Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

A cross-national network for environmental DNA monitoring of nearshore fish biodiversity across the Northeast Pacific

Jennifer M. Sunday

Jennifer.sunday@mcgill.ca

McGill University

Kate Sheridan

McGill University

Margot Hessing-Lewis

Hakai Institute

Rute B. G. Clemente-Carvalho

Hakai Institute

Salvador Babinet

McGill University

Jennifer L. Ruesink

University of Washington

Lillian R. Aoki

University of Oregon

Lia K. Domke

NOAA National Marine Fisheries Service Alaska Fisheries Science Center

Sue Velazquez

Hakai Institute

Deanna Beatty

University of California, Davis

Sarah Joy Bittick

Loyola Marymount University

Katharyn E. Boyer

San Francisco State University

Hannah Bregulla

Council of the Haida Nation

Margot V. Buchbinder

San Francisco State University

Alyssa A. B. Cooper

Sonoma State University

Ginny L. Eckert
University of Alaska Fairbanks

Scott S. Gabara
Channel Islands National Park

Olivia J. Graham
Cornell University

Drew Harvell
Cornell University

Kevin A. Hovel
San Diego State University

Joel Huckeba
Sonoma State University

Brent B. Hughes
Sonoma State University

Andrew Kinziger
Cal Poly Humboldt

Lynn C. Lee
Gwaii Haanas National Park Reserve, National Marine Conservation Area Re

Evan Morien
Hakai Institute

Ryan S. Mueller
Oregon State University

Zachary Monteith
Hakai Institute

Makenzie O'Connor
Morro Bay National Estuary Program

Angeleen M. Olson
McGill University

Kristin M. Robinson
Hakai Institute

Julie B. Schram
University of Alaska Southeast

Kathryn L. Sobocinski
Western Washington University

John J. Stachowicz
University of California, Davis

Fiona Tomas
Mediterranean Institute for Advanced Studies

Jennifer Yakimishyn
Pacific Rim National Park Reserve, Parks Canada Agency
Matthew A. Lemay
Hakai Institute

Article

Keywords: environmental DNA, biodiversity observation, seagrass monitoring, network, marine biogeography

Posted Date: July 9th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-10282151/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

A cross-national network for environmental DNA monitoring of nearshore fish biodiversity across the Northeast Pacific

Authors: Jennifer M. Sunday, Kate Sheridan, Margot Hessing-Lewis, Rute B. G. Clemente-Carvalho, Salvador Babinet, Jennifer L. Ruesink, Lillian R. Aoki, Lia K. Domke, Sue Velazquez, Deanna Beatty, Sarah Joy Bittick, Catherine E. Boyer, Hannah Bregulla, Margot V. Buchbinder, Alyssa A.B. Cooper, Stuart Crawford, Emmett Duffy, Ginny L. Eckert, Scott S. Gabara, Olivia J. Graham, Drew Harvell, Kevin A. Hovel, Joel Huckeba, Brent B. Hughes, Andrew Kinziger, Lynn C. Lee, Zachary Monteith, Evan Morien, Ryan S. Mueller, Lisa Needles, Makenzie O'Connor, Angeleen M. Olson, Kristin M. Robinson, Julie B. Schram, Kathryn L. Sobocinski, John J. Stachowicz, Fiona Tomas, Jennifer Yakimishyn, Matthew A. Lemay

Abstract

Biodiversity monitoring using environmental DNA (eDNA) stands to increase the scale at which we track biodiversity change, yet coordination is necessary to allow insightful comparisons. Here, we present an expansive fish biodiversity dataset from an international eDNA monitoring network – the *Pacific eDNA Coastal Observatory* – covering 26 degrees of latitude and four years of coordinated seawater sampling with metabarcoding using MiFish primers. The network is distributed across seagrass habitats in 16 regional nodes from San Diego, California to Juneau, Alaska. The dataset provides findings from 1041 samples collected between 2021 and 2024, with detections of 64 fish families and 136 identified species from northern temperate to subtropical fauna. We validate these data using beach seining conducted in 12 regions adjacent to eDNA collections, and find 77% of species captured by beach seine were detected by eDNA at the genus or species level (with a range from 49-100% within regions). Within-year biodiversity coverage varied among regions from 70-98%. These data demonstrate how the networked approach augments existing monitoring programs for comparative large-scale inference.

Keywords: environmental DNA, biodiversity observation, seagrass monitoring, network, marine biogeography

Introduction

Monitoring marine biodiversity at large spatial scales is needed to understand nature’s responses to global change (Duffy et al. 2013, Canonico et al. 2019). Environmental DNA offers a scalable, non-invasive means of detecting marine organisms from water samples, potentially improving monitoring of rare, cryptic, or difficult-to-sample taxa across broader spatial and temporal scales than many conventional survey methods.

Genomic analysis of environmental DNA (eDNA), for example, is a powerful tool that leverages the temporary stability of eDNA-containing particles in aquatic environments shed from live organisms. Isolated eDNA from a standard water sample (e.g., 1 litre) can contain the fingerprint of an entire nearby community (Shelton et al. 2016) especially when sampled in replicate to overcome undersampling of less common species (Grey et al. 2018, Beentjes et al. 2019). A myriad of eDNA sampling programs that focus on targeted species or whole communities is now emerging globally (Thomsen et al. 2024), yet few use a networked approach to allow data from across scales to be directly comparable (Santi et al. 2023, Sanchez et al. 2025).

Robust detection of biodiversity changes across time and space benefit from a spatially and temporally balanced sampling effort. It has long been established, for example, that methods of sampling biodiversity are biased towards more common and/or more detectable species, such that the level of sampling effort influences what kind and how many species are detected (Gotelli and Colwell 2001, Magurran and McGill 2011). If sampling effort varies across time and space it is challenging to disentangle sampling effects from true biological responses. Similarly, tracking biogeographical range shifts – a global fingerprint of climate change (Pecl et al. 2017, Lawlor et al, 2024) – requires large-scale comparable efforts of detection across time and space; unbalanced effort can easily confound range shift estimates (Bates et al. 2015). Although it is possible to combine data derived from different monitoring programs, the clearest comparisons come from coordinated programs, in which sampling follows the same methods. eDNA has emerged as a promising tool for detecting assemblage and species changes through time (Thomsen et al. 2024) but coordination of effort across space and time is needed for the full potential of biodiversity monitoring with eDNA to be reached (Canonico et al. 2019, Leinen et al. 2022, Santi et al. 2023). In addition, sharing gear, methods, and genomic processing costs can increase collective access to genomics technologies, improving spatial and temporal balance of monitoring efforts that can strengthen inference at larger scales (Stammnitz et al. 2024).

Here we present a new dataset resulting from a collaborative international eDNA observatory that has taken a networked approach to eDNA sampling: the Pacific eDNA Coastal Observatory (PECO). The observatory is composed of 18 partners in 16 regions spanning 26 degrees of latitude along the Northeastern Pacific coast, from San Diego, California, to Juneau, Alaska. Synchronized annual sampling since 2021 across this latitudinal transect now tracks four years of fish biogeography in coastal seagrass beds. Here, we describe the resulting biodiversity data that resolves to 64 families of fish, most to species-level. We take three approaches to evaluate eDNA sampling quality: overlap with adjacent capture surveys in one of the years; overlap with the available taxonomic reference databases; and how well the sampling design captured

sample, local, and regional richness. These data provide the foundation for an ongoing biomolecular time-series with potential to deepen our understanding of responses to global change.

Methods

Our methods include site selection, sampling kit deployment, seawater sampling and processing, laboratory processing of samples, and quality assurance and control. These are followed by several steps for data validation including beach seining surveys and assessment of biodiversity coverage. This paper reports on the data collected between 2021-2024; however, continued sampling following these protocols is planned for subsequent years.

Network development

Partnerships were established by a core leadership team reaching out to a collective and growing network of field ecologists, many of whom had ongoing nearshore field monitoring programs. We made early efforts when establishing the network to achieve a balanced sampling design across latitude (Fig.1), with at least two partner-sampled regions on either side of well-described biogeographic breaks at Point Conception, Cape Mendocino, Juan de Fuca Strait, and Hecate Strait (Valentine, 1966; Blanchette et al., 2008; Fenberg et al., 2015; Miller, 2023). Because eDNA metabarcoding can detect community differences across habitat types at scales of 100s of meters (Millard-Martin et al. 2024), we selected a similar focal habitat in each region to help ensure comparability of community assemblages downstream. Eelgrass meadows characterized by *Zostera marina* are known to be key habitat for coastal and benthic fish (Unsworth et al. 2019) and were chosen for our focal habitat sampling. Within regions, specific sampling sites were proposed by partners to increase feasibility, increase capacity for local knowledge, and maximize co-benefits to partners, for example by including sites visited regularly by partners for other sampling programs. Partners received sampling materials, protocols, and a protocol video; additionally a discussion of the protocol was conducted by teleconference each year. In some cases, more than one partnership was established within regions, or partners changed across years, in which case, we grouped partners within regions for the purposes of spatial analyses (Fig. 1).

The overall design was to visit five sites per region within the same one-month window (Jul 15 - Aug 15) every summer, with sites located 2-20 km apart, and collect four 1.5 L samples of seawater and one field blank at every site, with samples separated by 10 m along a line that followed the shoreline through the seagrass meadow. Some methods changed over the four years, and some deviations occurred within years at regions (see “Changes and Deviations in Field Methods” section below).

