

Eastern oyster (*Crassostrea virginica*) mortality, growth, and tissue sample metadata from the common garden experiments run as part of CviMVP project from 2023-2025 in the Chesapeake Bay

Website: <https://www.bco-dmo.org/dataset/1007522>

Data Type: Other Field Results, experimental

Version: 1

Version Date: 2026-09-10

Project

» [CAREER: Evaluation of machine learning algorithms for understanding and predicting adaptation to multivariate environments with a Model Validation Program \(MVP\)](#) (Model Validation Program)

Contributors	Affiliation	Role
Lotterhos, Katie	Northeastern University	Principal Investigator
Small, Jessica	The College of William & Mary	Co-Principal Investigator
Carnegie, Ryan	Virginia Institute of Marine Science (VIMS)	Scientist
Eppley, Madeline	Northeastern University	Scientist
Katsuki, Shelley	Virginia Institute of Marine Science (VIMS)	Scientist
Rumberger, Camille	Northeastern University	Scientist
Bajaj, Kiran	Northeastern University	Student
Mongillo, Nicole	Northeastern University	Student
Segnitz, Zea	Northeastern University	Student
York, Amber D.	Woods Hole Oceanographic Institution (WHOI BCO-DMO)	BCO-DMO Data Manager

Abstract

This dataset includes Eastern oyster (*Crassostrea virginica*) mortality, growth, and tissue sample metadata from the common garden experiments described below. This project is related to the CviMVP (Model Validation Program) project, which aims to assess Machine Learning Algorithms (MLAs) for understanding and predicting adaptation of organisms to multivariate environments from their DNA sequences. The data here is related to a field experiment validating the results of MLAs predicting responses to common garden conditions at two sites in the Chesapeake Bay, Lewisetta and York River, and includes metadata on experimental setup, metadata on experimental individuals, and experimental mortality and growth. Briefly, the experiment involved spawning oysters from two selection lines and six wild populations, and creating spawning treatments including monocultures of all these groups as well as two polyculture treatments. Juvenile oysters were then deployed at two common gardens at Lewisetta, VA and York River, VA and monitored over the course of two years. In depth mortality and phenotype data was collected every six months and used to assess fitness of individuals from different spawn treatments across the two gardens.

Table of Contents

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

Coverage

Location: Chesapeake Bay

Dataset Description

Organism identifiers:

Eastern oyster, *Crassostrea virginica*, urn:lsid:marinespecies.org:taxname:140657

Methods & Sampling

Experimental Design: Adult oysters from six wild populations and two selection lines were spawned to create juvenile oysters, which were then deployed for two years as two common gardens in the Chesapeake Bay: one location was in the Coan River, VA, herein “Lewisetta,” and the other in the York River, VA. Lewisetta is characterized by relatively low salinity (average 8-15 ppt) and moderate prevalence of the oyster parasites Dermo and MSX (Frank-Lawale et al. 2014). The York River, in contrast, has moderate salinity (average 15-23 ppt) and higher disease pressure. At both sites, mean daily temperatures range from 5°C to 30°C, with freezing common in the intertidal zone in winter months.

Experimental Population Sourcing: Between September and November of 2022, 50-80 oysters were collected from six sites distributed along the eastern oyster’s native range then shipped live to the Virginia Institute of Marine Science’s Aquaculture Genetics & Breeding Technology Center (ABC) where the animals were held in a quarantine system until late winter 2023, and subsequently conditioned in a closed -loop system for approximately eight weeks prior to spawning. Two additional proprietary ABC broods stock lines, LOLA and DEBY, were also used in this experiment. These lines have been selected for growth and survival at the common garden sites, and so we included them as populations that would in theory be adapted to local conditions. DEBY was established by consolidating several Delaware Bay lines (77%) with Mobjack Bay (14%) and Louisiana oysters (9%) over 10 years, and has been under selection for York River conditions for eight generations. LOLA was established later by introgressing Louisiana oysters (80%) with DEBY oysters (20%), and has been under selection to Lewisetta conditions for at least four generations (Ragone Calvo et al. 2003; Frank-Lawale et al. 2014; Puritz et al. 2022). In accordance with their selection history, DEBY has historically performed better in the moderate salinity in the York River and LOLA in the low salinity conditions in the Lewisetta (Roy and Kirchner 2000; Proestou et al. 2016).

Conditioning, Spawning, and Larval rearing: Oysters were held at ABC in a flow-through conditioning system. During conditioning, algal feed levels and temperature were manipulated to induce gamete production. In May 2023, parental oysters were measured prior to spawning. For each population, 10 males and 10 females were strip spawned according to standard hatchery protocol and crossed to make larval cultures. Larvae were reared in duplicated 60 L & 200L tanks containing aerated 1µm filtered seawater and maintained at 25-27°C. Larvae were fed a standard ration of live microalgae. At days 2, 7, and 14, the density of larval cultures were reduced to prevent overcrowding. Eyed larvae were set on microcultch in a downwelling system and reared until 1mm in shell length, then transitioned to an upwelling nursery system until 5-10mm in shell length.

Field Deployment & Monitoring: Experimental seed was deployed at two field sites, Lewisetta (37.98030, -76.46190) and York River (37.24728, -76.49937), in July 2023. For each experimental source population, approximately 1200 oysters per group were placed into each of six 4mm mesh oyster bags, three per common garden. Deployed bags were monitored every six months for bag-level survival and individual-level length (longest distance from shell hinge to shell edge). These data acted as proxies to track fitness through time. At each monitoring event, bags were removed from water for a standardized period of time and checked for survival. At the first monitoring event, the length and width of 40 oysters per bag were taken, then oysters were tagged with shellfish tags (Hallprint) held in place by coral epoxy for future identification. These tagged oysters were re-identified and re-measured at each subsequent monitoring event to track length through time. To reduce density-dependent growth, bags were thinned to a density of 225 oysters per bag in May 2024. To account for this thinning, survival at each monitoring timepoint (S_t) was calculated recursively as the number

of live oysters at that monitoring event ($alive_t$) divided by the number of alive and dead oysters at that time point ($(alive+dead)_t$), scaled by cumulative survival at the previous monitoring event (S_{t-1}). Note that for $t = 1$, $S_{t-1} = 1$.

*Eqn. 1: $S_t = alive_t / (alive_t + dead_t) * S_{t-1}$*

Data Processing Description

Survival at each monitoring timepoint (S_t) was calculated recursively as the number of live oysters at that monitoring event ($alive_t$) divided by the number of alive and dead oysters at that time point ($(alive+dead)_t$), scaled by cumulative survival at the previous monitoring event (S_{t-1}). Note that for $t = 1$, $S_{t-1} = 1$.

Problem Description

NA

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

Related Datasets

IsRelatedTo

Rumberger, C., Lotterhos, K., Small, J., Eppley, M., Carnegie, R., Mongillo, N., Bajaj, K., Segnitz, Z. (2026) **Environmental Data from the common garden experiments run as part of CviMVP project from 2023-2025**. Biological and Chemical Oceanography Data Management Office (BCO-DMO). (Version 1) Version Date 2026-09-10 <http://lod.bco-dmo.org/id/dataset/1007572> [[view at BCO-DMO](#)]
Relationship Description: Data from the same common garden experiments.

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

Parameters

Parameters for this dataset have not yet been identified

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

Project Information

CAREER: Evaluation of machine learning algorithms for understanding and predicting adaptation to multivariate environments with a Model Validation Program (MVP) (Model Validation Program)

Coverage: East coast of North America

NSF Award Abstract:

Environmental change can be rapid and involve multiple aspects of the environment changing at the same time, such as warming and increased disease pressure. Rapid environmental change threatens the productivity of aquaculture and crops on which humans depend. Predicting organisms' vulnerabilities to rapid and multifactor environmental change, however, is a major scientific challenge. A hurdle to addressing this challenge arises from the complex and non-intuitive ways that organisms adapt, through changes at the level of the DNA sequence, to many environmental stresses at the same time. Thus, there is a need for new approaches to understand and predict adaptation in multivariate environments. To address this need, this project integrates research and education with a Model Validation Program (MVP). The research is developing and evaluating

Machine Learning Algorithms (MLAs) for understanding and predicting adaptation of organisms to multivariate environments from their DNA sequences. To evaluate MLAs, this research combines both data simulation and an empirical test in the field with the Eastern Oyster, which provide important ecosystem services and support a multi-million dollar industry. For oysters, this research is studying how temperature, disease pressure, and salinity interact with evolutionary history to determine fitness in the field. This research advances efforts toward addressing the major scientific challenge of predicting adaptation in complex environments by integrating concepts across the frontiers of marine, evolutionary, and statistical sciences in a new way. Machine learning and model validation are not traditionally taught in the marine and environmental sciences, but are becoming increasingly relevant to these fields. As part of a broader education program, this research is developing MVP Learning Modules for high school students and undergraduates, which help students build the foundational knowledge they need to critically evaluate and apply models. Modules are being disseminated to hundreds of students in the greater Boston area and are being made available online for widespread use. The MVP mentoring program is training graduate students, undergraduates, and high school students in marine evolutionary ecology, statistical genomics, and machine learning. This research addresses a pressing societal need to more informatively match genotypes to environments for restoration, farming, and assisted gene flow efforts. Results are being disseminated to stakeholders in the oyster industry.

The goal of this research is to evaluate if MLAs, which can model non-linearities, can be used to understand and predict adaptation to multivariate environments under a wide range of scenarios. In Objective 1, the Principal Investigator (PI) is creating simulated datasets with different aspects of realism, and using them to evaluate and refine the MLAs. This novel set of simulations is studying genome evolution under high gene flow in complex, multivariate environments. In Objective 2, the PI is building on their expertise with the Eastern oyster to evaluate the MLAs in a field setting. The PI is first developing a comprehensive seascape genomic dataset and using it to train MLAs to predict an individual's multivariate environment based on a single nucleotide polymorphism genotype. Then, the PI is testing if the MLA prediction can predict the fitness of different genotypes from across the species range when raised in common garden field conditions. In Objective 3, the PI is integrating research and education by using the data obtained from Objs. 1 and 2 to develop a series of original "MVP Learning Modules" with interactive web apps for persons at different levels of understanding, using the relatable example of an oyster restoration project. This research lays the foundation for future studies by producing datasets that could become classical examples for developing and benchmarking innovative modeling approaches.

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
NSF Division of Ocean Sciences (NSF OCE)	OCE-2043905

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