

RESEARCH ARTICLE

Incorporating environmental DNA metabarcoding for improved benthic biodiversity and habitat mapping

Rylan J. Command¹ , Shreya Nemani², Benjamin Misiuk^{2,3} , Mehrdad Hajibabaei^{4,5}, Nicole Fahner⁴, Emily Porter⁴, Greg Singer⁴, Beverly McClenaghan^{4,6} & Katleen Robert¹

¹School of Ocean Technology, Fisheries and Marine Institute of Memorial University of Newfoundland, St John's, Newfoundland, Canada

²Department of Geography, Memorial University of Newfoundland, St John's, Newfoundland, Canada

³Department of Earth Science, Memorial University of Newfoundland, St John's, Newfoundland, Canada

⁴Centre for Environmental Genomics Applications, St John's, Newfoundland, Canada

⁵Centre for Biodiversity Genomics & Department of Integrative Biology, University of Guelph, Guelph, Ontario, Canada

⁶David Suzuki Foundation, Vancouver, British Columbia, Canada

Keywords

Benthic habitat mapping, eDNA, joint species distribution modelling, marine biodiversity, metabarcoding, seafloor imagery

Correspondence

Rylan J. Command, School of Ocean Technology, Fisheries and Marine Institute of Memorial University of Newfoundland, St John's, NL, Canada. E-mail: rylancommand95@gmail.com

Funding information

This research was funded by Fisheries and Oceans Canada Coastal Environmental Baseline Program. Shreya Nemani received baseline funding from Memorial University's School of Graduate Studies as part of the Master of Science program. Additional funds were received by Principal Investigator K. Robert's Canada Research Chair in Ocean Mapping.

Editor: Dr. Kylie Scales

Associate Editor: Dr. David Curnick

Received: 10 May 2025; Revised: 4 October 2025; Accepted: 4 November 2025

doi: 10.1002/rse2.70048

Remote Sensing in Ecology and Conservation 2026; **12** (3):444–461

Abstract

Complex coastal seascapes harbor high marine biodiversity from which humans derive numerous ecosystem services. Maps of benthic habitats are important tools used to inform coastal development and conservation efforts. Seafloor imagery is commonly used to collect information about the distribution of benthic organisms, but these data are often limited to low taxonomic resolutions and may systematically underrepresent local biodiversity. Recent advances in genomics enable rapid and accurate detection of taxa with high taxonomic resolution from environmental DNA (eDNA) extracted from water samples, but there are few examples of broad-spectrum eDNA biodiversity data in nearshore benthic habitat mapping. We combined an eDNA-based biodiversity assessment with concurrently collected high-resolution video ground-truth data to assess the benefit of metabarcoding data for improving benthic habitat mapping in the sub-Arctic coastal embayment of Mortier Bay, Newfoundland and Labrador, Canada. Features derived from acoustic bathymetry and backscatter data were used to develop full-coverage habitat and biodiversity maps using a joint species distribution-modeling framework. The predicted taxonomic richness spatial patterns were similar between video-only, eDNA-only and combined datasets, suggesting diversity patterns were accurately represented by both methods. However, 226 additional taxa (72 species, 109 genera) were identified using eDNA compared to the 46 detected by video ground-truthing. Averaged over all taxa, the video-only model performed best in terms of discriminating presences from absences; however, we found that most sessile taxa were better predicted by the combined dataset compared to video data alone. These results highlight the limitations of imagery-only datasets for biodiversity surveys and demonstrate the utility and limitations of metabarcoding data to improve benthic habitat and diversity maps in complex coastal habitats. This study highlights opportunities to fill gaps that could improve spatial modeling of seafloor assemblages derived from metabarcoding data, including sources and sinks of DNA in the environment and water column properties that control its dispersal.

Introduction

Benthic habitat mapping characterizes the seabed based on biophysical characteristics and supports effective management of marine ecosystems (Brown et al., 2011;

Buhl-Mortensen et al., 2015). Development of habitat maps typically relies on data from remote-sensing instruments to derive spatially continuous information on depth (bathymetry) and seabed texture (backscatter), which can be used to infer species–environment

relationships (Brown et al., 2011; Lecours et al., 2015). Underwater videos, bottom trawling or sediment grabs are often used to ground-truth these acoustic data and derive biological information related to species presence and abundance (Lathrop et al., 2006; Misiuk & Brown, 2024). Species distributions at ground-truthing stations can then be modeled as a function of the remotely sensed terrain features, allowing for the generation of a full coverage map over the entire area of interest.

Biological seafloor data has historically been collected using bottom-contact fishing gear or grab sampling (e.g., Van-veen, box corer, Remotely Operated Vehicle) (Misiuk & Brown, 2024). However, the use of underwater imagery has increased in recent years (Mallet & Pelletier, 2014), and is now a standard tool for ground-truthing habitat maps (e.g., Misiuk & Brown, 2024; Nemani et al., 2022). It is particularly useful for exploring hard-bottom marine habitats where other physical sampling methods are less reliable. Imagery is a less-destructive sampling method than grabs or bottom trawling and can provide important *in situ* ecological context. However, it suffers from several limitations. First, rare, cryptic or motile species can be missed or misidentified in imagery (Ruppert et al., 2019). Second, video quality can be influenced by environmental conditions such as turbidity, sediment clouds or dense vegetation (Durden et al., 2016). Finally, imagery-based taxonomic resolution relies heavily on morphology and is therefore inherently limited (Durden et al., 2016; Howell et al., 2014). Although some of these challenges can be mitigated by careful planning, sampling design and training of image annotators, complementary data collection methods can help address the limitations of imagery alone.

One such method is the use of environmental genomics, a rapidly developing tool for biodiversity characterization that involves collecting and analyzing environmental DNA (eDNA) shed by organisms into their environment (Baird & Hajibabaei, 2012; Deiner et al., 2017). DNA can be extracted from sediment or water samples, amplified using polymerase chain reaction (PCR) and sequenced using high-throughput DNA sequencing technology. DNA sequences are then taxonomically assigned based on reference databases containing genetic information from known specimens. This process, known as metabarcoding, can be tailored to target specific taxonomic groups or species of interest (Compson et al., 2020), or can be used to characterize entire ecological communities at a high taxonomic level (e.g., eukaryotes). By identifying taxa that cannot be captured using other methods and by differentiating closely related or cryptic species, eDNA can yield higher resolution taxonomic data compared to traditional sampling

methods (Bonfil et al., 2021; Cahill et al., 2018; Gibson et al., 2015). eDNA has been used for individual species distribution modeling (Eddings, 2019; Kutti et al., 2020; Riaz et al., 2020); however, incorporation of eDNA metabarcoding into benthic species assemblage predictive mapping frameworks remains rare (but see Jaquier et al., 2021; Pichler & Hartig, 2021).

Coupling video and eDNA collection and analysis simultaneously may compensate for the above limitations and improve taxon detection, and many studies have compared eDNA to morphological and image-based datasets (Cahill et al., 2018; Muenzel et al., 2024; Stat et al., 2019; Valdivia-Carrillo et al., 2021). Indeed, eDNA metabarcoding can provide complementary information to morphological identifications (Cahill et al., 2018) and contribute to a holistic understanding of species distribution within marine benthic communities. However, to the best of our knowledge, the combination of video and eDNA datasets has not yet been applied to generate habitat maps for nearshore coastal benthic assemblages. Our goal was to use co-located eDNA sampling to improve a benthic habitat map created using only video ground-truth data (Nemani et al., 2022). Specifically, we were interested in three questions: (1) Does eDNA detect a different set of taxa than video ground-truthing? (2) Do these assemblages have any relationships with remotely sensed environmental data or their derivatives? (3) Can combining eDNA with video ground-truth data yield better predictions of assemblage distribution than models based on video or eDNA alone? Our ultimate goal was to produce a full-coverage benthic habitat map for the epibenthic megafauna of Mortier Bay, Newfoundland and Labrador that incorporated the most complete biodiversity data available for the region.

Materials and Methods

Study site

Placentia Bay, Newfoundland, is an ecologically and biologically significant area that is subject to increasing human activity, mainly from shipping and aquaculture (Fig. 1A). Previous surveys were conducted by the Canadian Hydrographic Service (CHS) and the Geological Survey of Canada (GSC) between 1989 and 2006 to characterize the seafloor throughout Placentia Bay. These surveys resulted in full-coverage backscatter and bathymetry data and a set of 1:50,000 scale thematic maps (e.g., Potter & Shaw, 2009a, 2009b), and a seascape map classifying substrate type and depth throughout Placentia Bay, including several coastal embayments (Shaw et al., 2011, 2013). Mortier Bay is one such embayment located on the western edge of Placentia Bay (Fig. 1B,C). Nemani