Preparation and distribution of sample kits

Each partner was provided with a standardized sampling kit, prepared at the Hakai Institute on Quadra Island in British Columbia (BC), Canada. The sampling kit materials and protocol are published separately (Sunday et al. 2024), and we summarize these briefly here. The kit included sterile 1.5 L Nalgene bottles, a custom-made bottle-mount from PVC piping that could be mounted on the end of a PVC pipe or stick, labelled Sterivex filters (0.22 µm) for each sample, an air-pressure spray bottle modified to allow pressurized air to flow into air tubing, an inline HEPA filter, a modified Nalgene bottle lid that allowed air pressure to displace sample water into the tubing (see Fig. S1), tubing, 50 mL syringes, storage buffer, sterile gloves, and bench pads. All non-sterile materials were sterilized with 10% bleach, rinsed, dried, and subject to UV light treatment before being packed into Ziploc and Whirl-Pak bags, and further packed into a sterile extra-large tote. The air-pressure pump was not sterilized with bleach because it contained metal parts that would react, and therefore was wiped down with 70% Ethanol between years.

Sample collection

Full details of the sampling protocol are described in the PECO eDNA field sampling protocol (Sunday et al. 2024). Briefly, at each site, four water samples of 1.5 L were collected along a transect parallel to shore, 10 m apart, within 1.5 m above the bottom with *Z. marina* present. In addition, an identical bottle of filtered water was opened and closed at each site to serve as a negative control. Samples were filtered into 0.22 µm Sterivex filters using a low-cost sterile positive pressure pumping device (see protocol; Fig. S1). Samples were kept in the dark at 4°C and shipped within two weeks of sampling to the Hakai Institute Genomics Laboratory on Quadra Island, BC, Canada. Upon arrival, samples were stored at -20°C until processing began within two months of field collection.

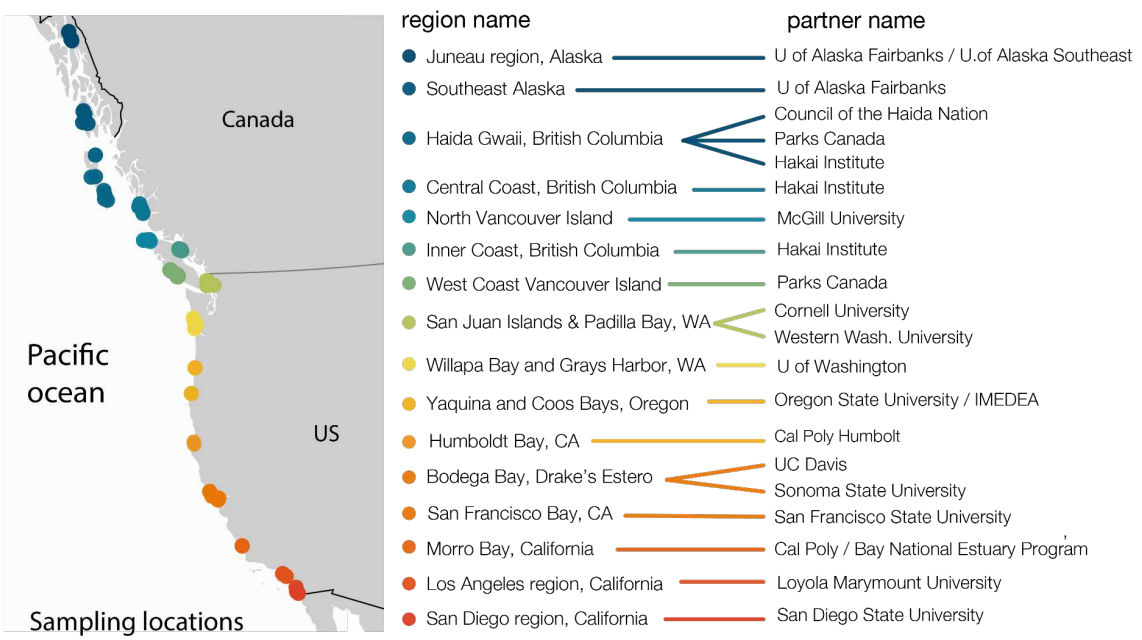


Figure 1. Sampling locations and partner network of the Pacific eDNA Coastal Observatory. Points indicate approximate sampling sites combining all years.

Changes and deviations in field methods

Across four years of networked seawater sampling for eDNA, some elements of our field design changed. Due to constraints of field access or doubling-up of partners within regions, the number of sites within regions was sometimes greater or less than 5, but over the course of the program we made efforts to balance this out to 5 sites per region (Fig. S2). Similarly, in the first year (2021), all sampling was aimed to be conducted between the dates of Aug 1 and 30, but through discussion as a network, it was decided to move this 1-month window to Jul 15 to Aug 15 in subsequent years.

Our methods sometimes deviated from the protocols in terms of timing, number of samples collected, and total volume of water filtered per sample. Some collections deviated from the target time windows over the years (Fig. S3). Although we initially set out to filter 1 L of seawater per sample, the turbidity of seawater at many sites prevented this volume from being achieved given the system we used for pressurizing samples, and most partners reduced volume to 500 mL (Fig. S4). We provide all specific details as region-level and sample-metadata for investigation of the sensitivity of results.

Lab protocols

We extracted DNA from Sterivex filters using a modified version of the Qiagen Blood and Tissue Kit, described in detail here: *DNA Extraction from 0.22 µm Sterivex Filters -*

Qiagen Blood and Tissue Kit (Kellogg et al. 2024). DNA extracts were used to amplify a fragment of the 12S rRNA gene according to the dual-PCR metabarcoding protocol described by Bourlat *et al.* (Bourlat et al. 2016). Briefly, the first-step PCR was performed using MiFish-U primers (Miya et al., 2015) containing the Illumina overhang adapters. In the second-step PCR, dual indexes were attached to the amplicon generated from the first PCR round. We introduced a single modification to the original protocol during this second step: Illumina DNA/RNA UD Indexes were used instead of Nextera indexes. This substitution was implemented to ensure unique dual index combinations for each sample, and mitigate the risk of index hopping during sequencing. Final pooled libraries typically yield two distinct bands: the target amplicon (~350 bp) and an undesired bacterial amplification product (~450 bp). To isolate the target, the final library was run on a 3% agarose gel. The target band was then excised and purified using the Wizard SV Gel and PCR Clean-up Kit (Promega) according to the manufacturer's instructions. We sequenced each pooled library (one library for each year of sampling) on a NextSeq 1000 instrument (Illumina) using the P1 600 kit. To optimize performance for a low-diversity library, we followed Illumina technical support recommendations by loading the library at 650 pM with a 40% PhiX spike. For the full detailed workflow please see the protocol described by Carvalho et al. (2024).

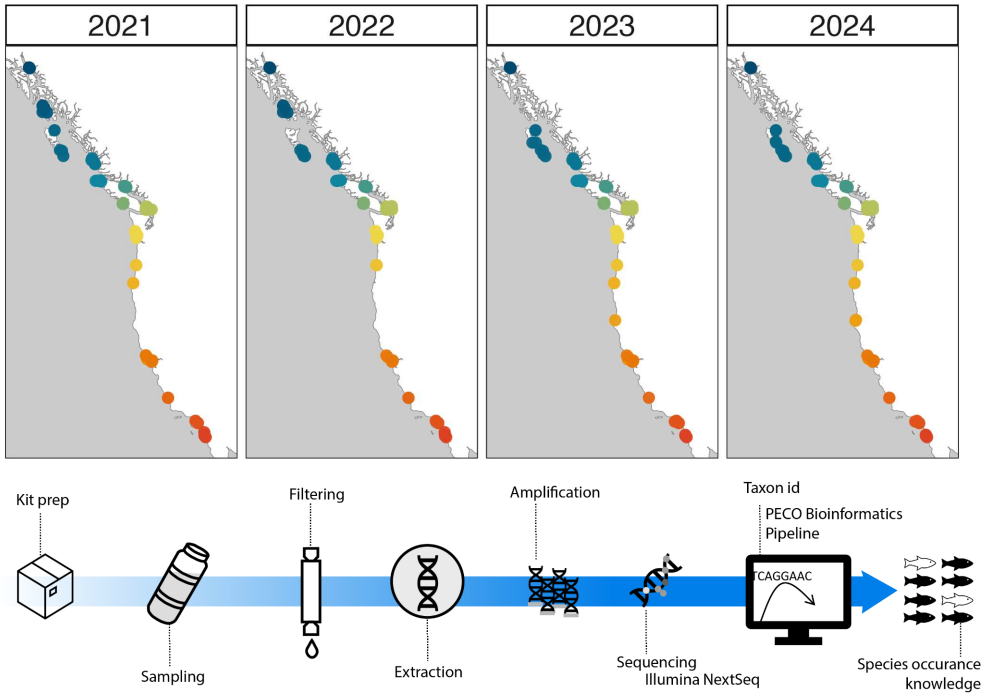


Figure 2. Annual map and pipeline of activity of the PECO network. Points represent locations of eDNA samples in each annual sampling event year.

Bioinformatics

The first step of our processing pipeline was applied to the dataset from each year separately while later steps were applied to all years together (Fig. S5). First, sequence reads with ambiguity ('*N*'s in place of nucleotides) or with fewer than 50bp were removed during primer trimming with cutadapt v4.4 (Martin 2011) and remaining reads were filtered and trimmed, then denoised with dada2 (Callahan et al. 2016) keeping only reads >110bp, and truncation quality (truncQ) of 2 or higher. PhiX control sequences were removed. After denoising, de-duplication, merging of paired-end reads, and amplicon sequence variant (ASV) inference with default dada2 parameters, ASVs present at extremely low abundance (<0.0001%) and found only within a single sample were discarded (both conditions needed to be met for removal at this stage), and chimeras were removed with dada2's removeBimeraDenovo function, using the "pooled" option. Read counts throughout these quality control steps are shown in Table S1. Each step was conducted in R v4.3.1 (R Core Team 2023).

