

DNA metrics from minimally-lethal genotyping with eastern oysters (*Crassostrea virginica*) conducted in 2022 and 2023

Website: <https://www.bco-dmo.org/dataset/1005320>

Data Type: experimental

Version: 1

Version Date: 2026-08-17

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	Virginia Institute of Marine Science (VIMS)	Scientist
Leung, Elisabeth	Northeastern University	Student, Contact
Soenen, Karen	Woods Hole Oceanographic Institution (WHOI BCO-DMO)	BCO-DMO Data Manager

Abstract

Understanding how selection acts on individual genotypes often requires genotyping at different life stages and tracking their survival under experimental conditions. However, in mollusks, such as the Eastern oyster (*Crassostrea virginica*), collecting sufficient high-quality DNA without causing mortality remains challenging. While prior studies have explored different noninvasive genotyping methods, an optimal method hasn't been identified. Here, we tested different techniques by combining two minimally invasive accession techniques (relaxation, shell notching), four types of cell sampling (swab, mantle biopsy, hemolymph, and extrapallial fluid), and three preservation techniques (flash freezing, ethanol, and FTA card) using nine different treatment groups. We monitored mortality for 11 days after cell sampling and quantified DNA quality and quantity from the extracted samples. This dataset includes DNA quality and quantification results from the different treatments and can be used to evaluate the effects of sampling and preservation techniques on DNA quality and quantity.

Table of Contents

- [Coverage](#)
- [Dataset Description](#)
 - [Methods & Sampling](#)
 - [BCO-DMO Curation Notes](#)
- [Related Publications](#)
- [Parameters](#)
- [Instruments](#)
- [Project Information](#)
- [Funding](#)

Coverage

Location: Experiment was conducted at the Virginia Institute of Marine Science (part of William & Mary) in Gloucester Point, Virginia, USA.

Methods & Sampling

Juvenile oysters ranging from 30mm to 60mm were sourced from the Aquaculture Genetics and Breeding Technology Center (ABC) at the Virginia Institute of Marine Science (VIMS). Each oyster was cleaned using a 1% bleach solution, air-dried for an hour, and labeled with a shellfish tag (Hallprint FPN 8x4 mm) using coral epoxy (Instant Ocean HoldFast® Epoxy Stick) (Figure 1A). Shell width and length were measured for each oyster before the experiment (Kraeuter et al., 2007). To store the oysters before the experiment, the oysters were stored at 5°C for up to 48 hours before experimentation.

Treatment groups

We planned nine treatments to determine which protocol would have the best DNA yield with the lowest mortality (Figure 1 of Leung et al. 2026 (in press)). Each treatment had at least 20 individual oysters. Groups of 10 oysters per each treatment were placed into separate trays in the sea tables. The treatments and short names are described in Table 1 of Leung et al. 2026 (in press).

Control Groups: We established two control groups in our experiment: one to monitor mortality and the other to quantify the amount of DNA extracted.

To establish a baseline for DNA quantification and quality, we sampled mantle, adductor, and gill tissues from 30 individuals and placed them into 2 mL O-ring vials with 95% molecular-grade ethanol. We stored the samples temporarily in a -20°C freezer before transferring them to a -80°C freezer.

Accessing the oyster tissue within the shell

Two methods were used to enter the oyster shell: creating a notch in the shell or relaxing the oyster. Creating a notch allowed for the extraction of fluid that contains genetic material (e.g., hemolymph or EPF), while relaxing the oyster allowed for tissue swabbing or direct sampling of the tissue. *Notch Protocol:* A variable-speed dremel was used to make a small, V-shaped notch in the oyster shell, just large enough for the needle (BD 23G PrecisionGlide Needle) to fit in at different angles. The size of the notch depended on how much new shell an oyster had recently produced, but was typically a few mm in size. The notch was created on the right valve of the oyster, approximately two-thirds to the top of the shell. The needle, attached to a syringe (BD 1mL Syringe), was used to extract fluid as described below for the *Hemolymph Protocol* and the *EPF Protocol*. *Relaxation Protocol:* For the relaxation protocol, 2 L of 1µm filtered seawater was mixed with 100g of Epsom salt until all the salt dissolved following protocol by Proestou et al. (2025). A maximum of ten oysters were placed in each replicate 2L container for 16 hours before the cell sampling (see *Treatment Groups* for experimental design). Oysters were fully submerged in the Epsom salt solution and held at room temperature. During relaxation, containers were aerated using air stones and airlines to maintain oxygen levels.