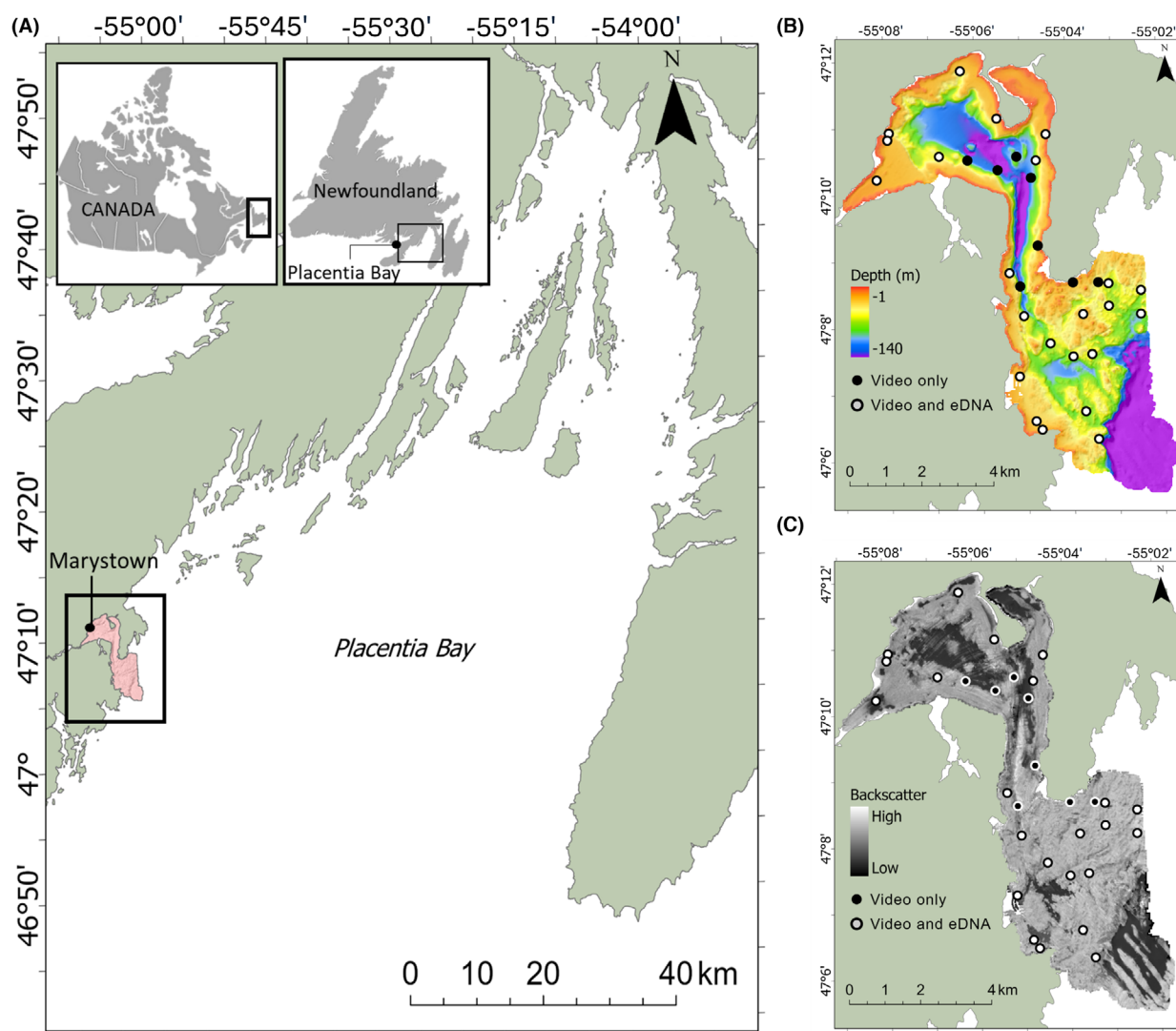


Figure 1. Map of study location and sampling sites (A), bathymetry (B), and backscatter (C), maps show the location of video ground truthing and the sites at which environmental DNA samples were collected.

et al. (2022) created the first full-coverage habitat maps of species assemblages for this area. The taxonomic and functional diversity in Mortier Bay has also been mapped (Nemani et al., 2024).

Seabed habitat data collection

We used two sources of remotely sensed acoustic data to describe the physical characteristics of the Mortier Bay seabed (Fig. 1B,C). First, we used backscatter data collected by the GSC in 1989 aboard the *RV Navicula* using a Klein-5500 side-scan sonar (SSS). Additionally, we used bathymetry data collected by the GSC and CHS in 1997 and 2004 aboard the hydrographic launch *Plover* using a hull-mounted Kongsberg-Simrad EM-3000 multibeam

echosounder (MBES) (Potter & Shaw, 2009a, 2009b). MBES bathymetry was gridded at 10 m, and the SSS backscatter was resampled from 5 m to 10 m to match the MBES data (Nemani et al., 2022). Secondary seafloor terrain features were derived from these primary data layers to further describe terrain variability and texture (see Nemani et al., 2022). In order to identify both fine- and broad-scale variability in seabed terrain and composition, we calculated derivatives across multiple scales using moving-window sizes from 3×3 to 21×21 cells (≈ 10 – 210 m), which resulted in 120 features in total (Table 1). Secondary features were derived using the Benthic Terrain Modeler (BTM) toolbox (Walbridge et al., 2018), the Terrain Attribute Selection for Spatial Ecology (TASSE) toolbox (Lecours, 2017) and the ArcGIS

Table 1. Environmental features calculated across multiple scales used to model species assemblages.

Source layer	Derivative layer	Description	Units	Scales (m)	Method
<i>Bathymetry</i>	Source bathymetry surface		meters	10	
	Bathymetric position index (BPI) – fine*	Topographic highs and lows	meters	30, 50, 70, 90, 110, 130, 150, 170, 190, 210	BTM
	Bathymetric position index (BPI) – broad*	Topographic highs and low	meters	30, 50, 70, 90, 110, 130, 150, 170, 190, 210	BTM
	Mean curvature		meters ⁻¹	30, 50, 70, 90, 110, 130, 150, 170, 190, 210	Curvature tool – ArcGIS
	Profile curvature	Rate of change of a slope (convex and concave contours in a surface)	meters ⁻¹	30, 50, 70, 90, 110, 130, 150, 170, 190, 210	Curvature tool – ArcGIS
	Planar curvature	Perpendicular to slope (defines ridges and valleys)	meters ⁻¹	30, 50, 70, 90, 110, 130, 150, 170, 190, 210	Curvature tool – ArcGIS
	Eastness	Orientation of the seabed (sine transformed measure of aspect)	unitless	30, 50, 70, 90, 110, 130, 150, 170, 190, 210	TASSE
	Northness	Orientation of the seabed (cosine transformed measure of aspect)	unitless	30, 50, 70, 90, 110, 130, 150, 170, 190, 210	TASSE
	RDMV – relative difference to the mean value	Measure of the variability of terrain ruggedness and rugosity	unitless	30, 50, 70, 90, 110, 130, 150, 170, 190, 210	TASSE
	Slope	Steepness of gradient	degrees	30, 50, 70, 90, 110, 130, 150, 170, 190, 210	TASSE
	SD	Standard deviation of terrain	meters	30, 50, 70, 90, 110, 130, 150, 170, 190, 210	TASSE
<i>Backscatter</i>	Source backscatter surface		dB	10	
	Backscatter homogeneity	Textural indices derived from GLCM (gray-level co-occurrence matrices). Homogeneity quantifies the amount of local similarity	unitless	30, 50, 70, 90, 110, 130, 150, 170, 190, 210	R – GLCM package
	Backscatter contrast	Textural indices derived from GLCM (gray-level co-occurrence matrices). Contrast quantifies the amount of local variation	unitless	30, 50, 70, 90, 110, 130, 150, 170, 190, 210	R – GLCM package
	Backscatter variability	Textural indices derived from GLCM (gray-level co-occurrence matrices). Variability quantifies the dispersion around the mean	unitless	30, 50, 70, 90, 110, 130, 150, 170, 190, 210	R – GLCM package
<i>Distance from Coast</i>			meters		Euclidian distance

Note: Scales represent the width of the window. Reproduced and modified with permission from Nemani et al. (2022).

*BPI fine was calculated using an inner and outer radius of 1 and 10 cells (a scale factor of 100 m); BPI broad was calculated using an inner and outer radius of 10 and 30 cells (a scale factor of 300 m) (Lundblad et al., 2006). Averaging over the increasing window sizes results in scales of 300, 500, 700, 900, 1100, 1500, 1700, 1900 and 2100 m for BPI fine, and 300, 900, 1500, 2100, 2700, 3300, 3900, 4500, 5100 and 6300 m for BPI broad.

Pro[®] curvature tool (with default parameters) in ArcGIS Pro[®] v.2.7, as well as the glcm package (Version 1.6.5; Zvoleff, 2020) in R Statistical Software (Version 4.1.0; R Core Team, 2021). All these tools allow the analysis window size to be specified (e.g., TASSE's 'Alternative Analysis Window' option; glcm package documentation: <https://cran.r-project.org/web/packages/glcm/glcm.pdf>).

Video data collection and processing

The video collection and analysis protocols are outlined fully in Nemani et al. (2022). Briefly,

two-minute video drifts were conducted at 31 sites in Mortier Bay using a custom drop camera system equipped with a Sony FDR-X3000 camera in a *Group B Inc.* housing, and two green-light lasers spaced 10 cm apart for scale. Videos were annotated using the Monterey Bay Aquarium Research Institute Video Annotation and Reference System (VARS) version 8.3.4 (Schlining & Jacobsen Stout, 2006). All visible taxa (> 2 cm) were identified to the lowest possible taxonomic level; taxa that could not be identified to species were grouped based on observed morphological similarity (morphotaxa).

eDNA data collection and processing

Sample collection

We collected water samples between November 18 and 20, 2019, at 25 of the 31 sites in Mortier Bay. We deployed a Niskin bottle open at both ends, lowered it to the seabed and then raised it slightly (using the vessel sonar for guidance) before triggering it to close and collect the water. Thus, water samples represent near-bottom water, typically within 1 m of the seafloor. The Niskin bottle was decontaminated between sites using ELIMINase (Decon Labs, Inc.). At each site, we conducted a single Niskin drop from which we collected three 250 mL subsamples; we also collected six field blanks (i.e., distilled water), for a total of 81 water samples. Samples were frozen immediately after collection, transported to the Centre for Environmental Genomics Applications (CEGA) and stored at -80°C until analysis.

Sample processing and analysis

We thawed water samples at 4°C and filtered each set of triplicate samples onto a single $0.22\text{ }\mu\text{m}$ PVDF Sterivex filter (MilliporeSigma) using a peristaltic pump to create 25 composite samples (one per site). We thawed and filtered field blanks individually. Prior to each filtration session, surfaces and equipment were decontaminated with ELIMINase and rinsed with deionized water. We then extracted DNA from all filter membranes using the DNeasy PowerWater Kit (Qiagen) and three DNA markers covering two genetic regions (cytochrome c oxidase I and 18S) were amplified in triplicate from each DNA extract using PCR (Table S1 Primers). This set of markers assesses a broad range of eukaryote diversity with a focus on metazoans. Each reaction contained 1X reaction buffer, 2 mM MgCl_2 , 0.2 mM dNTPs, $0.2\text{ }\mu\text{M}$ of each of the forward and reverse Illumina-tailed primers, 1.5 U Platinum Taq (Invitrogen) and $1.2\text{ }\mu\text{L}$ of DNA in a total volume of $15\text{ }\mu\text{L}$. See Table S1 Primers for PCR conditions for all primer sets. PCR replicates were pooled, cleaned with the QIAquick 96 PCR purification kit (Qiagen) and indexed with unique dual Nextera indexes (IDT; 8-bp index codes). Indexing PCR initiated at 95°C for 3 min, followed by 12 cycles of 95°C for 30 s, 55°C for 30 s, 72°C for 30 s and a final extension at 72°C for 5 min. The indexing reaction contained 1X reaction buffer, 2 mM MgCl_2 , 0.2 mM dNTPs, 2.5 mM of each of the forward and reverse Illumina Nextera indexing primers, 1.5 U Platinum Taq (Invitrogen) and $2\text{ }\mu\text{L}$ of DNA in a total volume of $25\text{ }\mu\text{L}$. Amplicons were quantified with Quant-iT PicoGreen dsDNA assay and pooled in equimolar concentration by DNA marker. Amplicon pools were cleaned using

AMPure XP cleanups, diluted to 4 nM and pooled into a single sequencing library with 10% PhiX. Negative controls generated during PCR, extraction, filtration and field collection were carried through to sequencing to monitor for contamination. The resulting library was sequenced on one MiSeq v3 600 cycle (Illumina) flow cell at a target depth of 100,000 sequences per sample per DNA marker. Raw sequence reads are available in NCBI's sequence read archive under project PRJNA1206193.