We combined ASV sequence tables across years and collapsed them to unique variants before assigning each a unique identifier (e.g. ASV99). This assured that identical sequence variants encountered across years were given the same alphanumeric ID, were handled in the same way during taxonomic assignment, and could co-contribute to the clustering algorithm (in the next step). We clustered the denoised ASVs into MOTUs using SWARM v2 (using default settings, D=1, Mahé et al. 2015), and curated the resulting clusters with LULU (Frøslev et al. 2017), which merges the daughter and parent OTUs based on co-occurrence in samples.

To assign each ASV and MOTU to a taxonomy, we used the CRABS pipeline (Jeunen et al. 2023) to build a custom database of 12S sequences from MitoFish (Sato et al. 2018) and NCBI, and a subset of likely non-fish matches (i.e. bacteria, invertebrates, cows, whales, humans, algae, etc.). The non-fish matches were sourced from BLAST results from other eDNA projects, and their inclusion allowed higher taxonomic levels to be inferred for the majority of non-fish sequences. While pipeline defaults were sufficient for most steps, during the pairwise global alignment step of the CRABS pipeline, we set the threshold for percent identity and coverage to 0.75 in order to retain the maximum number of sequences that were not identified by the previous in-silico PCR step to explicitly contain the MiFish primer sequences. We BLASTed ASVs against this database and retained all matches above (90%) percent identity or the first 50, whichever was lower.

Before assignment, all taxonomy was converted to the WoRMS taxonomic backbone (Chamberlain and Vanhoorne 2024) using the *R* packages *taxize* 0.9.10 (Chamberlain et al. 2020) and *worrms* 0.4.3 (Chamberlain and Vanhoorne 2024), to collapse invalid

and synonymized names. Since it was possible for a sequence to align best with a fish species unlikely in the the northeast Pacific, we curated a list of 'non-regional' species according to the range descriptions from FishBase (Boettiger et al. 2012) or occurrence records from the Ocean Biodiversity Information System (OBIS)(OBIS 2025) to remove from the BLAST results. We used increasingly stringent thresholds to filter out BLAST hits and identify a "best match" species based on sequence percent identity. If there was only one putative species with 98.5% percent identity or higher, it was assigned. If an ASV was associated with one or more species at a level between 100 percent identity and above 97 percent identity, it was assigned to the species with the highest percent identity so long as it was 0.5% greater than other species. If there were multiple species with percent identity that were within 0.5% of each other, the ASV was assigned to the closest shared taxonomic level (i.e. the Lowest Common Ancestor method). Below 96.9 percent identity, sequences were assigned to the closest shared taxonomic level above species to permit retention of potentially unsequenced or cryptic species. MOTUs were assigned a taxonomy based percent identity of each ASV assignment within the cluster with the same method as above. In a few cases, the MOTU consensus assignment had a higher taxonomic rank than the direct ASV assignment; in all cases, we retained both assignments for different potential downstream uses. We recommend use of MOTU-based assignment as the default when a more conservative (less precise) taxonomic assignment is desired. After taxonomic assignment, sequences assigned to anything other than fish were removed for analysis.

Quality assurance and control

Quality control was conducted with negative controls at three levels using field blanks, DNA extraction blanks, and PCR blanks. Field blanks were collected alongside sample collections in the field by preparing a sample bottle with distilled or locally bottled water, opening the bottle above water in the field setting at each site and moving it gently for approximately five seconds, capping, and returning the bottle to be processed with other samples. At every site, the control water was filtered first into a Sterivex filter with the pump system. In the lab, additional DNA extraction blanks ("kit blanks") were processed alongside all samples to assure clean extractions, and additional negative PCR controls were processed alongside each PCR plate. DNA extraction blanks and PCR controls were each assessed for eDNA content using a QIAxcel Connect instrument after PCR amplification, and if any amplification was detected, the PCR was repeated before including this blank in the library pool. All field blanks were run through to the final sequencing step so that any PCR amplification could be assessed, see below.

Field blanks had either zero fish reads or very low fish richness, with the exception of 10 field blanks in the first year (2021) in which ASV richness was high, indicating a contamination event (Fig. S6). Among the 79 field blanks taken in 2021, 66% had 0

amplicons (i.e. clean field blanks), 22% had between 1-8 ASVs in the field blank, and 12% (10) had over 40 ASVs in the field blank (Fig. S6). The field blanks with >40 ASVs came from four regions across the network in BC, Oregon, and California and had similar community composition to each other, and with many associated samples that were also often high in ASV richness (Fig. S7). Although we could not pinpoint the exact source of contamination, the evidence was consistent through multiple sequencing preparation runs, indicating that it entered the pipeline upstream of sequencing, yet the richness was so high and compositionally similar across samples that it was unlikely to have entered at the field site itself.

To identify and remove samples that might be associated with this contamination event, we flagged all samples that belonged to the same cluster as those of the field blanks with >40 ASVs. These flagged samples were removed from all other display figures, and were not included in variance stabilized read estimation to remove their possible distorting effects (see below). We took this approach instead of read-subtraction because there were some samples that clustered with the putative contaminants without having high ASV richness in field controls, so full removal was deemed more conservative.

After removing flagged samples, samples remaining still contained some fish ASVs (1-9) that were also identified in the associated field blanks (Fig. S8). We did not perform read subtraction as a means of quality control from these remaining samples, because the fish ASVs common to field blanks and sampled belonged to species expected to occur in the region (occurred in other samples) yet field blanks often had higher read numbers than in samples, as can be expected when there are few competing DNA fragments for amplification. Read subtraction would therefore have removed common fish species (e.g., Pacific herring, surf perch) from samples for which we had independent evidence of occurrence in most regions.

Variance stabilized read estimation

To generate semiquantitative relative ‘abundance’ metrics that make use of relative read number while also accounting for batch effects, we used varianceStabilizingTransformation in DESeq2 (Love et al. 2017) with ‘year’ and ‘run’ as grouping variables as suggested by McMurdie and Holmes (2014). These provide a value potentially useful for comparisons within taxa similar to a read index (Kelly et al. 2019), but after removing batch effects that could affect variability in read number. We did not rarify samples to equal sampling depth or richness coverage in the dataset presented, in favour of retaining rare species when sampling depth was greater and to support analyses of community turnover which is less sensitive to sampling depth (McMurdie and Holmes, 2014).

Beach seining comparison

We conducted beach seining at 12 regions, allowing direct comparison between eDNA and beach seining results for nearshore fishes. The PECO project provided guidance on beach seining methods through a standard operating procedure (Sunday et al. 2022), similar to those described by Robinson and Yakimyshyn (Robinson and Yakimishyn 2013). In brief, beach seines were set in a round haul manner by boat or from shore adjacent to the seagrass meadow. The nets were pulled shoreward toward the edge of the meadow, during the window of low tide seagrass exposure around low tide. After the seine sets were completed, fish were held in buckets and totes until they were enumerated. Beach seine size (length, mesh size and drop in center), and replicate beach seine number varied across partners, due to either varying equipment available or legacy monitoring programs of each partner. The sampling effort, in terms of volume of water sampled, was therefore not uniform. Nevertheless, in all cases, fish were identified to species, counted, and returned rapidly to the water. Fish handling permits and specific details of methods of each partner are provided in the supplement (Table S1).

Most beach seining was conducted in 2021, although some partners conducted beach seines in subsequent years (two regions in 2022, two other regions in 2023). For exploration of overlaps between eDNA and beach seining results, we considered all within-region-year comparisons except for one region for which eDNA sample size became sparse after quality control in 2021 (West Coast Vancouver Island). To explore the coverage at the entire network scale, we also considered overlap between methods when all regions and years were pooled.

To assess the overlap of eDNA-derived data with beach-seining data, we associated each fish observation into a ‘set’ of observation type (seine only, seine and eDNA, eDNA only) and level of identification (species, genus, family). For ASVs identified to species in both the eDNA and beach seining, the overlapping set was straightforward. For ASVs identified only to genus or family in the eDNA, we associated these with fish species in the beach seining only if they were not already associated to eDNA at a lower taxonomic level. For example, when a seined species was not matched to an ASV identified to species but was matched to an ASV identified to genus, it was included in a new “seine-species to ASV-genus” set, and removed from the “seine only” category. We allowed associations between an observed species and an ASV in this way, even if only one ASV was observed but might represent multiple species (i.e., rockfish, *Sebastes* spp.), and denote these in our visualizations by connecting them.

For every fish species detected in beach seines, we estimated a site-level relative abundance as the total number of individuals of that species encountered across replicates within a site and year, divided by the total number of fish individuals encountered at that site and year.

Coverage of reference database

The region-specific cumulative taxonomic lists across 2021 to 2024 from eDNA were also compared to two databases of fish relevant to shallow-water coastal habitats of the Northeastern Pacific Ocean. The first was a custom database of fishes used for taxonomic assignment (see *Bioinformatics* above), and the second consisted of regionally-relevant fish from FishBase and OBIS (accessed May 3, 2026). To display overlap of eDNA-derived lists and these databases, we associated an observation type (eDNA, custom database, NCBI) with each fish observation and used the same database-species to ASV-taxon approach as taken in the comparison of eDNA and adjacent seine data.