Collection of genetic material

Cell collection for notched oysters: Cells were collected from notched oysters by sampling hemolymph or EPF fluids. For the preservation of fluids, we split the samples and placed half in liquid nitrogen and the other half in FTA cards (see below). *Hemolymph Protocol:* We inserted the needle through the notch diagonally down into the adductor muscle, which we identified through increased resistance. We collected 125 to 250 µL of hemolymph. *EPF Protocol:* We inserted the needle into the notch until it reached the mantle cavity of the shell and collected 80 to 500 µL of EPF from each oyster. If necessary, we removed and reinserted the needle at multiple angles for small oysters to obtain a minimum fluid amount of 80 µL.

Cell collection for relaxed oysters: Cells were collected from relaxed oysters by taking a mantle biopsy or swabbing the mantle tissue. *Mantle Biopsy Protocol:* We used forceps and scissors to remove a small piece (~10-20mg) of the mantle wrapped around the adductor muscle. We sampled the distal edge of the mantle tissue, approximately 2-3mm away from the visible edge of the adductor muscle.

After each sample was acquired, we sterilized the forceps and scissors using an ethanol lamp. *Swab Protocol:* Using an OmniSwap (Qiagen; WB100035), we swabbed in a circular motion, starting at the top layer of the mantle and moving down to the mantle around the adductor, then to the mantle's bottom layer, and then reversing this path. The motion was repeated five times for each sampled individual.

Preservation of genetic material: We used three preservation protocols for the genetic samples: ethanol, flash-freezing, and FTA cards. *Ethanol Protocol:* The mantle sampled from the relaxed oysters was preserved with 95% molecular-grade ethanol. It was temporarily stored in a -20°C freezer and then transferred to a -80°C freezer. *Flash-Freeze Protocol:* Liquid nitrogen was used to flash-freeze half of each EPF or hemolymph sample from the notched oysters. After the samples were flash-frozen, they were stored in a -80°C freezer. Additionally, the swabs used on the relaxed oysters were flash-frozen and then stored in a -80°C freezer. *FTA Cards Protocol:* The remaining half of the EPF and hemolymph samples were stored on QIAcard FTA Classic Cards (Qiagen; WB120305). Following the manufacturer's instructions, we put up to 125 µL of fluid on each circle in the FTA card. FTA cards were stored in WhirlPak bags with color-indicating silica beads (desiccants that absorb moisture and change color to indicate saturation) following the manufacturer's recommendations.

DNA extraction: A Qiagen DNeasy Blood & Tissue kit (Qiagen; 69504) was used for all DNA extractions, but with slight modifications for different cell types or preservation methods. The manufacturer's instructions were used for the control and the mantle biopsy groups with 20- 30mg of tissue, as well as the swabs (swabs were

horizontally cut in half before extraction so that they could be submerged fully in the buffer). We modified the protocol for the EPF and Hemolymph samples by using 200 µL of the flash-frozen sample as input (based on preliminary trials of 20, 100, and 200 µL that showed the latter amount yielded the most DNA) but otherwise followed the manufacturer’s instructions. For the Whatman FTA Cards, we used a 6.0mm biopsy punch (Cenmed™ PMD Sterile, Disposable Biopsy Punches) to cut a disk from the middle of the card. Per recommendation from Sigma Aldrich, 280 microliters of the ATL buffer was added to the disk for extraction, and the remainder of the manufacturer’s instructions were followed (*Reliable DNA Extraction from Whatman™ FTA Cards*, n.d.).

DNA quantification and quality assessment: We used a Qubit 3.0 Fluorometer to quantify DNA yield from the extractions. We used the High Sensitivity Kit (ThermoFisher; Q32851) for samples with a yield between 0 and 20 ng/ul and the Broad Range Kit (ThermoFisher; Q32850) for all samples with a yield higher than 20 ng/ul.

To test the quality of the DNA, we ran 3 µL of each DNA extract on a 1% agarose gel (0.6g of agarose with 60 mL of 1x TAE buffer), stained with GelRed at 130V for 30 minutes. Gels were visualized, and images were captured on a BioRad Gel Doc XR+ Molecular Imager machine. We categorized each sample on the gel into one of four levels: a bright band greater than 3000 bp with a slight smear (level 1), a completely smeared sample (level 2), a barely visible gel (level 3), or no band (level 4).