Base calling and demultiplexing were performed using Illumina's bcl2fastq software (v2.20.0.422). Primer sequences were trimmed using cutadapt v1.18. Sequences were filtered for quality and length ($\text{maxEE} = 2$, $\text{minQ} = 2$, $\text{truncQ} = 2$, $\text{maxN} = 0$) and then denoised using default parameters in DADA2 v1.16 to create exact sequence variants (ESVs), each of which represents a unique sequence from the sample (Callahan et al. 2016). ESVs with two or more reads were assigned taxonomy using the National Center for Biotechnology Information (NCBI) blastn tool (v2.11) and the nucleotide database (downloaded November 2020) with an e-value cut-off of 0.001. The taxonomy presented here matches the naming conventions used in NCBI's taxonomy database and was assigned based on a sequence similarity score (the product of the % sequence similarity and the % query coverage) using minimum score thresholds as follows: species (99%), genus (98%), family (95%), class (90%) and phylum (85%). The results from the three DNA markers were consolidated to create taxonomic lists, and eDNA occurrences are hereafter treated as presence-absence data. Taxonomic identifications of metazoans were verified against the World Register of Marine Species (WoRMs), the Global Biodiversity Information Facility and the Encyclopedia of Life (EOL) to ensure that spurious matches resulting from poor reference database coverage were not included in the list. Most ESVs shared between negative controls and samples did not pass the previously described filters (score thresholds) for inclusion in the analysis, which is focused only on the taxonomic data. The remaining ESVs identified in negative controls, assigned to humans, land plants, common insect pests and a freshwater rotifer were removed from the sample data. All non-target taxa (e.g., fungi and bacteria) were removed from taxonomic lists.

Dataset partitioning

We used three datasets for modeling. The 'video' dataset consisted of taxa identified via video ground-truthing. The 'eDNA' dataset consisted of all taxa where the adult stage is $> 2\text{ cm}$ and exhibits a benthic (or benthopelagic, in the case of Ammodytidae) lifestyle. This excluded many copepods, jellies and other zooplankton, as well as

Platyhelminthes and other meio- and macroinvertebrates that are too small to see in underwater video. We also merged the video and eDNA observations to create a 'combined' dataset (Table S2). A taxon was considered to be present at a given site if it was detected by at least one of the sampling methods (Table S3). Because taxonomic resolution differed between sampling methods, we pooled taxon identifications at the lowest common taxonomic rank shared between datasets (Table S3). In each dataset, rare taxa that occurred in fewer than three sample sites were excluded from the analysis (Ovaskainen & Abrego, 2020).

Data analysis

Variable selection

Prior to modeling, we selected predictor variables from the acoustic data using Redundancy Analysis (RDA) implemented via the 'vegan' package (Version 2.6–2; Oksanen et al., 2022). Given a matrix of predictor variables X , and a species matrix, Y , RDA finds the linear combination of X that best explain the variation in Y . This is a common ecological technique for reducing predictors while maintaining those important to explain variation in Y . Here, we used a scale-based variables selection approach, which was identified as the best performing variable selection method for the video ground truth data (Nemani et al., 2022). Because there were significantly more variables than observations, variables were split into two categories—bathymetry derivatives and backscatter derivatives. Within each category, variables were grouped by scale, according to the size of the focal window used to perform the calculation (e.g., 3×3 bathymetry features, 5×5 bathymetry features, etc.). We fit each species matrix as a function of feature categories for a given scale and determined the best set of variables for each scale-feature subset via forward selection, using the R^2_{adj} value of the full RDA as stopping criteria. We then ran a second RDA, fitting the species matrix as a function of the variables selected in the first round. Finally, we calculated Spearman's rank correlation between variables selected by the RDA process to ensure any remaining correlations were below 0.7 (Fig. S5).

This method was applied to each dataset (i.e., video, eDNA, combined) separately to determine the features that best explain variation in the taxa present.

Joint-species distribution modeling with HMSC

We fit joint-species distribution models using the Hierarchical Modeling of Species Communities (HMSC) implemented in the 'Hmsc' package (Version 3.0–12; Tikhonov

et al., 2022) in R. HMSC is a Bayesian generalized linear model framework that allows the distribution of species assemblages to be jointly modeled as a function of environmental covariates, functional traits, phylogenetic relationships and structured (e.g., spatial, temporal) and unstructured random effects (Ovaskainen et al., 2017). Unlike many other habitat mapping methods, HMSC assumes a continuous distribution of species along environmental gradients. The probability of species occurrences can be predicted over the entire study area and then classified using clustering methods to identify 'Regions of Common Profile' (RCP). As such, this approach can be considered a 'predict-and-assemble together' analogue of traditional habitat mapping (Foster et al., 2013).

For each dataset (video, eDNA, combined), we fit three models: one containing environmental covariates (variables selected by RDA) only ('environmental'), another containing a site-level spatial random effect only ('spatial') and a final model containing both environmental covariates and the spatial random effect ('full'). This approach allowed us to quantify the contribution of the spatial effect and environmental covariates separately and to investigate residual co-occurrences.

We explored the explanatory and predictive powers of the models through species-specific AUC (Pearce & Ferrier, 2000) and Tjur's R^2 (hereafter Tjur R^2 ; Tjur, 2009) values. Although AUC and Tjur R^2 are both metrics of discriminating presence from absence, they measure complementary aspects of model performance (Abrego & Ovaskainen, 2023). In the context of species distribution modeling, AUC is defined as the proportion of cases where a randomly selected occupied sampling unit has a higher occurrence probability than a randomly chosen unoccupied sampling unit (Hanley & McNeil, 1982). Tjur R^2 is defined as the difference in the average predicted probability between occupied and unoccupied sampling units. In other words, AUC measures a model's ability to discriminate whether a species is present or absent in all sampling units, whereas Tjur R^2 measures the model's ability to predict correctly only when a species is present with a high probability. A model with high AUC does well at predicting where a species is present and where it is absent. The same model with a low Tjur R^2 would fail to identify the sampling units where the species is present with high probability but would be very useful for identifying areas where the species is very unlikely to occur.

We performed a Leave-One-Out Cross-Validation to estimate predictive power from which we derived confusion matrices and classification metrics based on the Receiver-Operator Characteristic (ROC) curve (i.e., sensitivity and specificity). To evaluate the taxonomic richness predictions, we used simple regression metrics (i.e.,

Pearson's coefficient of correlation [r], mean absolute error (MAE) and per cent variation explained (%VE)) to evaluate how well each model explained and predicted the number of species occurring in Mortier Bay. More details on model specifications, diagnostics, variation partitioning and the role of environmental variables influencing species distributions (Ovaskainen & Abrego, 2020) can be found in Appendix S1.

Benthic habitat and diversity maps

We produced separate predictions from all three models ('environment', 'spatial', 'full') over the entire study area by pooling the parameter estimates across all four Markov Chain Monte Carlo (MCMC) chains and making predictions for each of the 1000 posterior samples (Appendix S1). We averaged the predictions over all 1000 samples to get the posterior mean predicted occurrence probability for each taxon and took the standard deviation across the posterior samples to obtain a depiction of the uncertainty of the predictions (Ovaskainen & Abrego, 2020). HMSC can also make predictions of community-level metrics, so we performed the same averaging procedure to derive posterior mean predictions of taxonomic richness and its standard deviation for the entire study area to generate a full-coverage map of local biodiversity for Mortier Bay.

Both the spatial extent and location of underwater video samples are known based on the camera field of view and the length of the video drift. However, eDNA is ostensibly transported through the water column before being sampled, making its provenance and transport distance unknown. To estimate the extent of eDNA transport, we calculated Jaccard dissimilarity from eDNA community data (subset used for modeling) between sample sites and plotted it as a function of distance. This yielded an empirical estimate of the rate of community turnover at the scale of the study area that is attributable to spatial autocorrelation—similar to the univariate empirical variogram. We repeated this analysis for video-derived community data for comparison.

To investigate taxonomic co-occurrences, we grouped predictions into regions of similar taxonomic profile using self-organizing maps (SOMs) with a 5×5 window (after Tikhonov et al., 2020) using the *kohonen* package in R (Wehrens & Buydens, 2007; Wehrens & Kruisselbrink, 2018). SOMs are unsupervised neural networks that reduce high-dimensional data into two dimensions by maximizing global and local distances between groups (Kohonen, 1982). Here, the algorithm grouped sites together that have similar taxonomic composition and are close together in real space. We then mapped the nodes into the U and V components of the YUV color space.

Depiction of taxonomic composition as a continuous gradient avoids imposing assumptions about the structure of taxonomic assemblages on the predictions.

Results

eDNA metabarcoding

A total of 12,957,844 sequences were generated from the 25 composite samples analyzed for this project, yielding an average of 106,259 ($\pm 60,084$ SD) sequences per sample per marker after bioinformatic filtering. We grouped these sequences into 13,700 ESVs across the three DNA markers with a median of 540 ESVs per sample per marker. Overall, 99% (13,498) of ESVs matched a reference sequence, 47% (6,417) were attributed to eukaryotes and 25% (3,421) met the selection criteria for taxonomy assignment for high-quality eukaryote matches.

We identified 18 metazoan phyla across the 25 sites (Table S2). Phyla Annelida, Arthropoda, Bryozoa, Chordata and Mollusca were detected at all 25 sampling locations, whereas Cnidaria, Ctenophora, Echinodermata and Porifera were detected in $> 80\%$ of sampling locations. Using the number of ESVs as a proxy for diversity within each phylum, Annelida ($n_{ESV} = 124$) had the greatest diversity across all samples, followed by Arthropoda ($n_{ESV} = 66$), Mollusca ($n_{ESV} = 61$), Cnidaria ($n_{ESV} = 53$), n_{ESV} and Porifera ($n_{ESV} = 43$) which were also highly diverse.

Within the 18 metazoan phyla, we identified 35 classes, 67 orders, 124 families, 119 genera and 77 species. The families Mopaliidae (Mollusca) and Terebellidae (Annelida) were detected at all 25 sites, and Oithonidae (Arthropoda), Temoridae (Arthropoda), Ammodytidae (Chordata), Obeliidae (Cnidaria), Pleurobrachiidae (Ctenophora) and Spirastrellidae (Porifera) were found at 80% or more of the sites (Fig. S1). In addition to metazoans, we identified 18 phyla and 50 classes of non-animal eukaryotes (i.e., algae and protists).

Considerably more taxa were identified via eDNA ($n_s = 237$) compared to video ($n_s = 46$). Species accumulation curves indicate that the number of taxa identifiable using video was nearly saturated (Fig. 2). However, saturation was not reached for the eDNA method, indicating missing taxa that were not sampled in Mortier Bay (Fig. 2).