We pulled species data from OBIS to generate a species list for the Northeastern Pacific Ocean. We first used shapefiles from the MEOW nearshore ecoregions and PPOW pelagic provinces, including the following set: Aleutian Islands, Gulf of Alaska, North American Pacific Fjordland, Puget Trough/Georgia Basin, Oregon, Washington, Vancouver Coast and Shelf, Northern California, Southern California Bight, California Current, Eastern Tropical Pacific. We then used these to pull OBIS records of fish using the R package, *rOBIS*. To remove species with spurious OBIS occurrences in these ecoregions, we further filtered this list to only species associated with the following FAO areas according to FishBase: Pacific, Northeast; Pacific, Eastern Central; America, North - Inland waters.

Effect of sampling effort at different hierarchical levels

We assessed how species richness increased as a function of sampling effort at three hierarchical levels. First, within samples, we assessed how species richness and sample coverage (using Chao2 operator, defined as 1 minus the probability that a new species would be encountered in the next sequence, Chao et al, 2020), increased with sequencing depth (number of sequences recovered for each sample). We used raw read number after quality control, so only including reads of ASVs associated with fish, and removed singleton reads within samples before rarefaction. We next assessed at the site level how species richness and sample coverage increases with number of water samples to the maximum of four samples within sites every year. Finally we assessed the relationship between species richness in a region and the number of sites visited. This allowed us to assess which aspects of our sampling effort, if any, were

incomplete and would benefit most from greater sampling effort. Calculations of coverage via Chao2 were carried out with the iNext package in R (Hsieh et al. 2016).

Culturally Sensitive Data

We developed a plan on how data would be made available through consultation with partners, including First Nations partners of our network on whose traditional territory data were collected. Data co-created with the Council of the Haida Nation (CHN) were provided for the sole purpose of this research. Geolocations of samples collected on Haida territory, including collections by the CHN and by the CHN and Parks Canada in Gwaii Haanas National Park Reserve, National Marine Conservation Area Reserve, and Haida Heritage Site, were obfuscated to 0.1 decimal degrees. Data from these regions require permission before further use (see Data Availability section).

Data Record

The final dataset has 1052 samples (distributed across 94 sites, 16 regions, and 4 years). Across all sites and years combined, there were 1145 unique amplicon sequence variants (ASVs) that were clustered into 265 unique molecular taxonomic units (MOTUs). Through taxonomic assignment, 982 unique ASVs (86%) were resolved to the level of species, 127 only to genus (11%), 32 only to family and 4 only to order. Similarly, 209 MOTUs (79%) were resolved to the level of species, 43 only to genus (16%), 10 to family and 3 to order.

Taxonomic assignment before vs. after clustering

Taxonomic assignments of MOTUs were sometimes at odds with taxonomic assignment of ASVs and we therefore provide both results in the data record. For some ASVs (n=5), clustering to MOTUs led to a more refined taxonomic assignment than direct assignment of the ASV on its own, because it was clustered into a group with a single closer match to the reference database. However, in other cases (n=21), at least one ASV was clustered into an MOTU with other ASVs that had different associations to the reference database and therefore it was brought down in taxonomic resolution. This occurred for example in the genus Syngnathus (pipefish), in which *Syngnathus exilis* and *Syngnathus californiensis* clustered together, so the resolution of the MOTU was to genus, reflecting the poor taxonomic resolution of this clade at the 12S locus. We recommend using MOTU-level assignment for most downstream biodiversity assessment data uses in which a conservative approach to taxonomic confidence is needed. However, when exploring overlap between eDNA and other records, we recommend using ASV-level taxonomies to maximize potential resolution.

Distribution of fish families

Across all years and sites, there were 63 distinct fish families recorded in total. These showed a distinct biogeographic pattern, with 14 families having southerly ranges, 24 families with more northerly ranges (mode at or north of latitude 50), and the rest (14) that span the entire study area (Fig. 3).

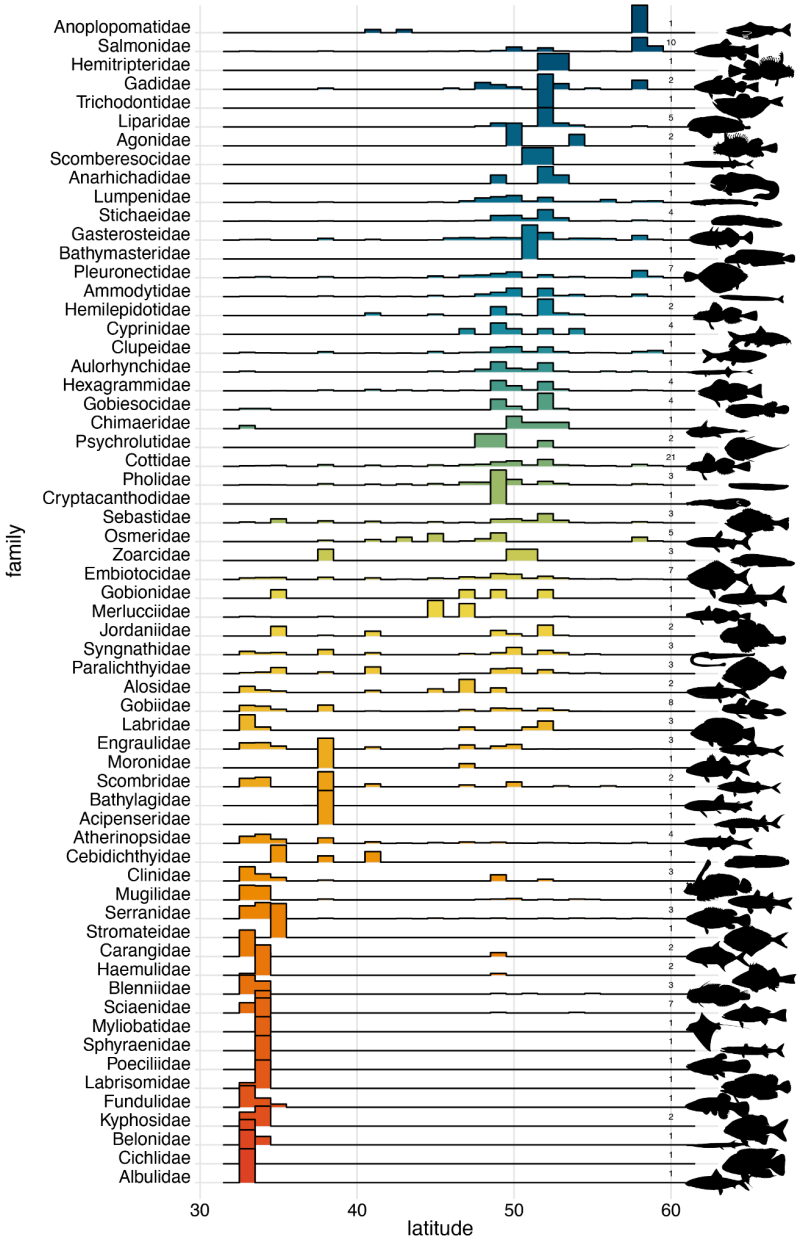


Figure 3. Fish families detected in the eDNA transect survey for all 4 years combined. Histogram shows relative number of samples in which ASVs assigned to family occurred, by binned latitude; number indicates total number of species per family. Most silhouettes retrieved from Phylopic, others created originally.

Technical validation

Comparison between eDNA results and adjacent beach seining provides some validation of recovered taxa. We found substantial overlap in taxa detected using each method (Fig. 4). A total of 81 fish species were detected across all 15 paired surveys using beach seining, and 50 of these (61%) were identified by eDNA at the species-level, 62 (12 more, or 77%) were detected by eDNA at the species or genus level, and 68 (84%) were detected by eDNA at the species, genus, or family level (Fig. 4a). Within regions, species in beach seines that were also detected by eDNA at species or genus-levels varied from 49% (Southeast Alaska region) to 100% (Juneau, Alaska region; Fig. 4b). There were 13 species in beach seines that were never detected at the species-to-family levels in the eDNA (dark grey set in Fig. 4a), and these typically had low site-level relative abundances in beach seines, with maximum site-level relative abundances that typically varied from 0.001% to 2.4% of the number of captured individuals within regions (Fig. 4a, *inset*), with one outlier being *Psychrolutes paradoxus* which was numerically dominant at 24% of beach seined fish at one site in Padilla Bay. This species had been detected in one of the eDNA samples that was deemed as contaminated before removal, and was also detected in eDNA samples in subsequent years. Many more taxa were detected by eDNA that were not detected in adjacent beach seines; we detected between 3 and 24 additional species within each region (41 total), as well as up to 5 additional genera (7 total; Fig. 4a).

