ratio was ~1.9 and the 260/230 ratio fell between 2.0 and 2.2, following Sun et al. 2020.

Library prep and sequencing
N. lapillus long reads were sequenced using ONT platforms and libraries were prepared with the ONT Ligation sequencing kit (SQK-LSK114, ONT, Oxford, UK) and NEBNext Companion Module (E7180S NEB). Standard manufacturer's protocol was implemented with a few exceptions (see Ford et al., 2025 for details). Sequencing was done on a PromithION. Six flow cells in total were used to generate 103,553,219,099 bp raw data. PromethION flow cells (FLO-PRO114M) were primed and loaded per the standard manufacturer's protocol. To increase the yield of each flow cell, runs were paused and flushed using the EXP-WSH004 (ONT) kit and reloaded. High quality base calling was performed with Dorado 0.7.1 (ONT).

Genome size and heterozygosity
Genome size was estimated using JELLYFISH v2.2.10 to count canonical 41-mers from high quality ONT reads (min quality: 5) and computed a histogram of k-mer occurrence (Marcais and Kingsford 2011). The histogram was used to estimate heterozygosity with GenomeScope (Ranallo-Benavidez et al. 2020).

Data Processing Description

Assembly and annotation
Briefly, we assembled the genome using all reads of 2kb in length or greater with Hifiasm (0.25.0-r726). BlobTools2 (v4.4.0, Challis et al. 2020) was used to visually assess the assembly and filter contigs. RepeatModeler (v2.0.6, Flynn et al. 2020) and RepeatMasker (Smit et al.) identified and soft-masked repetitive regions in the genome. RNASeq data was mapped to the soft-masked genome with HISAT2. This information, along with a custom protein database, was supplied as evidence for Braker3 . TSEBRA was used to merge Braker outputs. Functional annotation was carried out with InterProScan and Funannotate. For full details, please see Ford et al. (2025) paper.

[ ]

Parameters

Parameters for this dataset have not yet been identified

[ ]

Project Information

Local adaptation and the evolution of plasticity under predator invasion and warming seas: consequences for individuals, populations and communities (evolution of plasticity)

NSF Award Abstract:
Over the past two decades, the Gulf of Maine has experienced unprecedented warming that, among other things, has further enabled the invasive green crab to expand its range in rocky shore habitats. The adverse ecological impacts of this invasive predator have been documented worldwide. This study examines how

geographic variation in the capacity of two common prey species to respond to the combination of this predator and warming ocean temperatures can shape prey feeding and performance and impact community structure and dynamics. Hence, this research enhances understanding of the evolution of phenotypes, their plasticity, and the nature of adaptation and its role in eco-evolutionary dynamics. More broadly, it informs understanding of how organisms and marine communities may respond to future environmental change. In addition, this project makes contributions to the STEM pipeline by providing middle and high school, undergraduate, and graduate students with cross-disciplinary training in evolutionary and community ecology. In collaboration with an institutional outreach program, the investigator is also developing web-based multimedia projects and teacher resource materials based on this research.

A central principle in ecology is that species residing in the middle of food chains must balance the benefits of eating with the risk of being eaten by their predators. Solving this foraging-predation risk trade-off often involves plasticity in prey traits with consequences for the evolution of adaptation and species interactions that drive community-level processes. Hence, the foraging-predation risk trade-off provides a powerful conceptual framework that links evolutionary and community ecology. Yet at the same time, other environmental stressors like temperature can shape this trade-off, adding complexity that makes it difficult to predict the capacity of organisms to adapt to environmental change and the consequences for communities. The investigator is conducting this study in rocky shore habitats of the Gulf of Maine (GOM) which have long been influenced by strong latitudinal temperature gradients and non-native species invasions. The overarching hypothesis is that predation risk and temperature are factors shaping geographic variation in plasticity and adaptation, with consequences for individuals, populations, and communities. First, the investigator is conducting field experiments to document geographic variation in the trait plasticity of two common prey species in the green crab's diet. Second, he is using reciprocal transplant experiments to examine trait plasticity in response to risk and water temperature, generating data to compare with similar experiments conducted in the late 90s prior to recent ocean warming and expansion in range of green crabs. Third, he is conducting a laboratory common garden experiment to evaluate the effects of risk and water temperature on trait plasticity. Finally, he is using reciprocal transplant experiments in the field to understand the interactive effects of risk and water temperature on prey foraging rates and the abundance of a species that plays an important role in intertidal community structure and dynamics.

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.

Funding

Funding Source	Award
NSF Division of Ocean Sciences (NSF OCE)	OCE-2017626