Of the 237 taxa identified by eDNA, we removed 28 due to their pelagic life history, and 86 due to their size (adult stage < 2 cm); 7 taxa had both characteristics. Finally, 71 taxa occurred at fewer than three sites and were removed due to low occurrence. Thus, we included 57 taxa in the eDNA model, consisting of 9 phyla, 15 classes, 29 orders, 37 families, 34 genera and 20 species.

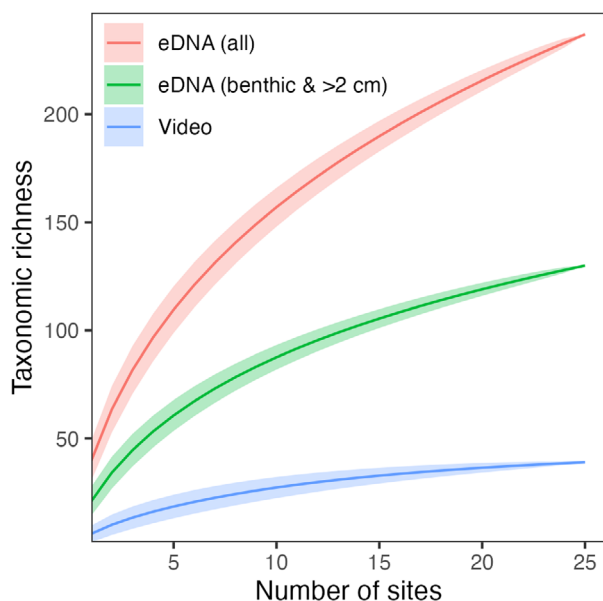


Figure 2. Species accumulation curves for the video ground-truth sampling (blue) and eDNA sampling methods (red = all detected fauna, green = detected benthic fauna >2 cm) at the 25 sites in Mortier Bay where both methods were conducted. Ribbons indicate the 95% confidence interval around the expected taxonomic richness.

HMSC model fit

The ‘combined’ model fit the corresponding data well for explanatory power (Tjur $R^2 = 0.20$, $AUC = 0.88$), but exhibited poor predictive power (Tjur $R^2 = 0.027$, $AUC = 0.54$). Overall, the video-only model performed better at discriminating presence from absence in the corresponding dataset, compared to the eDNA and combined models (Table 2). However, for some individual taxa (e.g., *Cyclostomatida* sp.1, *Ascidacea* sp.1, *Halisarca dujardini*), the combined model was best (Fig. 3). The video model performed very poorly when predicting taxonomic richness, whereas the eDNA and combined models performed moderately (Table 2).

Predictive power for individual taxa

Combined versus video-only

At least one of the combined or video-only models was able to discriminate presence from absence for all taxa detected via video ground-truthing ($AUC_{CV} > 0.5$, Tjur $R^2_{CV} > 0$), except for *Arctica islandica* and *Cancer irroratus*, which were poorly predicted by both models (Table S4). The sponge *Porifera* sp.1 and mussel *Mytilidae* were predicted similarly well by both models (Table S4). The sand-dollar *Echinarachnius parma* was also predicted equally well by both models, with higher Tjur R^2_{CV} values for the video-only model but higher AUC_{CV} values for the combined model (Table S4).

The combined model made better predictions (i.e., yielded higher AUC_{CV}) than the video-only model for 5 of 18 taxa, all five of which were sessile filter-feeders: the tunicates *Ascidacea* sp.1 and *Tunicata* sp.1, the bryozoans *Bryozoa* sp.1 and *Cyclostomatida* sp.1, and the bivalve *Arctica islandica* (Table S5). However, a few sessile taxa (e.g., *Hormathia*, *Mytilidae*, *Metridium senile*) were better predicted by the video-only model.

Overall, the video-only model better predicted all mobile taxa, including all sea stars, the sea urchin *Strongylocentrotus droebachiensis*, the rock crab *Cancer irroratus* and the hermit crabs *Pagurus* sp. (Table S4). The video-only model predicted brittle stars (*Ophiuroidea*) with higher sensitivity and the sea star *Pteraster militaris* with higher specificity than the combined model. For other taxa, sensitivity and specificity were generally higher for the combined model compared to the video model, although no clear patterns in these ROC metrics were evident to help explain the differences in predictive power for other mobile taxa.

Combined versus eDNA-only

After grouping taxa for the combined dataset, there were 51 taxa common to the eDNA-only and combined models. At least one of the combined or eDNA-only

Table 2. Model performance metrics for explanatory and cross-validated predictive (subscript ‘CV’) power of the video, eDNA and combined datasets.

	Presence–Absence				Richness		
	AUC	AUC_{CV}	Tjur R^2	Tjur R^2_{CV}	Pearson’s r_{CV}	MAE_{CV}	$\%VE_{CV}$
Video	0.94	0.67	0.37	0.13	−0.26	1.96	6.60
eDNA	0.83	0.51	0.15	0.02	0.44	3.97	19.70
Combined	0.88	0.54	0.20	0.03	0.46	4.37	21.00

Note: Pearson’s r is the correlation between the observed and cross-validated predictions for taxonomic richness. The mean absolute error (MAE) is the average absolute difference between the observed and predicted taxonomic richness for each model. $\%VE_{CV}$ is the percentage of variance in taxon occurrences explained by each model based on cross-validation.

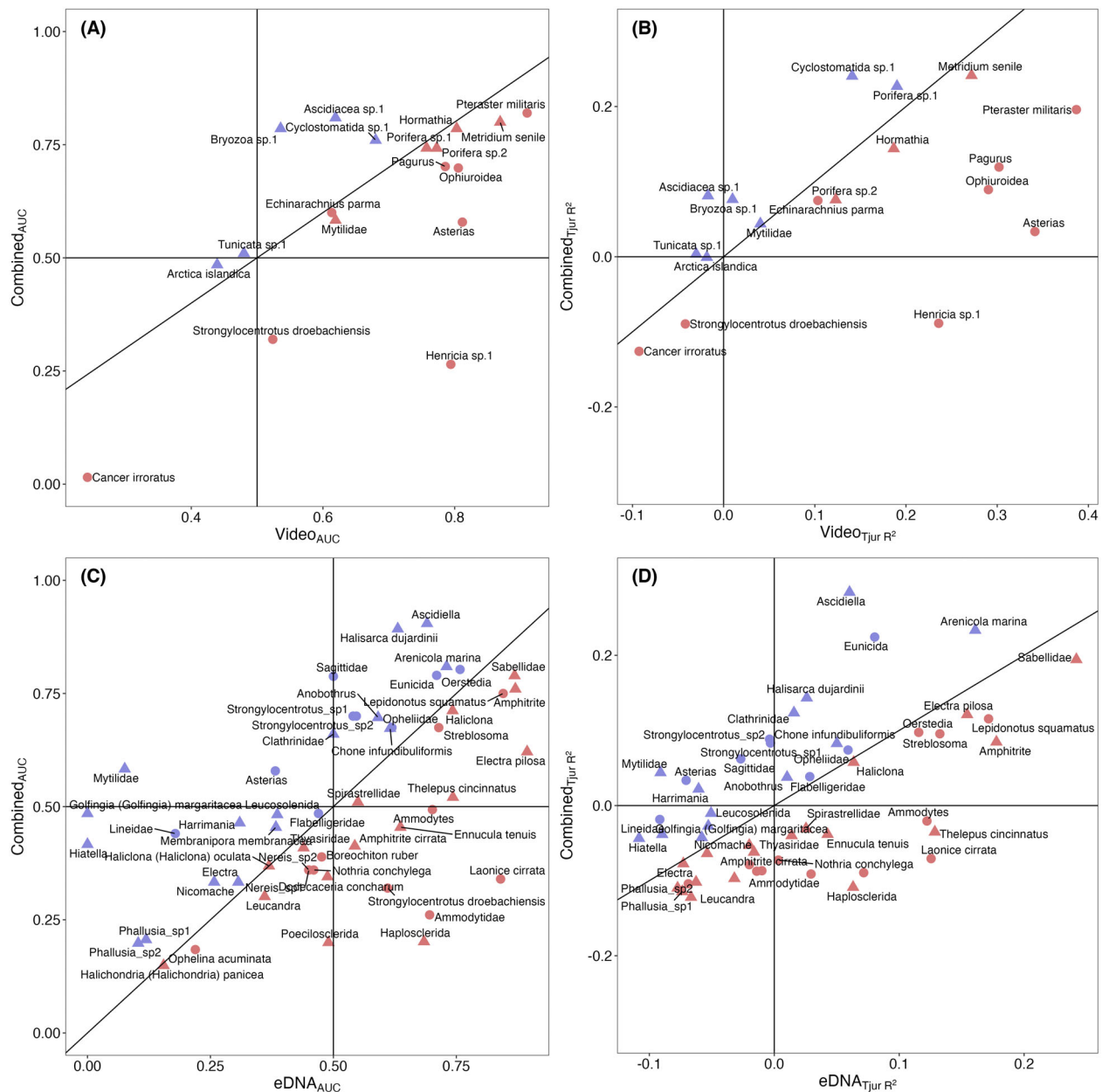


Figure 3. Cross-validated AUC (A, C) and Tjur R^2 (B, D) values for taxa present in both video (A, B) and eDNA (C, D) datasets, modeled using the combined dataset. For each panel, the top-left quadrat indicates taxa for which predictions were poor for the single method but better than average for the combined model. The bottom-right quadrant indicates taxa for which predictions were poor for the combined model but better than average for the single method model. The top-right quadrant indicates taxa that were predicted well by both combined and eDNA-only models. The combined model better predicted taxa above the 1:1 diagonal line, while the eDNA-only model better predicted taxa below the 1:1 diagonal line. Taxa with predominantly mobile (sessile) adult stages are indicated with triangles (circles).

models was able to discriminate presence from absence for 29 of the 51 taxa detected by eDNA ($AUC_{CV} > 0.5$, $Tjur R^2_{CV} > 0$); the other 22 of 51 taxa were poorly predicted by both models, exhibiting no clear taxonomic pattern (Table S4).

The combined and eDNA-only models performed similarly with respect to the number of taxa they best predict. The eDNA-only model predicted 26 of 51 taxa better than the combined model, whereas the combined model predicted 25 of 51 taxa better than the eDNA-only model.

The majority of taxa best predicted by the combined model were sessile, including the ascidian *Ascidella*, and for some sponges including *Halisarca dujardini*, Clathrinidae, Leucosolenida and *Haliclona* (*Haliclona*) *oculata*, although the latter was poorly predicted overall ($AUC_{CV} < 0.5$, $Tjur R^2_{CV} < 0$). Several polychaete taxa were also better predicted by the combined model, including *Arenicola marina*, *Eunicida*, *Opheliidae* and *Chone infundibuliformis*.