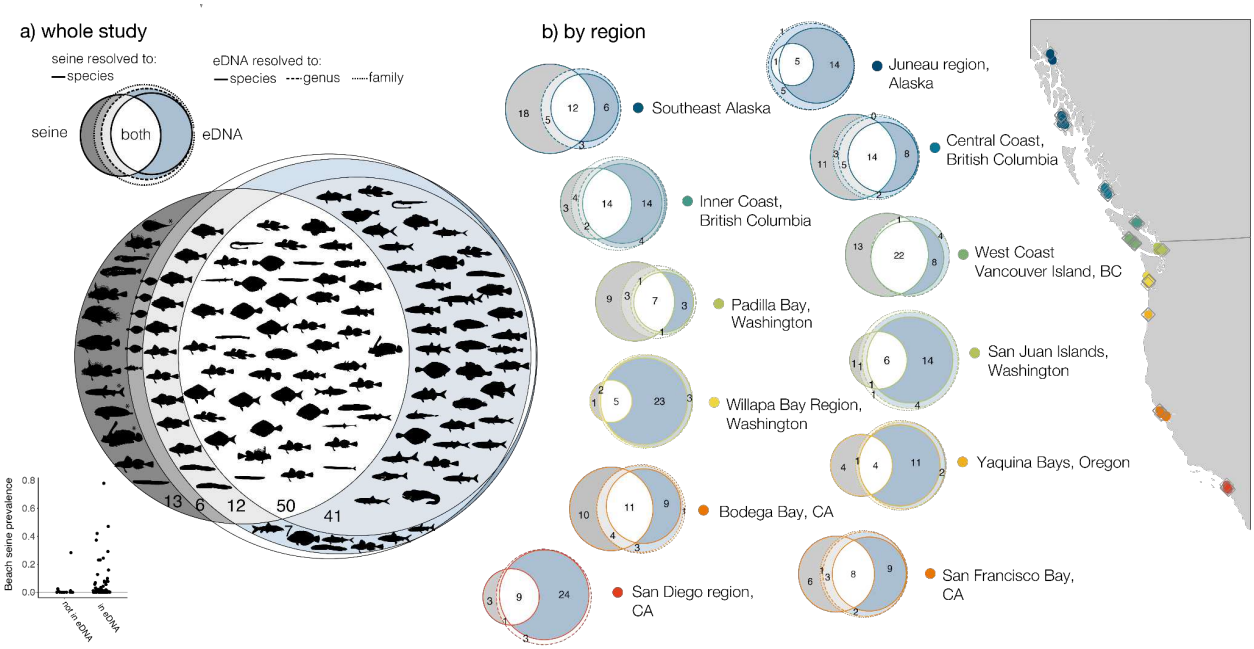


Figure 4. Comparison of fish diversity assessed using beach seines and eDNA metabarcoding across the PECO network. Map inset shows locations in which beach seining (diamonds) accompanied eDNA sampling (points). Data used in comparisons were from site-years in which more than 10 eDNA samples remained after quality control. Euler plots indicate species observed only in the beach seine (darkest grey),

only in eDNA (light blue) and in both methods (white) at the whole-study level (a), and within each sampled region (b). Dashed and dotted lines and light grey shading represent sets of additional taxa that could only be identified to genus (dashed) and family (dotted) levels in the eDNA. When overlap was found between ASVs identified to genus or family in the eDNA and species detected in the beach seines, the number of species detected in beach seining is represented. Of the 13 fish species in (a) only observed in beach seines, 5 were observed in eDNA in other years (shown with an asterisk). Inset in (a) shows maximum site-level relative abundance across beach seines in which each species was present, against the level at which it was detected in eDNA. Note, Padilla Bay and San Juan Island are displayed separately for the benefit of partners, but represent the same region.

Biodiversity coverage

Species richness of fish ranged from 1 to 22 species within samples, and appeared to be saturated at the sequencing read numbers that were achieved during metabarcoding, with Chao2 sample coverage of 1 (maximum) in all samples investigated (Fig. 5a,e; Fig. S9). At the site level, species richness ranged from 2 to 33, and rarefaction of samples (typically four per site) showed that richness did not always reach a plateau after four samples, with Chao2 coverage varying from x to y (Fig. 5b,e, S10). At the regional level within years, species richness varied from f to r, and rarefaction of sites (typically 5 per region) showed that richness was even further from a plateau with Chao2 sample coverage varying from z to w (Fig. 5c,e, S11). We also found that species richness continued to increase with increasing years of sampling (4 years total; Fig. 5d).

Comparison to regional databases

Overlap between a) the broad list of Northeastern Pacific fish (1425 species), b) the fish species detected at our sites with eDNA (136 identified to species), c) the fish detected at our sites beach seining (82 species), and d) those for which we had reference sequences at the 12S locus (12652 species after CRABs pipeline) gives a sense of biodiversity coverage possible in this region using our methods. While 75% of the broad list of Northeastern Pacific fish had 12S sequence in NCBI (Fig. 5f), all species detected by beach seining had sequences in the NCBI database (Fig. 5f). Four species detected in eDNA were not within the OBIS occurrence records: three have native ranges in Europe and were not expected to be detected (*Tautogolabrus adspersus, Gobio gobio, Squalius cephalus*), and one was a non-indigenous freshwater species known to live in the region (*Rhinichthys cataractae*). In addition to these, we also identified in the eDNA five additional non-native fishes that were included in the broad NE Pacific list but listed or known to be non-native species living in the region: *Tridentiger trigonocephalus,*

Alosa sapidissima, Salmo salar, Morone saxatilis, Gambusia affinis. We discuss these potential non-native detections below.

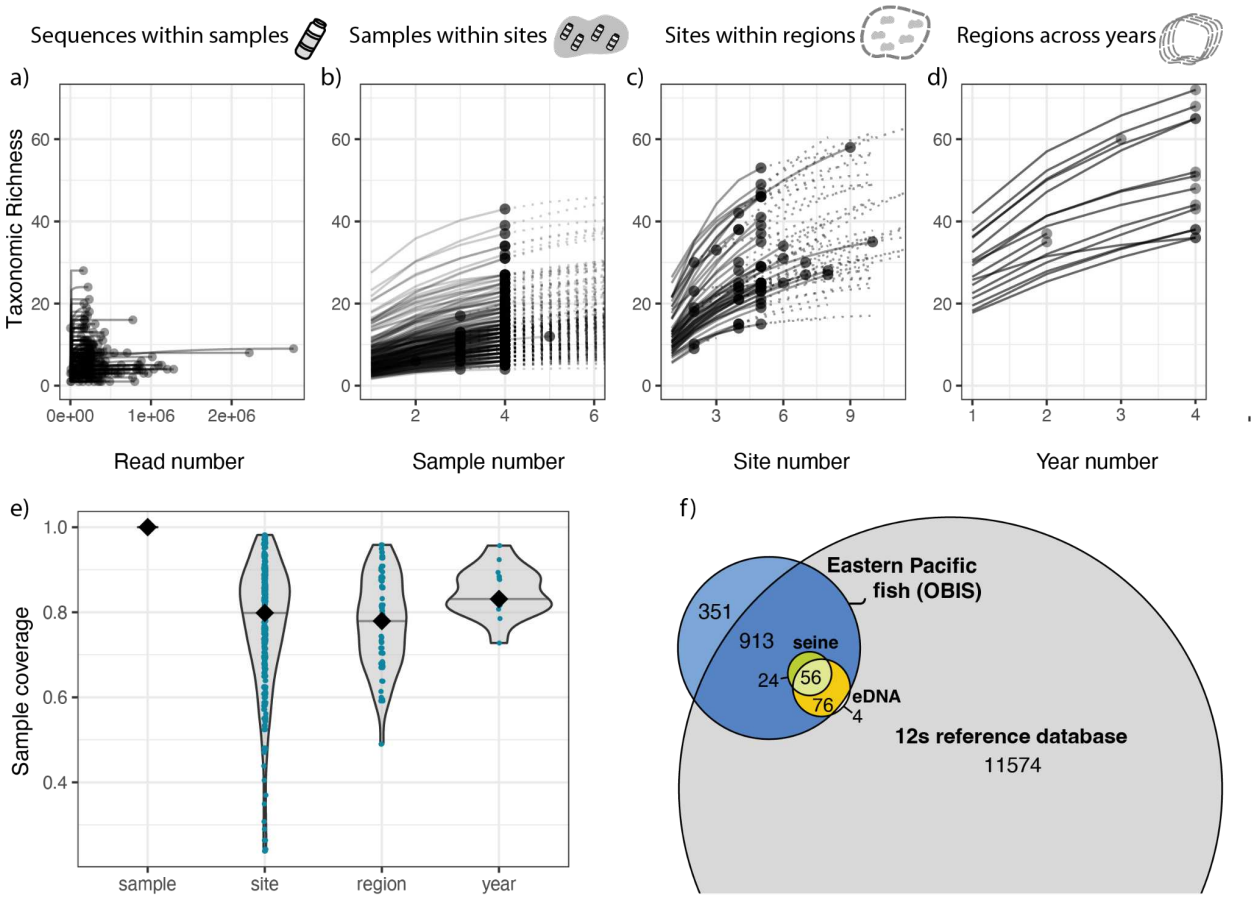


Figure 5. Richness and coverage of fish species detected in the eDNA metabarcoding network at different levels of sampling effort. (a) Sample-level taxonomic richness rarefied as a function of sequencing depth within samples, shown for a random set of 20% of total samples in study (n=210). (b) Site-level taxonomic richness rarified as a function of sample number within sites and years (typically 4 replicates per site, range 2-6). (c) Region-level taxonomic richness rarefied as a function of number of sites sampled within regions and years (typically 5 sites per region, range 2-10). (d) Region-level richness observed within regions across all years, rarefied as a function of years visited at each region (typically 4 years at each region). Taxonomic richness represents the number of uniquely-identified MOTUs, to represent species-level units even if not all identified at the species level. Solid lines represent mean rarefied richness across iterative draws, points show observed total richness at each sampling level, and dashed lines show extrapolated estimates based on Chao2 procedures. (e) Estimated Chao2 sample coverage at each level of sampling; coverage ranges between 0 and 1 and indicates the probability that the next sample will contain previously-sampled taxa. Grey

polygons show density distribution of data; diamond and horizontal lines show median. (f) Overlap between fish species detected in eDNA, fish species detected in beach seines, species with known occurrences in the North Eastern Pacific regions of OBIS, and fish species in the 12S reference dataset. Six species in the eDNA were not within the OBIS occurrence records, see text for details. Beach seine and eDNA species represent pooled data from all PECO samples across years.