BCO-DMO Curation Notes

- Loaded file DNA_quantification_data.csv, treated empty strings, "nd", and "NA" as missing values
- Renamed table from dna_quantification_data to 1005320_dnametrics

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

Related Publications

KRAEUTER, J. N., FORD, S., & CUMMINGS, M. (2007). OYSTER GROWTH ANALYSIS: A COMPARISON OF METHODS. *Journal of Shellfish Research*, 26(2), 479–491. [https://doi.org/10.2983/0730-8000\(2007\)26\[479:ogaaco\]2.0.co;2](https://doi.org/10.2983/0730-8000(2007)26[479:ogaaco]2.0.co;2) [https://doi.org/10.2983/0730-8000\(2007\)26\[479:OGAACO\]2.0.CO;2](https://doi.org/10.2983/0730-8000(2007)26[479:OGAACO]2.0.CO;2)
Methods

Leung, E., Small, J., & Lotterhos, K. (in press). A comparison of non-lethal genotyping methods for juvenile Eastern oysters. *Ecology and Evolution*.
Results

Proestou, D. A., Delomas, T. A., & Small, J. M. (2025). Genetic parameters for dermo challenge survival in a Virginia, U.S. eastern oyster breeding population. *Aquaculture*, 609, 742816.
<https://doi.org/10.1016/j.aquaculture.2025.742816>
Methods

elisabethleung. (2026). *elisabethleung/non_lethal_oysters: 02/15/2026* (Version oysters) [Computer software]. Zenodo. <https://doi.org/10.5281/ZENODO.18654417> <https://doi.org/10.5281/zenodo.18654417>
Software

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

Parameters

Parameter	Description	Units
Overall_Trт	All three treatments combined into one term to easily identify all at once.	unitless
Vial_Number	The number of the vial where tissue or fluid samples were placed.	unitless

extraction_label	The id given to an extraction. This consists of the vial number, the date of the extraction, and	unitless
kit_used_shortcode	The short name of the DNA extraction kit used for the extraction.	unitless
sample_amount	The amount of sample	Either in mg, mm or uL depending on the treatment
elution_order	If the elution was the first or the second one	unitless
elution_amount	The amount in which the DNA was eluted into	uL (microliter)
pre_treatment_analysis	if there was anything done to the sample before doing the DNA extraction	unitless
lyse_treatment_hours	the number of hours the samples was lysed for	Hours
dna_conc_ng_ul	The reading from the Qubit of ng/ul of DNA in the sample	ng/ul (nanogram per microliter)
dna_conc_ng_total	The reading from the Qubit of the total ng of DNA in the sample.	ng (nanogram)
method_dna_conc	How the DNA was quantified (all Qubit in this case)	unitless
qubit_kit_name	The name of the Qubit kit, either High Sensitivity (HS) or Broad Range (BR)	unitless
qubit_standard_1	The ID of the first standard used for the Qubit kit	unitless
qubit_standard_2	The ID of the second standard used for the Qubit kit	unitless
qubit_date	The date that the sample was run on the qubit	unitless
publish	description	units
gel_date	The date that the gel was run	unitless
gel_score	The score of the gel	unitless

notes	Any notes	unitless
-------	-----------	----------

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

Instruments

Dataset-specific Instrument Name	Qiagen DNeasy Blood & Tissue kit (Qiagen; 69504)
Generic Instrument Name	DNA Analysis Kit
Generic Instrument Description	A laboratory kit containing reagents and materials used to extract, purify, detect, measure, or otherwise evaluate DNA from a sample.

Dataset-specific Instrument Name	Qubit High Sensitivity Kit (ThermoFisher; Q32851)
Generic Instrument Name	DNA Analysis Kit
Generic Instrument Description	A laboratory kit containing reagents and materials used to extract, purify, detect, measure, or otherwise evaluate DNA from a sample.

Dataset-specific Instrument Name	Qubit Broad Range Kit (ThermoFisher; Q32850)
Generic Instrument Name	DNA Analysis Kit
Generic Instrument Description	A laboratory kit containing reagents and materials used to extract, purify, detect, measure, or otherwise evaluate DNA from a sample.

Dataset-specific Instrument Name	Qubit 3.0 Fluorometer
Generic Instrument Name	Qubit fluorometer
Generic Instrument Description	Benchtop fluorometer. The Invitrogen Qubit Fluorometer accurately and quickly measures the concentration of DNA, RNA, or protein in a single sample. It can also be used to assess RNA integrity and quality. Manufactured by Invitrogen, Carlsbad, CA, USA (Invitrogen is one of several brands under the Thermo Fisher Scientific corporation.)

[[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](#)]