Interestingly, the combined model made better predictions than the eDNA-only model for the sea star *Asterias*, although it did not perform as well as the video-only model. The sea urchins *Strongylocentrotus* sp.1 and sp.2 followed the same trend.

Predictive power for taxonomic richness

In terms of community-level predictions, the combined model performed slightly better than the eDNA-only model, which in turn performed better than the video model; each model explained 21.0%, 19.7% and 6.6% of the variation in taxonomic richness, respectively. The spatial predictions made for each model were similar (Fig. 4A–C). However, the combined and eDNA models predicted higher taxonomic richness than the video model in each area; this result was expected based on the higher number of species identified by eDNA. Additionally, the standard deviation around the mean posterior predicted taxonomic richness was higher for the

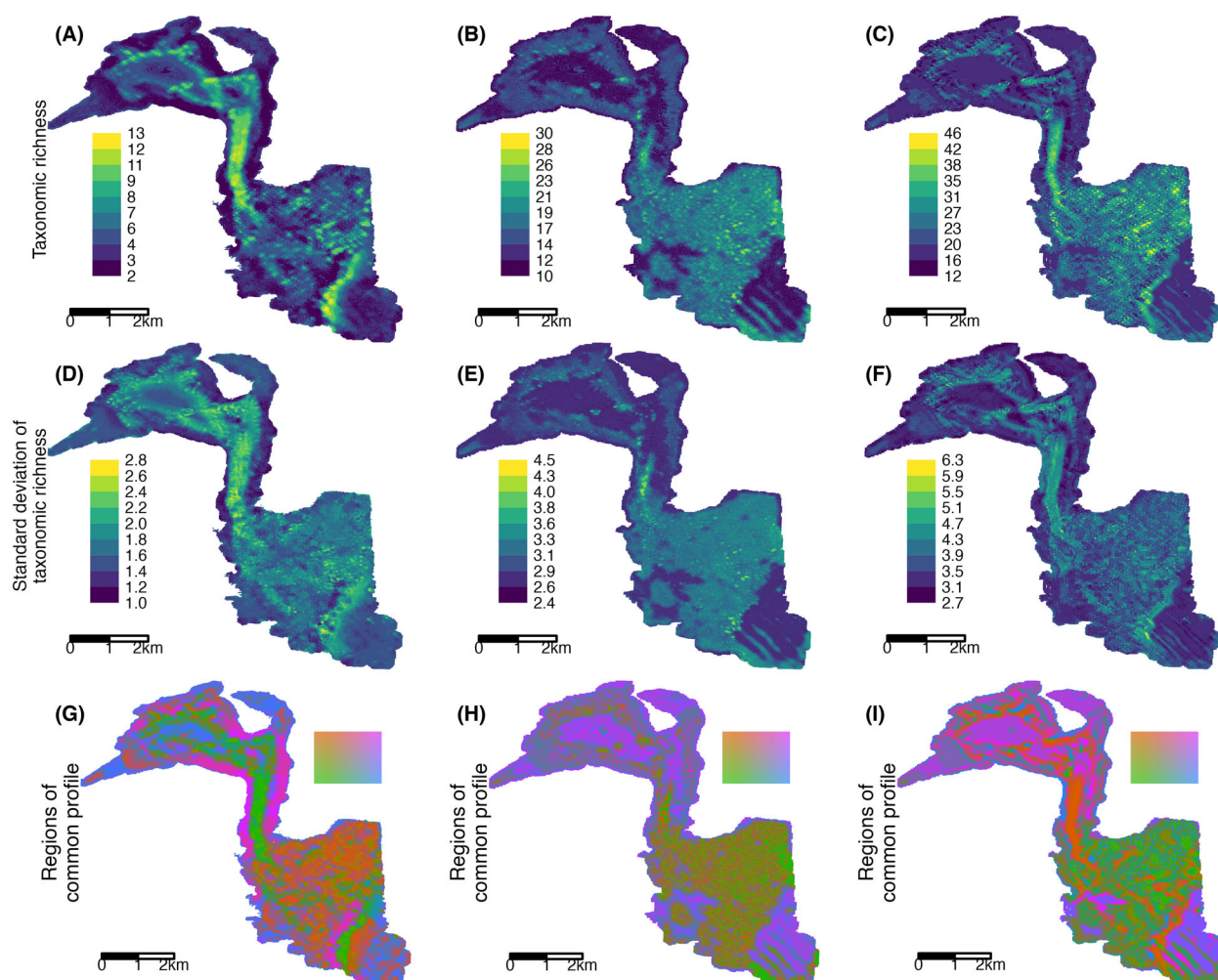


Figure 4. Spatial predictions for the video-only model (A, D, G), the eDNA-only model (B, E, H) and the combined data model (C, F, I). The top row (A–C) represents mean posterior predicted taxonomic richness. The middle row (D–F) depicts the standard deviation around the mean posterior predicted taxonomic richness. The bottom row (G–I) represents predicted regions of common profile, with nodes of the 5×5 self-organizing map mapped to YUV color space. In G, H and I, the colors correspond to different regions of common profile because different species are included in each dataset.

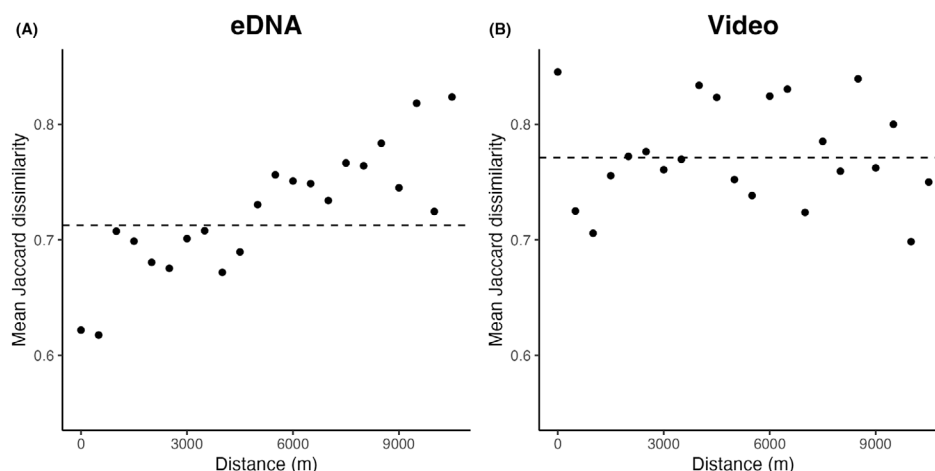


Figure 5. Mean Jaccard dissimilarity in community composition as a function of distance, aggregated to 500 m lags, between (A) eDNA samples and (B) video samples.

combined and eDNA models compared to the video model (Fig. 4D–F).

Spatial scale and community dissimilarity

Jaccard community dissimilarity was relatively constant at all inter-site distances (~ 200 m–1 km) when calculated from video data but was low at proximal sites (< 1 km) and increased consistently with distance between sites when calculated from eDNA data (Fig. 5). Dissimilarity measured from eDNA data did not appear to plateau at the full extent of the study area (~ 11 km), suggesting that even the most distal sites may not be spatially independent. Conversely, the lack of spatial structure in community dissimilarity observed from video data suggests fine-scale spatial autocorrelation, which occurs at extents less than the minimum distance between sample sites (~ 200 m). Jaccard dissimilarity from eDNA samples at the most distal sites (> 9 km) was comparable to dissimilarity calculated from the most proximal video samples.

Ecological inference: Regions of common profile

In order to contextualize our ecological inference, we reference the seascapes identified by Shaw et al. (2011). The regions of common profile (RCP) identified by each model were generally similar in geographic location and extent. The shallow muddy seascapes in the north of the bay and the deep muddy seascapes in the south were identified as similar RCPs by each model (Fig. 4G–I). Similarly, the taxonomic composition of the high relief,

hard, complex substrates in the north of the bay was distinguished from the muddy seascapes. The taxonomic composition of the northern seascapes was also distinct from the deeper southern bedrock seascapes, which were patchy and variable.

The models differed in the taxa associated with each RCP. The eDNA-only model closely followed the extent of the Shaw et al. (2011) seascapes, with hard boundaries evident between muddy and hard-bottom seascapes. However, the eDNA model did not generalize well; several regions exhibit colors from opposite axes of the YUV color space (Fig. 4H), indicating that the model cannot discriminate between assemblages. In contrast, the RCPs identified from the combined model, and to a lesser extent, the video model, were more spatially continuous. This was especially evident in homogeneous muddy seascapes, although this pattern was also observed in bedrock and high relief seascapes (Fig. 4I).

Discussion

Spatial patterns captured by video were a decent proxy of the benthic community, and the distribution of many taxa was best explained by video alone. However, video ground-truthing provided an incomplete benthic species catalogue for Mortier Bay. Indeed, near-bottom water samples analyzed using eDNA metabarcoding techniques revealed an additional 237 taxa. Among the detections were a large number of sponges and infaunal taxa (i.e., polychaetes), which are difficult or impossible to identify using video alone. Failing to observe and account for ecologically important species could result in inaccurate habitat maps which, when used for decision-making, could

ultimately lead to erroneous conclusions and suboptimal management outcomes.

Additionally, cryptic and rare species often missed in visual surveys may be detected by eDNA, improving overall coverage of the benthic community (Ruppert et al., 2019). However, the detection of rare species by metabarcoding methods may be limited by available reference sequences. It is also important to note that rarity could be due to sampling bias associated with gear type (e.g., cameras, bottom trawls) rather than a truly uncommon species. Although taxa present in fewer than three sites were removed from our spatial analysis, considerably more taxa were detected with eDNA metabarcoding than video in our study. Indeed, our species accumulation curve for eDNA detections suggests that more sampling is needed to complete the species catalogue for Mortier Bay, potentially driven in part by the challenge of adequately sampling complex coastal ecosystems.

Ecological inference and method comparison

The ability of each model to differentiate presence and absence is evaluated based on the dataset used to generate that model. As such, a model that performs better at this task may not necessarily provide a better reflection of true ecological patterns. For example, if video data does not capture a particular species, we cannot assess the performance of the video model for that species. The same can be said for different life stages. However, patterns in community structure were generally similar among eDNA and video methods and were consistent with the modeling approach of Nemani et al. (2022).

For some taxa, combining metabarcoding results with video ground-truthing improved spatial predictions. This was especially true for sessile, substrate-associated, filter-feeding taxa (e.g., sponges, ascidians, bryozoans), which are difficult to identify in video, particularly in complex habitats. Conversely, mobile taxa were better predicted by the video-only model compared to the combined model, possibly due to DNA dilution since mobile species move between habitats, reducing the probability of detection. Video likely captures these species in the habitats in which they spend most of their time, whereas eDNA could theoretically capture them in any habitat they pass through. Furthermore, organism morphology may also have contributed to this result. Most of the epibenthic mobile invertebrates observed in our video footage were decapod crustaceans with hard exoskeletons, which were poorly represented in the eDNA data. Several studies have demonstrated low or variable detection rates for exoskeleton-shedding fauna, such as crustaceans

(Crane et al., 2021; Korbel et al., 2024; Lancaster et al., 2025; Trimbo et al., 2020). In contrast, the sessile fauna that were well predicted were the sponges, tunicates and bryozoans, which consistently expulse DNA (Lancaster et al., 2025).