Discussion

The Pacific eDNA Coastal Observatory (PECO) dataset represents a spatially balanced and quality-controlled four-year dataset of fish biodiversity data across 26 degrees of latitude along the west coast of North America. The level of matching to beach seining detections within regions and years contextualizes the quality of the data record; all abundant and most rare species in seines were detected by eDNA, as well as many other species known to live in the region. The distinct biogeographic pattern at the family level provides confidence in biogeographical-scale information and suggests family-level fidelity to major thermal regimes (Stuart-Smith et al. 2015). PECO thus provides a new resource for future assessments of biodiversity change and contributes to the advancement of eDNA technology for ocean biodiversity monitoring as envisioned for the UN Ocean Decade (Canonico et al. 2019, Leinen et al. 2022), joining other large-scale networks in this goal (Santi et al. 2023, Sanchez et al. 2025, Sioud et al. 2026). Here we discuss sensitivity of the survey, the hierarchical and networked approach to achieving scale, and the value of the network to overcome both method and data interoperability as technologies develop.

Sensitivity of eDNA surveys of fish biodiversity

The quality of an eDNA biodiversity assessment hinges on its ability to detect species when they are present (sensitivity) while reducing erroneous detections as false positives (specificity). Our validation steps assess the sensitivity of the PECO network, by addressing three main sources of missing species: a possible lack of coverage in the reference database, insufficient sequencing depth to detect fragments when present in a sample, and insufficient environmental sampling to capture DNA fragments in the site or region. Our assessments of coverage suggest that the PECO eDNA dataset is unlikely to miss eelgrass-associated fish due to lack of coverage in the reference database, as all fish detected in the beach seines as well as most shallow-water marine species were in the 12S reference database. In addition, our rarefaction curves suggest that sequencing was sufficiently deep to capture fish species (coverage of 1 in all samples investigated). Instead, our results show that sample coverage of both sites and regions is where we likely miss some species; median sample coverage 0.8 suggests that more sampling at either level would reveal more fish species (20% chance of a new species in the next sample or site; Chao et al. 2020).

While our comparison to beach seining showed substantial overlap, there were thirteen species (16%) found in beach seines that were never detected at or below the family level in the eDNA. However, these were generally lower-biomass species, supporting previous work that eDNA has good overlap with capture survey data in the NE Pacific, and typically misses a small percentage of fish that tend to have lower biomass (He et al. 2022, Millard-Martin et al. 2024, Rehill et al. 2024, Domke et al. 2025). For species only detected by the eDNA at the genus and family levels, we cannot be certain of their match and these include the rockfishes (genus *Sebastes*), which are culturally and economically important across this region (Acharya-Patel et al. 2024). These near-matches will need to be ameliorated to make use of species-level information as functional trait and distribution data. Yet because they are distinct operational taxonomic units, they still contribute to biodiversity accounting. Future work could better resolve the taxonomic level of these identifications, e.g., using *miSebastes* (Min et al. 2021, Acharya-Patel et al. 2024).

Our annotation method allowed non-native species to be retained in the final dataset if we had independent evidence for their occurrence in the region at large. Hence our final dataset has 8 non-native species for which we had independent evidence that they could occur in the area. Although we retain these in the dataset, they should be verified with direct organismal observations. The occurrences of *Salmo solar* (Atlantic salmon) we caution likely does not indicate the presence of *S. solar* in these regions, but rather a long history of open-net farms on the coast which has released a high abundance of genetic material into the surrounding seawater and sediment.

Our discovery of a contamination event was made possible by processing field blanks throughout our pipeline. Although in most years field blanks contained genetic evidence of only a few ubiquitous fish species that were of minor concern, the field blanks were essential for identifying when a high number of fish species were found in a small number of field blanks and samples, leading us to remove those samples from the analyses. The “extra” fish species detected in these samples were not unique to the survey as a whole but were often unique to each sampled region and therefore failing to remove them would have led to spurious biodiversity patterns. It seems unlikely that these field blanks were affected by technical flaws in the field, as we see no clear mechanism by which diverse fish from other regions could appear within a single field site. Methods evidently improved after the first year, possibly connected to reducing the number of times extracted DNA was accessed as methods evolved.

Hierarchical networked sampling facilitates biogeographic-scale observatory

PECO provides comparable results about fishes across 26° of latitude that span several known biogeographic breaks (Tosetto et al. 2025). Given four years of biodiversity sampling, the dataset is on its way to becoming a *datacube*, tracking individual species over space and time (Groom et al. 2025).

The PECO dataset can be interpreted at different hierarchical scales. Samples were collected within sites, sites nested within regions, and regions were repeatedly sampled across years. Our hierarchical design allows repeat surveys within a site that can be useful in detectability modelling, while also providing information across sites within regions. The saturated species rarefaction curves with sequencing depth suggests that adequate sequencing was conducted within every sample, and that patterns of species occupancy across samples represent true heterogeneity in eDNA captured by 1L samples. Our finding of slightly greater coverage at the site level compared to the region level suggests that more effort allocated to new sites would likely be more efficient than increasing samples within sites, in line with previous work on allocation of effort in trend detection (Andersen and Steidl 2020).

Networks and shared protocols enable comparable biodiversity assessments
Sharing methods allowed immediate comparability and immediate assessment of biogeographical patterns. Had all 18 partners conducted their own eDNA surveys, methods likely would have been variable and therefore difficult to integrate (Takahashi et al. 2023). Instead, sharing technology and local knowledge across our partnership helped to achieve a spatially comprehensive design.

The Pacific eDNA Coastal Observatory is among the first large collaborative eDNA networks in North America. Since a key goal of biodiversity monitoring is to track change through time, we plan to keep this monitoring program going throughout the UN Ocean Decade, with an immediate goal of reaching a 10 year data series. Through species annotations, this dataset is ready to be connected with species traits, foodwebs, and associated environmental/human drivers. For instance, community composition may follow latitudinal gradients or shift primarily at biogeographic boundaries, or may vary with regional habitat heterogeneity around each seagrass habitat (Proudfoot et al. 2023). Regionally-warm areas or interannual thermal variation, including marine heatwaves, could lead to short-term spatiotemporal variation in fish communities (Robinson et al. 2022). As samples are cryopreserved, new insights can also be gained in the future by reanalyzing samples, such as using miSebastes to improve resolution in rockfish, or targeted PCR to track species of economic or social interest. The repeated design of this network can further allow forecasting of change, which is then tested with subsequent sampling.

We have made this dataset public and followed developing metadata standards to make the data findable, accessible, interoperable, and reproducible, using international metadata standards (FAIR eDNA, Takahashi et al. 2025), making the data machine-readable and interoperable with other datasets. Given the relatively fast uptake of eDNA sampling methods globally, including other monitoring programs that have taken place within the study domain of our network, integrating data from diverse datasets will be an important next goal and test of eDNA data standards. We acknowledge that these biodiversity data have been collected on Indigenous lands, and especially for data co-acquired with First Nations partners within our network, we have sought informed permission before data use, and set up future use of these data to follow the same steps (CARE principles; Carroll et al. 2021). PECO's networked approach to biodiversity sampling opens up opportunities to explore spatiotemporal variation at nested scales in a taxonomic group of highly-mobile organisms that are difficult to sample underwater. The data overview here shows the value of a network for addressing questions of biodiversity change.

Data availability

Raw DNA sequences generated in this study will be deposited at NCBI upon acceptance (https://www.ncbi.nlm.nih.gov/bioproject/PRJNA1275805/) to be publicly available as of the date of publication.

Some of the raw DNA sequences are derived from samples collected on Haida territory by the Council of the Haida Nation (CHN). These samples were provided for the sole purpose of this research, and require permission before further use. All research conducted on, or in relation to, Haida territory is subject to a Council of the Haida Nation Research Permit, with the exception of research carried out in partnership with the CHN or Gwaii Haanas Parks Canada.

The ASV-and-taxon-by-sites table, along with site- and sample-level metadata and a data dictionary has been deposited Zenodo. This will be made public with publication of the paper, and we provide a private link for review.

Code Availability

The bioinformatics pipeline used to process data from raw files to derived data has been deposited in Zenodo. This will be made public with publication of the paper, and we provide a private read-only link for review. We also include a cleaned version of our sequence data to run through post-cleaning scripts. We provide a a private link for review.

Local BLAST database and CRABS pipeline have been deposited together at Zenodo and is publicly available at https://zenodo.org/records/20384787.

The code needed to reconstruct figures in this manuscript has been deposited in Zenodo. This will be made public with publication of the paper, and we provide a private read-only link for review.

Authorship contribution statement

J.M.S., M.H-L., M.A.L. S.V. and K.S. designed the research; A.B.C., A.K., A.M.O., B.B.H., D.B., D.H., F.T., G.L.E., H.B., J.B.S., J.H., J.J.S., J.L.R., J.M.S., J.Y., K.A.H., K.E.B., K.L.S., L.C.L., L.K.D, L.N., L.R.A., M.H-L., M.O., M.V.B., O.J.G., R.S.M., S.C., S.J.B., S.S.G., and Z.M. led regional field collections and conducted data quality control; O.J.G., J.L.R., L.K.D, L.R.A., A.B.C., A.M.O., J.Y., K.L.S., L.C.L., M.V.B., S.S.G., and Z.M. led beach seining collections within regions; R.B.G.C-C and K.M.R. conducted laboratory processing and quality control, S.V. prepared and administered shipment of sampling kits; K.S., J.M.S., and S.B. conducted data analysis; M.H-L., L.C.L., S.J.B., J.Y., K.E.B., E.M., M.A.L. and J.M.S. conducted administrative, technical, and supervisory support; K.S., J.M.S., S.B., and E.M. developed software; K.S. developed bioinformatics pipeline, J.M.S. and S.B. prepared figures, A.B.C., A.K., A.M.O., D.H., H.B., J.B.S., J.H., J.J.S., J.L.R, J.Y., K.L.S., K.S., L.C.L., L.K.D, L.N., L.R.A., M.A.L., M.H-L., M.V.B., O.J.G., R.B.G.C-C, R.S.M., S.B., S.J.B., S.S.G., Z.M. provided feedback in group meetings; J.M.S. and K.S. wrote the first draft; A.K., A.M.O., D.H., E.M., F.T., H.B., J.B.S., J.L.R, J.Y., K.E.B., K.L.S., K.M.R., K.S., L.C.L., L.K.D, L.R.A., M.A.L., M.H-L., M.V.B., O.J.G., R.B.G.C-C, R.S.M., S.B., S.J.B., S.S.G., S.V., Z.M. provided comments/feedback/ideas on manuscript draft.