The few studies that have combined eDNA with other sampling methods report varying results. In rocky benthic habitats, for example, eDNA captured regional-scale (e.g., between and within seas) trends in macrobenthic biodiversity reasonably well when compared with morphological assessments (Cahill et al., 2018). Conversely, Czechowski et al. (2024) report that eDNA metabarcoding detected fewer fish species compared to concurrently deployed baited cameras, owing to the poor availability of relevant DNA sequences. These results align with our conclusion that eDNA approaches may be best suited to detect sessile, rather than mobile, taxa. However, eDNA-based detections of fish species from trawl-water runoff have yielded more accurate food web reconstructions compared to trawl survey samples alone (Cicala et al., 2024), highlighting the limitations of collection-based sampling schemes for assessing regional patterns. Collectively, these studies highlight the need for further research into the limitations of eDNA methods while underscoring the utility of metabarcoding as a complementary tool in a wide variety of systems (Deiner et al., 2017; Goldberg et al., 2016).

Performance and limitations of environmental DNA for habitat mapping

Overall, the eDNA-only and combined models performed poorly for discriminating presence and absence, but it is important to emphasize that the eDNA study design applied here was relatively small-scale (one composite sample per site, 25 sites, three markers, MiSeq 100 K sequencing depth). eDNA sensitivity or detection probability generally increases with larger water volumes, separate replicates, more sites, more markers and deeper sequencing—or some combination of these factors. eDNA may also be horizontally advected by currents, leading to ‘signal smearing’ and high spatial-autocorrelation between distant sample sites (Goldberg et al., 2016; present study). Empirical models of mesopelagic DNA transport suggest that DNA is generally detectable within 10 vertical meters of the depth at which it was shed (Allan et al., 2021). However, complex seabed topographies and fine-scale variation in substrate and texture may influence local water flow (Su et al., 2008), potentially transporting eDNA further from its source. Collecting water samples as close to the seabed as possible reduces this effect and provides the highest detectability of benthic fauna (Antich et al., 2021). Furthermore, collecting DNA from sediment

or other physical samples may provide more localized indicators of species presence (Hajibabaei et al., 2019).

Signal smearing can also be described as differences in spatial and temporal resolution between sampling methods. Video captures high spatial nuance on the seabed, making it possible to visually differentiate between adjacent habitats (e.g., sandy patch next to a boulder). However, video typically only provides a single view of the seabed, and taxa can easily be missed, especially in complex heterogeneous habitats. Additionally, video data shows what is present only at the moment of sampling, whereas eDNA shows what is there and what was there recently. This wider temporal scope creates some smoothing to the eDNA signal compared to the instantaneous measures derived from video. As such, eDNA provides a more comprehensive view of the ecological community, at the expense of spatial precision.

Another possible explanation for the poor performance of the eDNA and combined models is that eDNA could be capturing multiple life stages (Goldberg et al., 2016). Sequence matches due to pelagic larval stages would result in a taxonomic match but would not necessarily be associated with the benthic adult stage; thus, benthic terrain variables may be poor predictors of occurrence for these detections. As such, models may be improved by including water column parameters that influence larval dispersal, such as local current speed and direction (Kutti et al., 2020).

More research is needed on the sources and sinks of DNA in the marine environment (Barnes et al., 2014). DNA decay and shedding rates are active areas of research and are both species- and environment-specific (e.g., Barnes et al., 2014; Goldberg et al., 2016; Scriver et al., 2023). Characteristics of DNA molecules, abiotic environmental conditions (e.g., temperature, pH, dissolved oxygen, ultraviolet radiation exposure, etc.) and biotic interactions (e.g., extracellular enzymes, microbial metabolism), all affect DNA degradation rates (Barnes et al., 2014; Strickler et al., 2015), while DNA shedding rates are influenced by water temperature and organism abundance or biomass (Andruszkiewicz et al., 2021), as well as organism characteristics and responses to environmental conditions (Stewart, 2019). Much of this research has been conducted in mesocosm experiments, which may not accurately represent *in situ* ocean conditions.

Conclusions

Our study provides the first comparison between concurrently collected environmental DNA samples and drop camera videos for the application of habitat mapping. We highlight the utility of eDNA as a complementary tool for

surveying marine biodiversity and its potential to provide valuable insights for habitat and diversity mapping. First, our results highlight that incorporating metabarcoding data into biodiversity surveys can help mitigate some of the limitations of imagery-only datasets. Second, many sessile fauna that were difficult to distinguish in visual surveys, such as sponges, tunicates and bryozoans, were better represented by eDNA and best predicted by either the eDNA or combined models, although overall model predictive performance remained low. Finally, predictions from all three models highlighted similar areas of high taxonomic richness and similar RCPs, indicating that general patterns of biodiversity in Mortier Bay were recovered with opportunistic eDNA sampling.

In addition to these insights, we identified several limitations and opportunities for research to improve the utility of eDNA for habitat and diversity mapping. Further research on sources and sinks of DNA in the environment, comprehensive flow models, directed sampling methods and more robust reference sequence libraries would likely improve predictive modeling capacity where metabarcoding approaches are incorporated. Additionally, collecting eDNA from physical samples of surface-layer sediment to acquire DNA samples that are highly localized to the seafloor terrain variables could add fauna occurrences that might be missed with water samples and improve the predictive performance of models for habitat mapping. Finally, the spatial resolution of eDNA should be evaluated against the research questions under consideration for a given study; this could be measured in a pilot study to determine the magnitude and scale of sampling required. Although eDNA cannot yet replace video surveys for habitat mapping, it is a powerful tool to bolster traditional survey methods and warrants further study for benthic habitat and diversity mapping in complex coastal habitats.

Acknowledgements

This research was funded by Fisheries and Oceans Canada Coastal Environmental Baseline Program. Shreya Nemani received baseline funding from Memorial University's School of Graduate Studies as part of the Master of Science program. Additional funds were received by Principal Investigator K. Robert's Canada Research Chair in Ocean Mapping. We thank Adam Templeton and Kirk Regular for their support in data collection onboard the RV *D. Cartwright* provided by Memorial University's Fisheries and Marine Institute (Holyrood campus), Newfoundland and Labrador, Canada. We also thank Dr. John Shaw for providing processed bathymetry and backscatter data, ZC for brainstorming the initial concept, as well as ABYS, EB, MEC and ZMK for critical feedback on early

drafts of the paper. We thank Eman Elbakry, Kerry Hobrecker, Avery McCarthy and Hoda Rajabi for their assistance in processing the eDNA samples.

Author Contributions

Rylan J. Command: Conceptualization; methodology; formal analysis; visualization; project administration; writing – original draft; writing – review and editing. **Shreya Nemani:** Conceptualization; methodology; visualization; project administration; writing – review and editing; formal analysis. **Benjamin Misiuk:** Conceptualization; methodology; formal analysis; visualization; writing – review and editing. **Mehrdad Hajibabaei:** Methodology; supervision; funding acquisition; writing – review and editing. **Nicole Fahner:** Methodology; formal analysis; writing – review and editing. **Emily Porter:** Methodology; formal analysis; writing – review and editing. **Greg Singer:** Methodology; formal analysis; writing – review and editing. **Beverly McClenaghan:** Methodology; formal analysis; writing – review and editing. **Katleen Robert:** Conceptualization; supervision; funding acquisition; project administration; writing – review and editing.

Conflict of Interest Statement

RJC, SN, BM and KR declare they have no known conflicts of interest. At the time of writing, BMCC, NF, GS and EP are employees of eDNatec Inc. MH is the founder and Chief Scientific Officer of eDNatec Inc. This does not alter our adherence to policies on sharing data and materials.

Data Availability Statement

The video data are archived in the Federated Research Data Repository (<https://doi.org/10.20383/102.0687>). Genetic sequences are available in NCBI GenBank (PRJNA 1206193). All processed data and code required to reproduce the analyses in the paper are available on GitHub (<https://github.com/rjcommand/edna-mapping>).

References

- Abrego, N. & Ovaskainen, O. (2023) Evaluating the predictive performance of presence–absence models: why can the same model appear excellent or poor? *Ecology and Evolution*, **13**, e10784. <https://doi.org/10.1002/ece3.10784>
- Allan, E.A., DiBenedetto, M.H., Lavery, A.C., Govindarajan, A.F. & Zhang, W.G. (2021) Modeling characterization of the vertical and temporal variability of environmental DNA in the mesopelagic ocean. *Scientific Reports*, **11**(1), 21273. <https://doi.org/10.1038/1598-021-00288-5>

- Andruszkiewicz, E.A., Zhang, W.G., Lavery, A. & Govindarajan, A. (2021) Environmental DNA shedding and decay rates from diverse animal forms and thermal regimes. *Environmental DNA*, **3**(2), 492–514. <https://doi.org/10.1002/edn3.141>
- Antich, A., Palacín, C., Cebrian, E., Golo, R., Wangensteen, O.S. & Turon, X. (2021) Marine biomonitoring with eDNA: can metabarcoding of water samples cut it as a tool for surveying benthic communities? *Molecular Ecology*, **30**(13), 3175–3188. <https://doi.org/10.1111/mec.15641>
- Baird, D.J. & Hajibabaei, M. (2012) Biomonitoring 2.0: a new paradigm in ecosystem assessment made possible by next-generation DNA sequencing. *Molecular Ecology*, **21**(8), 2039–2044. <https://doi.org/10.1111/j.1365-294x.2012.05519.x>
- Barnes, M.A., Turner, C.R., Jerde, C.L., Renshaw, M.A., Chadderton, W.L. & Lodge, D.M. (2014) Environmental conditions influence eDNA persistence in aquatic systems. *Environmental Science & Technology*, **48**(3), 1819–1827. <https://doi.org/10.1021/es404734p>
- Bonfil, R., Palacios-Barreto, P., Vargas, O.U.M., Ricaño-Soriano, M. & Díaz-Jaimes, P. (2021) Detection of critically endangered marine species with dwindling populations in the wild using eDNA gives hope for sawfishes. *Marine Biology*, **168**(5), 1–12. <https://doi.org/10.1007/s00227-021-03862-7>
- Brown, C.J., Smith, S.J., Lawton, P. & Anderson, J.T. (2011) Benthic habitat mapping: a review of progress towards improved understanding of the spatial ecology of the seafloor using acoustic techniques. *Estuarine, Coastal and Shelf Science*, **92**(3), 502–520. <https://doi.org/10.1016/j.ecss.2011.02.007>
- Buhl-Mortensen, L., Olafsdottir, S.H., Buhl-Mortensen, P., Burgos, J.M. & Ragnarsson, S.A. (2015) Distribution of nine cold-water coral species (Scleractinia and Gorgonacea) in the cold temperate North Atlantic: effects of bathymetry and hydrography. *Hydrobiologia*, **759**(1), 39–61. <https://doi.org/10.1007/s10750-014-2116-x>
- Cahill, A.E., Pearman, J.K., Borja, A., Carugati, L., Carvalho, S., Danovaro, R. et al. (2018) A comparative analysis of metabarcoding and morphology-based identification of benthic communities across different regional seas. *Ecology and Evolution*, **8**(17), 8908–8920. <https://doi.org/10.1002/ece3.4283>
- Callahan, B.J., McMurdie, P.J., Rosen, M.J., Han, A.W., Johnson, A.J.A. & Holmes, S.P. (2016) DADA2: high-resolution sample inference from Illumina amplicon data. *Nature Methods*, **13**(7), 581–583. <https://doi.org/10.1038/nmeth.3869>
- Cicala, D., Maiello, G., Fiorentino, F., Garofalo, G., Massi, D., Sbrana, A. et al. (2024) Spatial analysis of demersal food webs through integration of eDNA metabarcoding with fishing activities. *Frontiers in Marine Science*, **10**, 1209093. <https://doi.org/10.3389/fmars.2023.1209093>

- Compson, Z.G., McClenaghan, B., Singer, G.A.C., Fahner, N.A. & Hajibabaei, M. (2020) Metabarcoding from microbes to mammals: comprehensive bioassessment on a global scale. *Frontiers in Ecology and Evolution*, **8**, 581835. <https://doi.org/10.3389/fevo.2020.581835>
- Crane, L.C., Goldstein, J.S., Thomas, D.W., Rexroth, K.S. & Watts, A.W. (2021) Effects of life stage on eDNA detection of the invasive European green crab (*Carcinus maenas*) in estuarine systems. *Ecological Indicators*, **124**, 107412. <https://doi.org/10.1016/j.ecolind.2021.107412>
- Czechowski, P., De Lange, M., Heldsinger, M., Kardailsky, A., Rayment, W., Hepburn, C. et al. (2024) Comparison of traditional and molecular surveys of fish biodiversity in southern Te Wāhipounamu/Fiordland (Aotearoa/New Zealand). *Environmental DNA*, **6**(1), e514. <https://doi.org/10.1002/edn3.514>
- Deiner, K., Bik, H.M., Mächler, E., Seymour, M., Lacoursière-Roussel, A., Altermatt, F. et al. (2017) Environmental DNA metabarcoding: transforming how we survey animal and plant communities. *Molecular Ecology*, **26**(21), 5872–5895. <https://doi.org/10.1111/mec.14350>
- Durden, J.M., Schoening, T., Althaus, F., Friedman, A., Garcia, R., Glover, A.G. et al. (2016) Perspectives in visual imaging for marine biology and ecology: from acquisition to understanding. *Oceanography and Marine Biology: An Annual Review*, **54**, 1–72. <https://doi.org/10.1201/9781315368597>
- Eddings, J.B. (2019) *The utility of environmental DNA and species distribution models in assessing the habitat requirements of twelve fish species in Alaskan north slope Rivers*. MSc, Logan, Utah: Utah State University.
- Foster, S.D., Givens, G.H., Dornan, G.J., Dunstan, P.K. & Darnell, R. (2013) Modelling biological regions from multi-species and environmental data: modelling regions of common profiles. *Environmetrics*, **24**(7), 489–499. <https://doi.org/10.1002/env.2245>
- Gibson, J.F., Shokralla, S., Curry, C., Baird, D.J., Monk, W.A., King, I. et al. (2015) Large-scale biomonitoring of remote and threatened ecosystems via high-throughput sequencing. *PLoS One*, **10**(10), e0138432. <https://doi.org/10.1371/journal.pone.0138432>
- Goldberg, C.S., Turner, C.R., Deiner, K., Klymus, K.E., Thomsen, P.F., Murphy, M.A. et al. (2016) Critical considerations for the application of environmental DNA methods to detect aquatic species. *Methods in Ecology and Evolution*, **7**(11), 1299–1307. <https://doi.org/10.1111/2041-210x.12595>
- Hajibabaei, M., Porter, T.M., Robinson, C.V., Baird, D.J., Shokralla, S. & Wright, M.T.G. (2019) Watered-down biodiversity? A comparison of metabarcoding results from DNA extracted from matched water and bulk tissue biomonitoring samples. *PLoS One*, **14**(12), e0225409. <https://doi.org/10.1371/journal.pone.0225409>
- Hanley, J.A. & McNeil, B.J. (1982) The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology*, **143**(1), 29–36. <https://doi.org/10.1148/radiology.143.1.7063747>
- Howell, K.L., Bullimore, R.D. & Foster, N.L. (2014) Quality assurance in the identification of deep-sea taxa from video and image analysis: response to Henry and Roberts. *ICES Journal of Marine Science*, **71**(4), 899–906. <https://doi.org/10.1093/icesjms/fsu052>
- Jaquier, M., Albouy, C., Bach, W., Waldock, C., Marques, V., Maire, E. et al. (2021) Environmental DNA diffusion of reef fishes along a distance gradient from four isolated islands of the Western Indian Ocean. *ARPHA Conference Abstracts*, **4**, e65060. <https://doi.org/10.3897/aca.4.e65060>
- Kohonen, T. (1982) Self-organized formation of topologically correct feature maps. *Biological Cybernetics*, **43**(1), 59–69. <https://doi.org/10.1007/BF00337288>
- Korbel, K.L., Hose, G.C., Karwautz, C., Greenfield, P., Wang, H., Chariton, A.A. et al. (2024) Detection, movement and persistence of invertebrate eDNA in groundwater. *Scientific Reports*, **14**, 17151. <https://doi.org/10.1038/s41598-024-67349-3>
- Kutti, T., Johnsen, I.A., Skaar, K.S., Ray, J.L., Husa, V. & Dahlgren, T.G. (2020) Quantification of eDNA to map the distribution of cold-water coral reefs. *Frontiers in Marine Science*, **7**, 446. <https://doi.org/10.3389/fmars.2020.00446>
- Lancaster, E.R., Grey, E.K., Brady, D.C. & Frederich, M. (2025) Squishy and crunchy invasive invertebrates: environmental DNA is not shed equally. *Environmental DNA*, **7**, e70072. <https://doi.org/10.1002/edn3.70072>
- Lathrop, R.G., Cole, M., Senyk, N. & Butman, B. (2006) Seafloor habitat mapping of the New York bight incorporating sidescan sonar data. *Estuarine, Coastal and Shelf Science*, **68**(1–2), 221–230. <https://doi.org/10.1016/j.ecss.2006.01.019>
- Lecours V. (2017) Terrain attribute selection for spatial ecology (TASSE). ArcGIS toolbox version, 1(10.13140).
- Lecours, V., Devillers, R., Schneider, D., Lucieer, V., Brown, C. & Edinger, E. (2015) Spatial scale and geographic context in benthic habitat mapping: review and future directions. *Marine Ecology Progress Series*, **535**, 259–284. <https://doi.org/10.3354/meps11378>
- Lundblad, E.R., Wright, D.J., Miller, J., Larkin, E.M., Rinehart, R., Naar, D.F. et al. (2006) A benthic terrain classification scheme for American Samoa. *Marine Geodesy*, **29**(2), 89–111. <https://doi.org/10.1080/01490410600738021>
- Mallet, D. & Pelletier, D. (2014) Underwater video techniques for observing coastal marine biodiversity: a review of sixty years of publications (1952–2012). *Fisheries Research*, **154**, 44–62. <https://doi.org/10.1016/j.fishres.2014.01.019>
- Misiuk, B. & Brown, C.J. (2024) Benthic habitat mapping: a review of three decades of mapping biological patterns on the seafloor. *Estuarine, Coastal and Shelf Science*, **296**, 108599. <https://doi.org/10.1016/j.ecss.2023.108599>
- Muenzel, D., Bani, A., De Brauwier, M., Stewart, E., Djakiman, C., Halwi et al. (2024) Combining environmental DNA and

- visual surveys can inform conservation planning for coral reefs. *Proceedings of the National Academy of Sciences*, **121** (17), e2307214121.
- Nemani, S., Cote, D., Misiuk, B., Edinger, E., Mackin-McLaughlin, J., Templeton, A. et al. (2022) A multi-scale feature selection approach for predicting benthic assemblages. *Estuarine, Coastal and Shelf Science*, **277**, 108053. <https://doi.org/10.1016/j.ecss.2022.108053>
- Nemani, S., Misiuk, B., Cote, D., Edinger, E., Mackin-McLaughlin, J., Templeton, A. et al. (2024) Incorporating functional traits with habitat maps: patterns of diversity in coastal benthic assemblages. *Frontiers in Marine Science*, **11**, 1141737. <https://doi.org/10.3389/fmars.2024.1141737>
- Oksanen, J., Simpson, G.L., Blanchet, F.G., Kindt, R., Legendre, P., Minchin, P.R., et al. (2022) vegan: community ecology package. Available from <https://CRAN.R-project.org/package=vegan>.
- Ovaskainen, O. & Abrego, N. (2020) *Joint species distribution modelling with applications in R*. Cambridge, United Kingdom: Cambridge University Press. 370. <https://doi.org/10.1017/9781108591720>
- Ovaskainen, O., Tikhonov, G., Norberg, A., Guillaume Blanchet, F., Duan, L., Dunson, D. et al. (2017) How to make more out of community data? A conceptual framework and its implementation as models and software. *Ecology Letters*, **20** (5), 561–576. <https://doi.org/10.1111/ele.12757>
- Pearce, J. & Ferrier, S. (2000) Evaluating the predictive performance of habitat models developed using logistic regression. *Ecological Modelling*, **133**(3), 225–245. [https://doi.org/10.1016/S0304-3800\(00\)00322-7](https://doi.org/10.1016/S0304-3800(00)00322-7)
- Pichler, M. & Hartig, F. (2021) A new joint species distribution model for faster and more accurate inference of species associations from big community data. *Methods in Ecology and Evolution*, **12**(11), 2159–2173. <https://doi.org/10.1111/2041-210X.13687>
- Potter, D.P. & Shaw, J. (2009a) *Shaded seafloor relief, Placentia Bay west, offshore Newfoundland and Labrador*. Natural Resources Canada: Geological Survey of Canada. "A" Series Map, 2144A. <https://doi.org/10.4095/247331>
- Potter, D.P. & Shaw, J. (2009b) *Shaded seafloor relief, Placentia Bay southwest, offshore Newfoundland and Labrador*. Natural Resources Canada: Geological Survey of Canada. "A" Series Map, 2146A. <https://doi.org/10.4095/247333>
- R Core Team. (2021) *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing.
- Riaz, M., Kuemmerlen, M., Wittwer, C., Cocchiara, B., Khaliq, I., Pfenninger, M. et al. (2020) Combining environmental DNA and species distribution modeling to evaluate reintroduction success of a freshwater fish. *Ecological Applications*, **30**(2), e02034. <https://doi.org/10.1002/eap.2034>
- Ruppert, K.M., Kline, R.J. & Rahman, S. (2019) Past, present, and future perspectives of environmental DNA (eDNA) metabarcoding: a systematic review in methods, monitoring, and applications of global eDNA. *Global Ecology and Conservation*, **17**, 547. <https://doi.org/10.1016/j.gecco.2019.e00547>
- Schlining, B. & Jacobsen Stout, N. (2006) MBARI's video annotation and reference system. In: Proceedings of the Marine Technology Society/Institute of Electrical and Electronics Engineers Oceans Conference, Boston, Massachusetts. <https://doi.org/10.1109/OCEANS.2006.306879>
- Scriver, M., Zaiko, A., Pochon, X. & Von Ammon, U. (2023) Harnessing decay rates for coastal marine biosecurity applications: a review of environmental DNA and RNA fate. *Environmental DNA*, **5**, edn3.405. <https://doi.org/10.1002/edn3.405>
- Shaw, J., Potter, D. & Kostylev, V.E. (2011) *Seascapes, Placentia Bay, Newfoundland and Labrador*. Geological Survey of Canada = Commission Géologique du Canada. Open File 6683. <https://doi.org/10.4095/288644>
- Shaw, J., Puig, P. & Han, G. (2013) Megaflores in a continental shelf setting, Placentia Bay, Newfoundland. *Geomorphology*, **189**, 12–25. <https://doi.org/10.1016/j.geomorph.2013.01.010>
- Stat, M., John, J., DiBattista, J.D., Newman, S.J., Bunce, M. & Harvey, E.S. (2019) Combined use of eDNA metabarcoding and video surveillance for the assessment of fish biodiversity. *Conservation Biology*, **33**(1), 196–205. <https://doi.org/10.1111/cobi.13183>
- Stewart, K.A. (2019) Understanding the effects of biotic and abiotic factors on sources of aquatic environmental DNA. *Biodiversity and Conservation*, **28**(5), 983–1001. <https://doi.org/10.1007/s10531-019-01709-8>
- Strickler, K.M., Fremier, A.K. & Goldberg, C.S. (2015) Quantifying effects of UV-B, temperature, and pH on eDNA degradation in aquatic microcosms. *Biological Conservation*, **183**, 85–92. <https://doi.org/10.1016/j.biocon.2014.11.038>
- Su, D.T., Liu, T.L. & Ou, C.H. (2008) Numerical investigation into effects of seabed topography on flows in and around artificial reefs. *Fisheries Science*, **74**(2), 236–254. <https://doi.org/10.1111/j.1444-2906.2008.01518.x>
- Tikhonov, G., Duan, L., Abrego, N., Newell, G., White, M., Dunson, D. et al. (2020) Computationally efficient joint species distribution modeling of big spatial data. *Ecology*, **101**(2), e02929. <https://doi.org/10.1002/ecy.2929>
- Tikhonov, G., Ovaskainen, O., Oksanen, J., de Jonge, M.M.J., Opedal, Ø.H., and Dallas, T. (2022) Hmsc: hierarchical model of species communities. Available from <https://www.helsinki.fi/en/researchgroups/statistical-ecology/hmsc>.
- Tjur, T. (2009) Coefficients of determination in logistic regression models—a new proposal: the coefficient of discrimination. *The American Statistician*, **63**(4), 366–372. <https://doi.org/10.1198/tast.2009.08210>
- Trimbos, K.B., Cieraad, E., Schrama, M., Saarloos, A.I., Musters, K.J.M., Bertola, L.D. et al. (2020) Stirring up the relationship between quantified environmental DNA

- concentrations and exoskeleton-shedding invertebrate densities. *Environmental DNA*, **3**, 605–618. <https://doi.org/10.1002/edn3.157>
- Valdivia-Carrillo, T., Rocha-Olivares, A., Reyes-Bonilla, H., Domínguez-Contreras, J.F. & Munguia-Vega, A. (2021) Integrating eDNA metabarcoding and simultaneous underwater visual surveys to describe complex fish communities in a marine biodiversity hotspot. *Molecular Ecology Resources*, **21**, 1558–1574. <https://doi.org/10.1111/1755-0998.13375>
- Walbridge, S., Slocum, N., Pobuda, M. & Wright, D. (2018) Unified geomorphological analysis workflows with benthic terrain modeler. *Geosciences*, **8**(3), 94. <https://doi.org/10.3390/geosciences8030094>
- Wehrens, R. & Buydens, L.M.C. (2007) Self- and super-organizing maps in R: the kohonen package. *Journal of Statistical Software*, **21**(5), 1–19. <https://doi.org/10.18637/jss.v021.i05>
- Wehrens, R. & Kruisselbrink, J. (2018) Flexible self-organizing maps in kohonen 3.0. *Journal of Statistical Software*, **87**(7), 1–18. <https://doi.org/10.18637/jss.v087.i07>
- Zvoleff, A. (2020) glcm: calculate textures from grey-level co-occurrence matrices (GLCMs) version 1.6.5 from CRAN. Available from: <https://rdrr.io/cran/glcm/>.

Supporting Information

Additional supporting information may be found online in the Supporting Information section at the end of the article.

Table S1. Details of the primers used to amplify DNA sequences.

Table S2. List of taxa identified by video ground truth and/or eDNA data. Blanks occur under “Authority” where no taxonomic authority is listed in the World Register of Marine Species.

Table S3. Method used to combine video and filtered eDNA datasets. In the combined dataset, a taxon was considered to be present if it was detected by either video or eDNA methods.

Table S4. Taxon-specific cross-validated discrimination metrics for each model.

Table S5. Taxon-specific cross-validated performance metrics for each model.

Table S6. Potential scale-reduction factors for each model and parameter.

Table S7. Environmental features selected by RDA for each model, and the percent variance explained by each feature in their respective model. For each dataset, the features are ordered by percent variance explained.

Figure S1. Taxonomic families identified by eDNA and Video in Mortier Bay. Only those taxa that exhibit benthic or benthopelagic lifestyle and are generally > 2 cm

are displayed. Vertical dashed line indicates the threshold (3/25 stations) used to filter out rare taxa prior to modeling.

Figure S2. Residual co-occurrence plots for the combined dataset, comparing the full model (A) with the spatial-only model (B). It can be seen in panel A that, after accounting for the environmental variables, the remaining co-occurrences have less than 90% posterior support.

Figure S3. Residual co-occurrence plots for the eDNA-only dataset, comparing the full model (A) with the spatial-only model (B). It can be seen in panel A that, after accounting for the environmental variables, the remaining co-occurrences have less than 95% posterior support.

Figure S4. Residual co-occurrence plots for the video-only dataset, comparing the full model (A) with the spatial-only model (B). It can be seen in panel A that, after accounting for the environmental variables, the remaining co-occurrences have less than 95% posterior support.

Figure S5. Spearman's rank correlations between variables selected by RDA for each model.

Figure S6. The responses of taxa to environmental covariates for the combined model. Responses that are positive with at least 95% posterior probability are shown by red, responses that are negative with at least 95% posterior probability are shown by blue, and responses that did not gain strong statistical support are shown by white.

Figure S7. The responses of taxa to environmental covariates for the video-only model. Responses that are positive with at least 95% posterior probability are shown by red, responses that are negative with at least 95% posterior probability are shown by blue, and responses that did not gain strong statistical support are shown by white.

Figure S8. The responses of taxa to environmental covariates for the eDNA-only model. Responses that are positive with at least 95% posterior probability are shown by red, responses that are negative with at least 95% posterior probability are shown by blue, and responses that did not gain strong statistical support are shown by white.

Figure S9. Histograms of the potential scale-reduction factors (PSRF) for the combined model. Top row are the PSRFs for the full model, middle row are the PSRFs for the environmental covariate-only model, bottom row are the PSRFs for the spatial-only model. Plots are paired, with each parameter estimate showing a point estimate and the upper 95% confidence interval.

Figure S10. Histograms of the potential scale-reduction factors (PSRF) for the eDNA-only model. Top row are the PSRFs for the full model, middle row are the PSRFs for the environmental covariate-only model, bottom row are the PSRFs for the spatial-only model. Plots are paired, with each parameter estimate showing a point estimate and the upper 95% confidence interval.

Figure S11. Histograms of the potential scale-reduction factors (PSRF) for the video-only model. Top row are the PSRFs for the full model, middle row are the PSRFs for the environmental covariate-only model, bottom row are the PSRFs for the spatial-only model. Plots are paired, with each parameter estimate showing a point estimate and the upper 95% confidence interval.

Figure S12. Predicted taxonomic richness as a function of each environmental feature from the combined model. Points are the observed taxonomic richness values for each site, lines are predicted response over the range of each covariate, with the grey ribbon indicating 95% credible interval. Level of posterior support for the hypothesis that the minimum richness is less than or greater than the maximum richness is found in the title of each plot.

Figure S13. Predicted taxonomic richness as a function of each environmental feature from the eDNA-only model. Points are the observed taxonomic richness values for each site, lines are predicted response over the range of each covariate, with the grey ribbon indicating 95% credible interval. Level of posterior support for the hypothesis that the minimum richness is less than or greater than the maximum richness is found in the title of each plot.

Figure S14. Predicted taxonomic richness as a function of each environmental feature from the video-only model. Points are observed taxonomic richness values for each site, lines are average response over the range of each covariate, with the grey ribbon indicating 95% credible interval. Level of posterior support for the hypothesis that the minimum richness is less than or greater than the maximum richness is found in the title of each plot.

Figure S15. Predicted occurrence probabilities for each taxa in each node of the 5×5 self-organizing map fit to the combined dataset. Colour represents the YUV colour space mapping of the nodes, and size represents the predicted occurrence probability for the given taxa.

Figure S16. Predicted occurrence probabilities for each taxa in each node of the 5×5 self-organizing map fit to the eDNA-only dataset. Colour represents the YUV colour space mapping of the nodes, and size represents the predicted occurrence probability for the given taxa.

Figure S17. Predicted occurrence probabilities for each taxa in each node of the 5×5 self-organizing map fitted to the video-only dataset. Colour represents the YUV colour space mapping of the nodes, and size represents the predicted occurrence probability for the given taxa.

Appendix S1.