Acknowledgements

This work would not be possible without the generous support of the Tula Foundation and the Hakai Institute through which all data was collaboratively generated. We acknowledge all Hakai project staff, including Silvio Carvalho, Carly Janusson, and Rosie Savage, for supporting the technical aspects of this work, as well as to Ben Millard-Martin for his role in developing the sampling kits. We are also grateful to many staff and volunteers from each region who supported field sampling across sites and years, including Ahmad Mersaghian, Sadie Kenny, Kirra Engbrecht, Mia Vafadari, Andrea Klishevich, Rebeka Shamis, Caroline Thorpe, Amalia Tamone, Regan Hickling, Charlotte Houston, Chavonne Guthrie, Fiona Francis, Gracelyn Ham, Jack Moseley, Jaydn Paquette, Carolyn Prentice, Jessica Whitney, Karl Koehler, Krista Bohlen, Krystal Bachen, Lianna Gendall, Michael McDermid, Rachael Karm, Julieta Gomez, Vini Souza, Joe Jackson, Abbey Dias, Megan Sixt, Maria Velazquez, Preslee Chase, Preston Sloan, Rumer Opie, Fiona Boardman, Mary K. English. We are grateful for financial

support from the National Science and Engineering Research Council of Canada and the Alfred P. Sloan Foundation; field work in the SE Alaska region was supported by the National Science Foundation (grant no. 1600230), and National Science Foundation Graduate Research Fellowships awarded to L.K.D.; fieldwork in the Willapa Bay region was supported by a sub-award to Pacific Shellfish Institute grant #20-31G with the Pacific States Marine Fisheries Commission, under award #NA18NMF4720007.

Disclaimer

The scientific results and conclusions, as well as any views or opinions expressed herein, are those of the author(s) and do not necessarily reflect the views of the National Marine Fisheries Service, the National Oceanic and Atmospheric Administration, or the Department of Commerce. Reference to trade names does not imply endorsement by the National Marine Fisheries Service, the National Oceanic and Atmospheric Administration, or the Department of Commerce.

References

Acharya-Patel N, Groenwold E, Lemay MA, Clemente-Carvalho R, Morien E, Dudas S, Rubidge E, Yang CL, Coombe L, Warren RL, Frid A, Birol I, Helbing CC (2024) Comparison of environmental DNA and SCUBA diving methods to survey keystone rockfish species on the Central Coast of British Columbia, Canada. Ecological Indicators 160: 111830. https://doi.org/10.1016/j.ecolind.2024.111830

Andersen EM, Steidl RJ (2020) Power to detect trends in abundance within a distance sampling framework. Michalski F (Ed.). Journal of Applied Ecology 57: 344–353. https://doi.org/10.1111/1365-2664.13529

Bates AE, Bird TJ, Stuart-Smith RD, Wernberg T, Sunday JM, Barrett NS, Edgar GJ, Frusher S, Hobday AJ, Pecl GT, Smale DA, Mccarthy M (2015) Distinguishing geographical range shifts from artefacts of detectability and sampling effort. Diversity and Distributions 21. https://doi.org/10.1111/ddi.12263

Beentjes KK, Speksnijder AGCL, Schilthuizen M, Hoogeveen M, Van Der Hoorn BB (2019) The effects of spatial and temporal replicate sampling on eDNA metabarcoding. PeerJ 7: e7335. https://doi.org/10.7717/peerj.7335

Boettiger C, Lang DT, Wainwright P (2012) rfishbase: exploring, manipulating and visualizing FishBase data from R. Journal of Fish Biology. Available from: https://doi.org/10.1111/j.1095-8649.2012.03464.x.

Bourlat SJ, Haenel Q, Finnman J, Leray M (2016) Preparation of Amplicon Libraries for Metabarcoding of Marine Eukaryotes Using Illumina MiSeq: The Dual-PCR Method. In: Bourlat SJ (Ed.), Marine Genomics. Methods in Molecular Biology. Springer New York, New York, NY, 197–207. https://doi.org/10.1007/978-1-4939-3774-5_13

Callahan BJ, McMurdie PJ, Rosen MJ, Han AW, Johnson AJ, Holmes SP (2016) DADA2: High-resolution sample inference from Illumina amplicon data. Nat Methods 13: 581–3. https://doi.org/10.1038/nmeth.3869

Canonico G, Buttigieg PL, Montes E, Muller-Karger FE, Stepien C, Wright D, Benson A, Helmuth B, Costello M, Sousa-Pinto I, Saeedi H, Newton J, Appeltans W, Bednaršek N, Bodrossy L, Best BD, Brandt A, Goodwin KD, Iken K, Marques AC, Miloslavich P, Ostrowski M, Turner W, Achterberg EP, Barry T, Defeo O, Bigatti G, Henry L-A, Ramiro-Sánchez B, Durán P, Morato T, Roberts JM, García-Alegre A, Cuadrado MS, Murton B (2019) Global Observational Needs and Resources for Marine Biodiversity. Frontiers in Marine Science 6: 367. https://doi.org/10.3389/fmars.2019.00367

Carroll SR, Herczog E, Hudson M, Russell K, Stall S (2021) Operationalizing the CARE and FAIR Principles for Indigenous data futures. Scientific Data 8: 108. https://doi.org/10.1038/s41597-021-00892-0

Chamberlain S, Vanhoorne B (2024) worrms: World Register of Marine Species (WoRMS) Client. Available from: https://docs.ropensci.org/worrms/.

Chamberlain S, Szoecs E, Foster Z, Arendsee Z, Boettiger C, Ram K, Bartomeus I, Baumgartner J, O'Donnell J, Oksanen J, Tzovaras BG, Marchand P, Tran V, Salmon M, Li G, Grenié M (2020) taxize: Taxonomic information from around the web. Available from: https://github.com/ropensci/taxize.

Chao A, Kubota Y, Zelený D, Chiu C, Li C, Kusumoto B, Yasuhara M, Thorn S, Wei C, Costello MJ, Colwell RK (2020) Quantifying sample completeness and comparing diversities among assemblages. Ecological Research 35: 292–314. https://doi.org/10.1111/1440-1703.12102

Domke LK, Ledger KJ, Whitney JM, Kachel SM, Larson W, Eckert GL, Baetscher DS (2025) A case for multi-gear assessments: Detection probabilities of nearshore fish with eDNA and seine nets vary by functional traits. https://doi.org/10.1101/2025.11.28.691209

Duffy JE, Amaral-Zettler LA, Fautin DG, Paulay G, Rynearson TA, Sosik HM, Stachowicz JJ (2013) Envisioning a Marine Biodiversity Observation Network. BioScience 63: 350–361. https://doi.org/10.1525/bio.2013.63.5.8

Frøslev TG, Kjøller R, Bruun HH, Ejrnæs R, Brunbjerg AK, Pietroni C, Hansen AJ (2017) Algorithm for post-clustering curation of DNA amplicon data yields reliable biodiversity estimates. Nature Communications 8. https://doi.org/10.1038/s41467-017-01312-x

Gotelli NJ, Colwell RK (2001) Quantifying biodiversity: procedures and pitfalls in the measurement and comparison of species richness. Ecology Letters 4: 379–391. https://doi.org/10.1046/j.1461-0248.2001.00230.x

Grey EK, Bernatchez L, Cassey P, Deiner K, Deveney M, Howland KL, Lacoursière-Roussel A, Leong SCY, Li Y, Olds B, Pfrender ME, Prowse TAA, Renshaw MA, Lodge DM (2018) Effects of sampling effort on biodiversity patterns estimated from environmental DNA metabarcoding surveys. Scientific Reports 8: 8843. https://doi.org/10.1038/s41598-018-27048-2

Groom Q, Abraham L, Adriaens T, Breugelmans L, Clarke DA, Di Musciano M, Dove S, Estupinan-Suarez LM, Faulkner KT, Fernandez M, Hendrickx LA, Hui C, Joly A, Kumschick S, Langeraert W, Martini M, Miller J, Oldoni D, Pereira H, Preda C, Robertson T, Rocchini D, Rodrigues A, Ryckewaert M, Seebens H, Sica YV, Teixeira H, Trekels M, Vandaele T, Wilson JRU, Zengeya T, Desmet P (2025) Creating the vision of rapid, repeatable, reactive data workflows for policy on biodiversity. Ecological Solutions and Evidence 6: e70113. https://doi.org/10.1002/2688-8319.70113

He X, Stanley RRE, Rubidge EM, Jeffery NW, Hamilton LC, Westfall KM, Gilmore SR, Roux L-MD, Gale KSP, Heaslip SG, Steeves R, Abbott CL (2022) Fish community surveys in eelgrass beds using both eDNA metabarcoding and seining: implications for biodiversity monitoring in the coastal zone. Canadian Journal of Fisheries and Aquatic Sciences 79: 1335–1346. https://doi.org/10.1139/cjfas-2021-0215

Hsieh TC, Ma KH, Chao A (2016) iNEXT: an R package for rarefaction and extrapolation of species diversity (H ill numbers). McInerny G (Ed.). Methods in Ecology and Evolution 7: 1451–1456. https://doi.org/10.1111/2041-210X.12613

Jennifer M. Sunday, Matthew A. Lemay, Margot Hessing-Lewis, Carolyn Prentice, Kate Sheridan, Scott Gabara, Nora Bartram-Forbes (2022) PECO Beach Seining Protocol. Available from: osf.io/z4fdu.

Kellogg C, Carvalho RuteC, Lemay M (2024) DNA Extraction from 0.22µm Sterivex Filters - Qiagen Blood and Tissue Kit v1. https://doi.org/10.17504/protocols.io.bp2l6221kgqe/v1

Kelly RP, Shelton AO, Gallego R (2019) Understanding PCR Processes to Draw Meaningful Conclusions from Environmental DNA Studies. Scientific Reports. https://doi.org/10.1038/s41598-019-48546-x

Leinen M, Chavez F, Meyer R, Buttigieg PL, Davies N, Casotti R, Fischer A (2022) The Ocean Biomolecular Observing Network (OBON). Marine Technology Society Journal 56: 106–107. https://doi.org/10.4031/MTSJ.56.3.20

Love M, Huber W, Anders S (2017) DESeq2. https://doi.org/10.18129/B9.BIOC.DESEQ2

Magurran, McGill (2011) Biological Diversity: Frontiers in Measurement and Assessment. Oxford University Press.

Mahé F, Rognes T, Quince C, De Vargas C, Dunthorn M (2015) Swarm v2: highly-scalable and high-resolution amplicon clustering. PeerJ 3: e1420. https://doi.org/10.7717/peerj.1420

Martin M (2011) Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet.journal; Vol 17, No 1: Next Generation Sequencing Data AnalysisDO - 10.14806/ej.17.1.200. Available from: https://journal.embnet.org/index.php/embnetjournal/article/view/200/479.

McMurdie PJ, Holmes S (2014) Waste Not, Want Not: Why Rarefying Microbiome Data Is Inadmissible. PLOS Computational Biology 10: e1003531. https://doi.org/10.1371/journal.pcbi.1003531

Millard-Martin B, Sheridan K, Morien E, Lemay MA, Hessing-Lewis M, Clemente-Carvalho RBG, Sunday JM (2024) Effect of environmental DNA sampling resolution in detecting nearshore fish biodiversity compared to capture surveys. PeerJ 12: e17967. https://doi.org/10.7717/peerj.17967

Min MA, Barber PH, Gold Z (2021) MiSebastes: An eDNA metabarcoding primer set for rockfishes (genus Sebastes). Conservation Genetics Resources 13: 447–456. https://doi.org/10.1007/s12686-021-01219-2

OBIS (2025) Ocean Biodiversity Information System. Intergovernmental Oceanographic Commission of UNESCO. Available from: https://obis.org.

Pecl GT, Araújo MB, Bell JD, Blanchard J, Bonebrake TC, Chen IC, Clark TD, Colwell RK, Danielsen F, Evengård B, Falconi L, Ferrier S, Frusher S, Garcia RA, Griffis RB, Hobday AJ, Janion-Scheepers C, Jarzyna MA, Jennings S, Lenoir J, Linnetved HI, Martin VY, McCormack PC, McDonald J, Mitchell NJ, Mustonen T, Pandolfi JM, Pettorelli N, Popova E, Robinson SA, Scheffers BR, Shaw JD, Sorte CJB, Strugnell JM, Sunday JM, Tuanmu MN, Vergés A, Villanueva C, Wernberg T, Wapstra E, Williams SE (2017) Biodiversity redistribution under climate change: Impacts on ecosystems and human well-being. Science 355. https://doi.org/10.1126/science.aai9214

Proudfoot B, Thompson PL, Bohlen K, Collyer M, Yakimishyn J, Robinson CLK (2023) Eelgrass Meadow Edge Habitat Heterogeneity Enhances Fish Diversity on the Pacific Coast of Canada. Estuaries and Coasts 46: 1326–1344. https://doi.org/10.1007/s12237-023-01203-z

R Core Team (2023) R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. Available from: https://www.R-project.org/.

Rehill T, Millard-Martin B, Lemay M, Sheridan K, Mueller A, Morien E, Clemente-Carvalho RBG, Hunt BPV, Sunday JM (2024) Detection differences between eDNA and mid-water trawls are driven by fish biomass and habitat preferences. Environmental DNA 6: e586. https://doi.org/10.1002/edn3.586

Robinson CLK, Yakimishyn J (2013) The persistence and stability of fish assemblages within eelgrass meadows (*Zostera marina*) on the Pacific coast of Canada. Rochet M-J (Ed.). Canadian Journal of Fisheries and Aquatic Sciences 70: 775–784. https://doi.org/10.1139/cjfas-2012-0339

Robinson CLK, Yakimishyn J, Evans R (2022) Minimal effects of the 2014-16 marine heatwave on fish assemblages found in eelgrass meadows on the southwestern coast of Vancouver Island, British Columbia, Canada. Frontiers in Marine Science 9: 980703. https://doi.org/10.3389/fmars.2022.980703

Sanchez L, Loiseau N, Albouy C, Bruno M, Barroil A, Dalongeville A, Deter J, Durand J-D, Faure N, Fopp F, Hocdé R, Jaquier M, Jiddawi NS, Jucker M, Juhel J-B, Kadarusman, Marques V, Mathon L, Mouillot D, Orblin M, Pellissier L, Seguin R, Sugeha HY, Valentini A, Velez L, Vimono IB, Leprieur F, Manel S (2025) eDNA surveys substantially expand known geographic and ecological niche boundaries of marine fishes. Tanentzap AJ (Ed.). PLOS Biology 23: e3003432. https://doi.org/10.1371/journal.pbio.3003432

Santi I, Beluche O, Beraud M, Buttigieg PL, Casotti R, Cox CJ, Cunliffe M, Davies N, De Cerio OD, Exter K, Kervella AE, Kotoulas G, Lagaisse R, Laroquette A, Louro B, Not F, Obst M, Pavloudi C, Poulain J, Præbel K, Vanaverbeke J, Pade N (2023) European marine omics biodiversity observation network: a strategic outline for the implementation of omics approaches in ocean observation. Frontiers in Marine Science 10: 1118120. https://doi.org/10.3389/fmars.2023.1118120

Sato Y, Miya M, Fukunaga T, Sado T, Iwasaki W (2018) MitoFish and MiFish Pipeline: A Mitochondrial Genome Database of Fish with an Analysis Pipeline for Environmental DNA Metabarcoding. Mol Biol Evol 35: 1553–1555. https://doi.org/10.1093/molbev/msy074

Shelton AO, O'Donnell JL, Samhouri JF, Lowell N, Williams GD, Kelly RP (2016) A framework for inferring biological communities from environmental DNA. Ecological Applications 26: 1645–1659. https://doi.org/10.1890/15-1733.1

Sioud I, Minegishi Y, Kajita T, Isowa Y, Tanabe AS, Yamakawa H, Berry O, Basyuni M, Cheah HL, Deagle B, Fall J, Habib KA, Kumbregama S, Klangnurak W, Leopardas VE, Mbaye NC, Peer N, Sivakumar K, Sivaipram I, Von Ammon U, Wee A, Woo SP, Kondoh M (2026) ANEMONE Global's First eDNA -Based Assessment: Insights Toward a Standardized Global Monitoring Network. Environmental DNA 8: e70252. https://doi.org/10.1002/edn3.70252

Stammnitz MR, Hartman Scholz A, Duffy DJ (2024) Environmental DNA without borders: Let's embrace decentralised genomics to meet the UN's biodiversity targets. EMBO Reports 25: 4095–4099. https://doi.org/10.1038/s44319-024-00264-w

Stuart-Smith RD, Edgar GJ, Barrett NS, Kininmonth SJ, Bates AE (2015) Thermal biases and vulnerability to warming in the world's marine fauna. Nature 528: 88–92. https://doi.org/10.1038/nature16144

Sunday J, Lemay M, Velazquez S, Hessing-Lewis M (2024) PECO eDNA field sampling protocol v2. https://doi.org/10.17504/protocols.io.ewov19d5olr2/v2

Takahashi M, Saccò M, Kestel JH, Nester G, Campbell MA, van der Heyde M, Heydenrych MJ, Juszkiewicz DJ, Nevill P, Dawkins KL, Bessey C, Fernandes K, Miller H, Power M, Mousavi-Derazmahalleh M, Newton JP, White NE, Richards ZT, Allentoft ME (2023) Aquatic environmental DNA: A review of the macro-organismal biomonitoring revolution. Science of The Total Environment 873: 162322. https://doi.org/10.1016/j.scitotenv.2023.162322

Thomsen PF, Jensen MR, Sigsgaard EE (2024) A vision for global eDNA-based monitoring in a changing world. Cell 187: 4444–4448. https://doi.org/10.1016/j.cell.2024.04.019

Unsworth RKF, Nordlund LM, Cullen-Unsworth LC (2019) Seagrass meadows support global fisheries production. Conservation Letters 12: e12566. https://doi.org/10.1111/conl.12566

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SupplementalmaterialPECOdatapaper.pdf