Title: Water volume, biological and PCR replicates influence the characterization of deep-sea pelagic fish communities.

Running title: Sampling eDNA in the deep sea.

Authors:

Pedro A. Peres[1], Heather Bracken-Grissom[1,2]

1 Institute of Environment, Department of Biological Sciences, Florida International University, 3000 NE 151st St, North Miami, FL, 33181 USA.

2 Department of Invertebrate Zoology, Smithsonian Institution, National Museum of Natural History, Washington, WA, 20013-7012 USA.

Contact information: Pedro A. Peres: pedro.peres27@gmail.com ; Heather Bracken-Grissom: hbracken@fiu.edu

Author contributions: PAP and HBG – conception and design of the study; PAP and HBG – data collection; PAP – formal analyses; HBG – supervision; PAP and HBG – funding acquisition; PAP – writing original draft; PAP and HBG – writing review and editing.

Data availability statement: All data used is available in Supplemental Materials.

Funding statement: This material is based upon work supported by the National Science Foundation under Grant No. HRD-1547798 and Grant No. HRD-2111661. These NSF Grants were awarded to Florida International University as part of the Centers of Research Excellence in Science and Technology (CREST) Program. This research was also made possible by a grant from the Office of Naval Research (#98288/1184488/2), the Gulf of Mexico Research Initiative (DEEPEND|RESTORE consortium), and the National Oceanic and Atmospheric Administration's RESTORE Science Program under award NA19NOS4510193 to Florida International University.

Conflict of interest disclosure: Authors declare no conflict of interest.

Ethics statement: Authors confirm that the manuscript has not been submitted elsewhere, and that all research meets the ethical guidelines of the study country.

Acknowledgments: We thank all researchers involved in the DEEPEND|RESTORE consortium for their support in sample collections. This is contribution #X from the Coastlines and Oceans Division of the Institute of Environment at Florida International University.

ABSTRACT

The pelagic deep sea is challenging to investigate due to logistical constraints regarding access and collection of samples, however environmental DNA (eDNA) can potentially revolutionize our understanding of this ecosystem. Although advancements are being made regarding technology and eDNA samplers, many countries and research groups still do not have access to modern equipment. At the same time, no investigation has been performed to assess the impact of different experimental designs using common gear found on research vessels (i.e., CTD mounted with Niskin bottles). Here, we investigated the effects of sampled water volume, biological and PCR replicates in characterizing deep-sea pelagic biodiversity. Samples were collected at 450m depth at night in the northern Gulf of Mexico using Niskin bottles, and we targeted the fish community using the MiFish primer (12S rRNA). Our results show that 1L is insufficient to characterize deep-sea pelagic fish communities. The treatments 5L and 10L detected similar communities and numbers of species, but 10L detected more exact sequence variants (ESVs). Five biological replicates can detect up to 80% of the species detected in the water collected in both 5L and 10L treatments. Extrapolation analysis indicates that 10 replicates

would also potentially recover 100% of the species. More biological replicates would be necessary to detect all ESVs. PCR replicates also had an important role in taxa and ESV detection, being the best strategy to perform more replicates if you filter less water (optimal number inversely proportional to the water volume filtered). We suggest that future studies collect at least 5L, 5 field replicates, and 5-10 PCR replicates to adequately investigate deep-sea pelagic biodiversity using eDNA, considering the limitations we mention. Our study provides guidance for future eDNA studies and a potential route to democratize eDNA studies at a global scale.

Key-words: deep sea, replicates, MiFish, Gulf of Mexico, ESV

INTRODUCTION

The use of DNA extracted from environmental samples (environmental DNA, eDNA) has become a common method to describe biodiversity and test ecological hypotheses (Dickie et al., 2018; Ruppert et al., 2019). Because this is a non-invasive approach, it has also gained popularity to inform conservation and management practices (Barnes & Turner, 2016; Beng & Corlett, 2020). eDNA can be particularly useful for the documentation of species diversity when a habitat is hard to access, or the target organisms are very small (Stat et al., 2017). Although this method is becoming a popular tool for coastal systems, it is less frequently employed to characterize deep-sea pelagic ecosystems (>200 m depth). The deep-sea, including the midwater, is the largest habitat on Earth but one of the least understood (Ramirez-Llodra et al., 2010). Deep-pelagic animals participate in diel vertical migration, the largest migration on the planet in terms of biomass, and play a critical role in the biological pump and transport of nutrients, acting

as the lifeline between the surface and deep ecosystems. In addition, there is potential economic importance represented by natural resources exploration, from fisheries to deep-sea mining (Thurber et al., 2014). Investigations of the deep sea and associated diversity are challenging due to logistical constraints regarding access, collecting, and filtering water samples. In this context, eDNA has the potential to revolutionize our understanding of this enigmatic and vast habitat.

Advances in robotics are improving sampling strategies, as seen using autonomous vehicles in combination with *in situ* filtration for both surface and deep-pelagic communities (Billings et al., 2017; Yamahara et al., 2019; Yoerger et al., 2021; Govindarajan et al., 2022). However, this type of technology might not be accessible for many scientific groups due to the costs associated with specialized equipment or trained personnel, limiting the global use of this approach (Govindarajan et al., 2022). Additionally, the use of specialized equipment and/or numerous materials for eDNA sampling/filtering is not possible on all vessels due to limited space and technological capabilities. Until advanced technologies are available for many research groups, there is a need to further investigate experimental designs that use common equipment with limited resources. This can facilitate the number of research groups performing eDNA in the deep sea, thus enabling large-scale collaborative studies.

Many eDNA studies, including the ones investigating the deep-pelagic habitat, use Niskin bottles to collect water samples, a common piece of equipment on research vessels. Usually, Niskin bottles are coupled with a conductivity temperature depth (CTD) rosette and triggered at a specific depth and location (McClenaghan et al., 2020; Govindarajan et al., 2021; Visser et al., 2021; Govindarajan et al., 2023). After collection, water samples are filtered, and the filters are

94 preserved for future processing. Although there is some general guidance for eDNA studies (e.g.,
95 Dickie et al., 2018), little is discussed about specific sampling schemes focused on mid- and
96 deep-water pelagic systems. If using Niskin bottles to collect water is a viable option, the amount
97 of water and the number of biological replicates becomes a limitation because these factors are
98 constrained by the size and number of bottles available. For instance, filtering many replicates of
99 large amounts of water can be a laborious and sometimes impossible task (i.e. time constraints,
100 filter clogging, and pump capacity). These limitations should be central in discussing deep-
101 pelagic eDNA explorations as they will dictate the experimental design, representing a first,
102 critical step in eDNA studies.
103
104 Besides equipment limitations, the pelagic deep-sea environment presents other unique
105 challenges to consider when designing eDNA-based investigations. For instance, the magnitude
106 of this habitat and often low density of individuals, can result in low levels of eDNA. Previous
107 work has suggested that higher volumes of water, biological and PCR replicates might be
108 necessary in offshore environments, where DNA can be sparsely distributed (Kumar et al., 2021;
109 Cote et al., 2023); however, no experimental study has proposed or investigated potential
110 universal standards (water volume, biological and PCR replication) for deep-pelagic biodiversity
111 studies. Although obvious, one should not assume that standard sampling methods from coastal
112 systems are sufficient or translatable for deep-pelagic waters.
113
114 Here, we investigate the current methodological knowledge gaps and provide guidance for future
115 deep-pelagic eDNA studies. We test the impact of sampled water volume, biological and PCR
116 replicates in assessing deep-sea fish pelagic biodiversity using common equipment found on

research vessels (CTD rosettes and Niskin bottles). We collected 1L, 5L, and 10L simultaneously within the deep-scattering layer (450 m depth) in the northern Gulf of Mexico which was targeted using acoustic methods. Each treatment was represented by five biological replicates (15 samples total). We used MiFish primer set and aimed for 10 PCR replicates per biological replicate and investigated species and exact sequence variant (ESV) detection. This experimental design allowed us to detect the optimal sampling scheme, acknowledging the challenges associated with pelagic deep-sea sampling and the limitations of using accessible gear.

MATERIALS AND METHODS

Experimental design

We investigated the effects of water volume and biological and PCR replicates by examining samples collected simultaneously at 450m depth (see below). Our experimental design consisted of three treatments: 1L, 5L, 10L. Each treatment consisted of five biological replicates (n = 15) plus one negative control. The negative control is represented by 1L of ultrapure water filtered using the same apparatus that was used for the biological samples (see below). To investigate the effects of technical replicates, 10 PCR replicates were run for each biological replicate and the negative control when possible (n = 160).

Sampling and shipboard processing

Water samples were collected during a cruise on R/V Point Sur in the Northern Gulf of Mexico (29º00.175N, 87º31.131W) as part of the DEEPEND|RESTORE consortium (Cook et al., 2020). Samples were collected at 450m depth (nighttime) during a cast of a CTD (Sea-Bird 911plus) mounted with eleven 10-L Niskin Bottles. To maximize the detection of individuals, multifrequency acoustic backscatter data was collected with a calibrated pole-mounted echosounder system (Simrad EK60 and EK80). The transducers were mounted in an enclosed housing and suspended 2.5 m below the water surface (see Boswell et al., 2020 for additional details). Backscatter data were collected simultaneously at four frequencies (18, 38, 70, and 120 kHz). This was used to detect the deep scattering layer (DSL) in real-time and guide us in the precise moment to trigger the Niskin Bottles. Because one of our objectives was to investigate the effect of filtering different water volumes, all samples were collected simultaneously during the same CTD cast to avoid any temporal or spatial variation in our results.

Once on board, seawater was transferred from Niskin bottles to 1-L sterile Nalgene bottles and immediately transported to the laboratory and filtered. We used a MilliporeSigma™ Chemical Duty Vacuum Pressure Pump (WP6111560) for filtration of all biological replicates and a negative control. Each replicate of specific seawater volume (1L, 5L, or 10L) was filtered through a sterile hydrophilic polyethersulfone (PES) filter (Pall Corporation; 47-mm diameter; 0.45-µm pore size). To prevent contamination, the filtering station was cleaned with 10% bleach, and nitrile gloves were worn throughout the filtering protocol and were changed frequently. Following sample filtration, the filters were placed in falcon tubes with silica beads, which function as a desiccator, drying out the filters and reducing DNA degradation (Bakker et al., 2017). Samples were stored at -80ºC in a freezer onboard. Upon arrival, samples were

162 transported in dry ice to Florida International University and stored in -80ºC freezer until sample
163 processing.
164
165 *eDNA: Sample Processing, DNA extraction, PCR, sequencing*
166
167 Genomic DNA from samples was extracted using the DNeasy Blood & Tissue Kit (250) (Cat.
168 No. / ID: 69506) according to the manufacturer's protocol. Whole filters (47 mm) were used for
169 genomic DNA extraction. Genomic DNA was eluted into 200µl and frozen at -20°C.
170 We decided to perform this investigation focusing on the fish community based on their
171 abundance and ecological relevance in the deep-pelagic environment (Irigoien et al., 2014;
172 Iglesias et al., 2023), and because of the availability of primers capable of giving species
173 resolution in deep-sea fish communities (McClenaghan et al., 2020; Canals et al., 2021; Polanco
174 F. et al., 2021). Portions of hyper-variable regions of the mitochondrial 12S ribosomal RNA
175 (rRNA) gene were PCR amplified from each genomic DNA sample using the MiFishUF and
176 MiFishUR primers with spacer regions (Miya et al., 2015). Both forward and reverse primers
177 also contained a 5' adaptor sequence to allow for subsequent indexing and Illumina sequencing.
178 Each 25 µL PCR reaction was mixed according to the Promega PCR Master Mix specifications
179 (Promega catalog # M5133, Madison, WI) which included 12.5ul Master Mix, 0.5 µM of each
180 primer, 1.0 µl of gDNA, and 10.5 µl DNase/RNase-free H2O. DNA was PCR amplified using
181 the following conditions: initial denaturation at 95ºC for 3 minutes, followed by 45 cycles of 20
182 seconds at 98ºC, 30 seconds at 60ºC, and 30 seconds at 72ºC, and a final elongation at 72ºC for
183 10 minutes. PCR amplification was performed in replicates of ten (PCR replicates) per biological
184 replicate. We had problems amplifying some 1L samples, resulting in few PCR replicates for two

185 of the biological replicates. We decided to reallocate the resources by increasing one PCR
186 replicate to the other biological replicates. PCR replicates were not pooled and were kept
187 separate. Each reaction was visually inspected using a 2% agarose gel with 5µl of each sample as
188 input to determine amplicon size and PCR efficiency. Successful PCRs were cleaned by
189 incubating amplicons with Exo1/SAP for 30 minutes at 37ºC following by inactivation at 95ºC
190 for 5 minutes and stored at -20ºC.
191
192 A second round of PCR was performed to complete the sequencing library construct, appending
193 with the final Illumina sequencing adapters and integrating a sample-specific,12-nucleotide
194 index sequence. The indexing PCR included Promega Master mix, 0.5 µM of each primer and 2
195 µl of template DNA (cleaned amplicon from the first PCR reaction) and consisted of an initial
196 denaturation of 95°C for 3 minutes followed by 8 cycles of 95°C for 30 sec, 55°C for 30
197 seconds and 72°C for 30 seconds. Final indexed amplicons from each sample were cleaned and
198 normalized using SequalPrep Normalization Plates (Life Technologies, Carlsbad, CA). 25µl of
199 PCR amplicon is purified and normalize using the Life Technologies SequalPrep Normalization
200 kit (cat#A10510-01) according to the manufacturer's protocol. Samples are then pooled together
201 by adding 5µl of each normalized sample to the pool. Library preparation was performed by
202 Jonah Ventures (Boulder, CO).
203
204 Sample library pools were sent for sequencing on an Illumina MiSeq (San Diego, CA) at the
205 Texas A&M Agrilife Genomics and Bioinformatics Sequencing Core facility using the v2 500-
206 cycle kit (cat# MS-102-2003). Necessary quality control measures were performed at the
207 sequencing center prior to sequencing.

208

209 *eDNA: Bioinformatics*

210

211 Raw sequence data were demultiplexed using pheniqs v2.1.0 (Galanti et al., 2021), enforcing
212 strict matching of sample barcode indices (i.e, no errors). Cutadapt v3.4 (Martin, 2011) was then
213 used remove gene primers from the forward and reverse reads, discarding any read pairs where
214 one or both primers (including a 6 bp, fully degenerate prefix) were not found at the expected
215 location (5') with an error rate < 0.15. Read pairs were then merged using vsearch v2.15.2
216 (Torbjørn et al., 2016), discarding resulting sequences with a length of < 130 bp, > 210 bp, or
217 with a maximum expected error rate > 0.5 bp (Edgar & Flyvbjerg, 2015). For each sample, reads
218 were then clustered using the unoise3 denoising algorithm (Edgar, 2016) as implemented in
219 vsearch, using an alpha value of 5 and discarding unique raw sequences observed less than 8
220 times. Counts of the resulting exact sequence variants (ESVs) were then compiled and putative
221 chimeras were removed using the uchime3 algorithm, as implemented in vsearch. For each final
222 ESV, a consensus taxonomy was assigned using a custom best-hits algorithm and a reference
223 database consisting of publicly available sequences (GenBank, Benson et al., 2005) as well as
224 Jonah Ventures voucher sequences records. Reference database searching used an exhaustive
225 semi-global pairwise alignment with vsearch, and match quality was quantified using a custom,
226 query-centric approach, where the % match ignores terminal gaps in the target sequence, but not
227 the query sequence. The consensus taxonomy was then generated using either all 100% matching
228 reference sequences or all reference sequences within 1% of the top match, accepting the
229 reference taxonomy for any taxonomic level with > 90% agreement across the top hits. Non-fish
230 species were removed from the final dataset. For downstream analyses, we kept only ESVs that

could be confidently assigned to a species (>98% similarity, Polanco F. et al.,2021). We kept species that were identified at the genus level if they had a reference deposited in public databases and an associated voucher number. The number of reads was coded as presence/absence. The final datasets consisted of 1) presence/absence of each species per biological replicate per treatment; 2) presence/absence of each species per PCR replicate per biological replicate per treatment; 3) presence/absence of each ESV per biological replicate per treatment; 4) presence/absence of each ESV per PCR replicate per biological replicate per treatment.

Biodiversity analyses

We compared the number of species (richness) and ESVs detected in each treatment (1L, 5L, 10L) by performing a 1-way analysis of variance (ANOVA) followed by a posthoc Tukey test. For comparisons of species and ESVs composition across treatments, we performed a Permutation Multivariate Analysis of Variance (PERMANOVA; Anderson, 2001). First, we built a dissimilarity matrix based on the Bray-Curtis index using the vegdist function in the vegan package (Oksanen et al., 2022). Because the package's default is to use quantitative data for all indices, we set binary = TRUE to account for the presence/absence structure of our dataset. The resulting matrix was used as the response variable and volume as the explanatory variable in the PERMANOVA model using the adonis2 function. Significance was accessed using 9999 permutations. To investigate the effects of each treatment, we performed a pairwise permutational test for multiple comparisons set to 9999 permutations using the pairwise.adonis2 function in the pairwiseAdonis (Arbizu, 2017). The relationships among each biological replicate

254 per treatment were visually represented by a Non-metric Multidimensional Scaling (nMDS) plot.
255 This analysis was done using the metaMDS function implemented in the vegan package.
256 Because convergence was not reached under default parameters, we used the argument trymax =
257 1000 to increase the number of random starts to find a stable solution. A permutation-based
258 similarity percentage (PER-SIMPER) analysis was performed to identify the taxa that most
259 contributed to the significant differences among treatments on both species and ESV datasets
260 (Gilbert & Escarguel, 2019). We also visualized the species and ESVs detected in each treatment
261 by plotting a Venn Diagram using ggvenn package (Yan, 2023), and used the gplots package
262 (Warnes et al., 2024) to find the exclusive and shared species/ESVs among treatments.
263
264 We investigated the efficiency of the number of biological replicates and PCR replicates in
265 detecting biodiversity by constructing rarefaction and extrapolation (R/E) plots using the
266 package iNterpolation/EXTrapolation (iNEXT, Hsieth et al., 2016). We followed a non-
267 asymptotic approach to plot a coverage-based R/E sampling completeness curve based on
268 sampling-unit-based incidence data. Depending on the analysis, the sampling unit was
269 considered the number of biological replicates (n = 5) per treatment (1L, 5L, 10L) or the number
270 of PCR replicates (n = 10) per biological replicate (n = 5) per treatment (1L, 5L, 10L). R/E plots
271 were constructed for both species and ESV datasets.
272 All analyses were run in R software (version 4.1.3).
273
274 RESULTS
275

We retained 2,061,989 reads that were assigned to the Class Actinopteri. After all filtering steps, we retained 1,487,217 reads in total that could be assigned to a species (>98% similarity). The number of reads per biological treatment was 1L = 48,793.4 ± 49,311.5; 5L = 100,506.2 ± 27,211.7; 10L = 148,143.8 ± 51,240.5. The number of reads per replicate per biological treatment was 1L = 4,900.3 ± 4,847.5; 5L = 9,136.9 ± 2,473.8; 10L = 13,467.6 ± 4,658.2. One biological replicate of the 1L treatment detected zero species. Negative controls had negligible amounts of contamination, with an average of 408.4 ± 1890.8 across PCR replicates. Only 6 species were found in the negative control, usually represented by one ESV in just one of the PCR replicates. We detected 98 taxa that could be confidently assigned to a species, representing a total of 661 ESVs (Supplementary Table 1).

The richness detected by the 1L treatment is lower than the number detected by the treatment filtering 10L (Tukey post-hoc, $p = 0.004$), but the number of species detected in 5L and 10L was not statistically different ($p = 0.109$; Table 1; Figure 1A). ESVs detected by 1L treatments are lower than the number of ESVs detected by the 10L treatment ($p = 0.004$); ESVs detected by 5L treatments are marginally lower than the 10L ($p = 0.048$; Table 1; Figure 1B). In both cases, 1L and 5L are not different (species, $p = 0.206$; ESVs, $p = 0.399$).

Table 1. Analysis of Variance (ANOVA) results comparing number of species (richness) found in each treatment. Permutation Multivariate Analysis of Variance (PERMANOVA) results comparing community structure among biological replicates per treatment (1L, 5L, 10L). DF: degrees of freedom; SS: sum of squares; F: F-value; Pseudo-F: pseudo-F-value; P: p-value.

	ANOVA - Richness							
	Species				**ESVs**			
Source of Variation	*DF*	*SS*	*F*	*P*	*DF*	*SS*	*F*	*P*
Volume	2	1493	8.112	0.005	2	1493	8.112	0.005
Residuals	12	1104			12	1104		
	PERMANOVA - Community Structure							
	Species				**ESVs**			
Source of Variation	*DF*	*SS*	*Pseudo-F*	*P*	*DF*	*SS*	*Pseudo-F*	*P*
Volume	2	0.71959	1.8027	0.004	2	0.9926	1.3744	0.002
Residuals	12	2.395			12	5.3258		

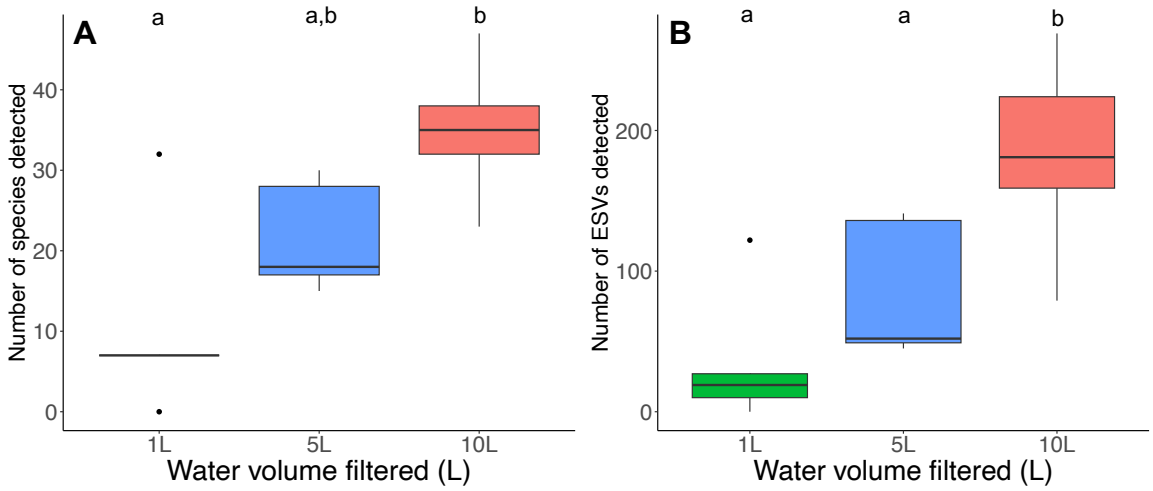


Figure 1. Box-plot representing the number of species (A) and ESVs (B) detected in each treatment (1L, 5L, 10L). Letters represent the result of the *post-hoc* Tukey test (similar letters are not statistically different). Colors represent different treatments.

Similarly, the PERMANOVA analysis indicated that volume affects community composition (Table 1). Considering the species dataset, the community detected by 1L marginally differs from the 5L (pairwise permutational test for multiple comparisons, $p = 0.053$) and differs from the 10L communities ($p = 0.017$), but 5L and 10L are not different ($p = 0.113$; Figure 2A). Considering the ESVs dataset, each treatment detected a different community composition (1L vs. 5L, $p = 0.04$; 1L vs. 10L, $p = 0.017$; 5L vs. 10L, $p = 0.031$, Figure 2B). When looking at the species contributing to the significant differences found in each comparison, the PER-SIMPER analysis indicates the top 15 species contributed 46.8% to the dissimilarity between 1L vs. 5L, and 34.2% between 1L vs. 10L (Table 2). Although each species contributes little to the dissimilarity found, PER-SIMPER indicates that 8 out of the 15 species had a role in both cases:

325 *Barbourisia rufa, Benthosema suborbitale, Caranx crysos, Cyclothone pallida, Diaphus lucidus,*
326 *Gonostoma elongatum, Lampanyctus alatus,* and *Rondeletia loricata.* Twenty-two species
327 (22.4%) were found in all treatments (Figure 3, Supplementary Table 1). Exclusive species
328 represent species detected in only one biological replicate in 100% (1L = 11, and 5L = 7) or 68%
329 (10L = 22) of the cases, with 10L being the treatment that had the highest number of exclusive
330 species (32)(Figure 3A, Supplementary Table 1). Regarding ESVs, the top 15 ESVs contribute to
331 33.6% (1L vs. 5L), 30.7% (1L vs. 10L), 25.6% (5L vs. 10L) (Supplementary Table 1). Most of
332 the ESVs contributing to the differences between 1L vs. 5L and 10L indicated by PER-SIMPER
333 are ESVs from the species detected in the species-level analysis (Table 2, Supplementary Table
334 1). In the case of the 5L vs. 10L comparison, all ESVs represent species shared between the two
335 treatments (Supplementary Table 1). Only twenty-seven ESVs were shared among all treatments
336 (4.1%), and the 10L treatment had the highest amount of exclusive ESVs (351, 53.3%) (Figure
337 3B)
338
339 Table 2. Permutation-based similarity percentage (PER-SIMPER) results showing the top 15
340 species contributing to the differences found between statistically different communities (see
341 PERMANOVA results).

1L vs. 5L		1 vs. 10L	
Species	%	Species	%
Cyclothone pallida	4.3	*Lagocephalus laevigatus*	3.2
Gonostoma elongatum	4.3	*Canthigaster rostrata*	2.8
Diaphus lucidus	3.6	*Cyclothone pallida*	2.8
Caranx crysos	3.3	*Diaphus lucidus*	2.8
Cyclothone pseudopallida	3.3	*Gonostoma elongatum*	2.8
Barbourisia rufa	3.3	*Benthosema suborbitale*	2.3
Lampanyctus alatus	3.1	*Notolychnus valdiviae*	2.3
Melamphaes simus	3.1	*Caranx crysos*	2.1
Diaphus perspicillatus	3.0	*Lampanyctus alatus*	2.1

Cyclothone obscura	2.8	*Barbourisia rufa*	2.0
Rondeletia loricata	2.8	*Cyclothone acclinidens*	1.9
Benthosema suborbitale	2.5	*Rondeletia loricata*	1.9
Lepidophanes guentheri	2.5	*Nannobrachium lineatum*	1.9
Seriola rivoliana	2.4	*Caranx hippos*	1.8
Diaphus rafinesquii	2.4	*Lutjanus campechanus*	1.8
Total	46.8		34.3

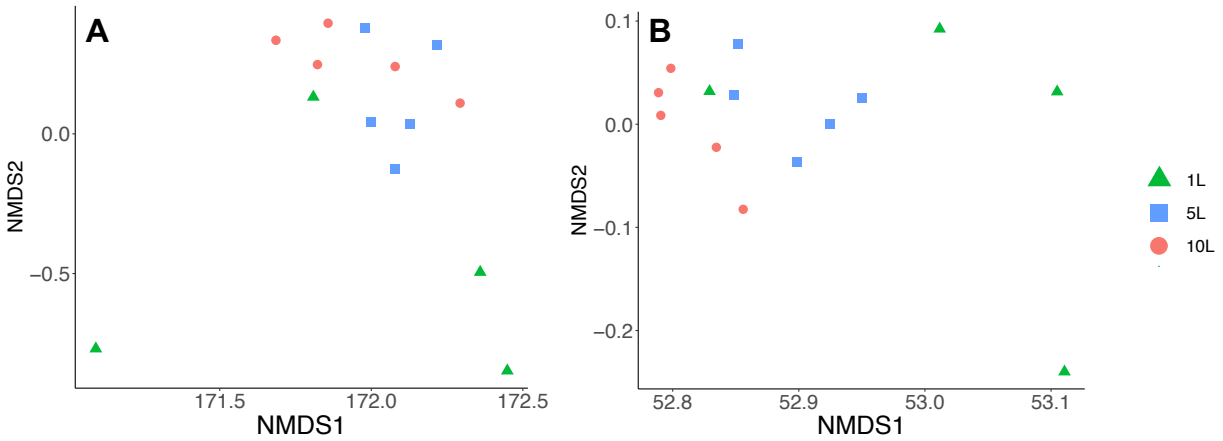


Figure 2. Non-metric multi-dimensional scaling (NMDS) plot representing the species (A) and ESVs (B) composition found in each treatment (1L, 5L, 10L). The 1L replicate that detected zero species/ESVs was excluded to facilitate visualization of the spatial distribution of samples. Colors represent different treatments.

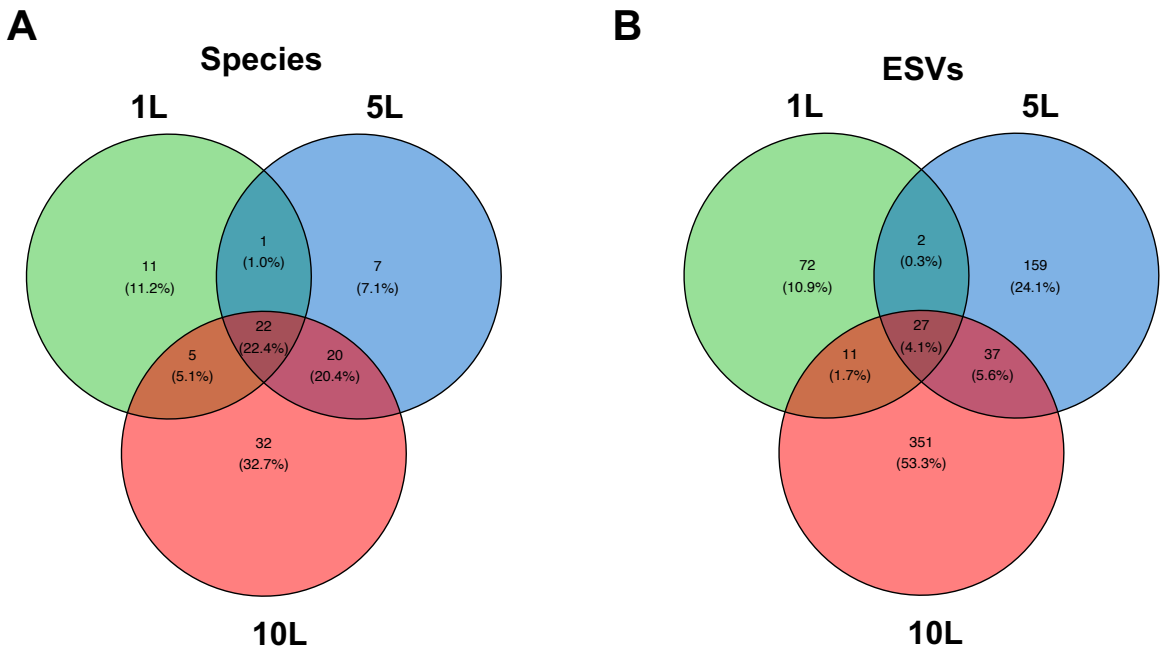


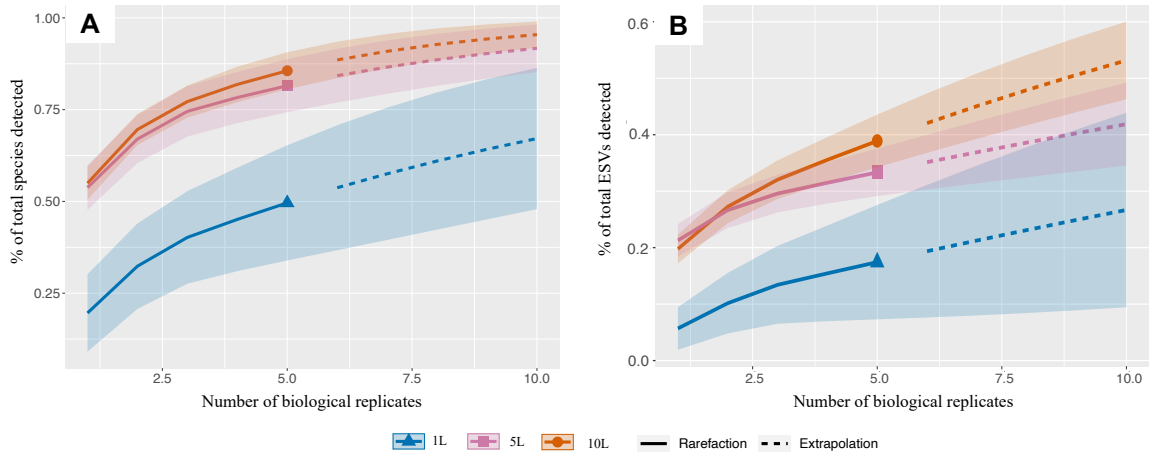
Figure 3. Venn diagram indicating the number of exclusive species in each treatment (1L, 5L, 10L), and shared species between two treatments or shared among all treatments (A); Venn diagram indicating the number of exclusive ESVs in each treatment (1L, 5L, 10L), and shared ESVs between two treatments or shared among all treatments.

When analyzing the impact of the number of biological replicates in detecting the total biodiversity across all of our samples, we show that 5 biological replicates could detect up to 80% of all 98 species in both 5L and 10L treatments. The extrapolation analysis indicates that 10 replicates would recover 100% of the species detected in the water collected (note: not total biodiversity from the region, but all 98 species). In contrast, 5 biological replicates of 1L treatment detected only 50% of the total biodiversity, and the extrapolation analysis indicates that more than 10 biological replicates would be necessary to detect 100% (Figure 4A). A different scenario is found when analyzing ESVs, showing that 5 biological replicates detected

365 ~40% of the 611 ESVs detected. Extrapolation analyses indicated that more than 10 biological

366 replicates would be necessary to detect all ESVs (Figure 4B).

367



368

369 Figure 4. Rarefaction and extrapolation plot depicting the increase of species (A) or ESVs (B)

370 detected as more biological replicates were added per treatment (1L, 5L, 10L).

371

372 When analyzing the impact of the number of PCR replicates per biological replicate, we find that

373 10 PCR replicates show variable biodiversity coverage for 5L, detecting between 60-85% of total

374 number of species depending on the biological replicate (Figure 5B). Similarly, 10L biological

375 replicates show a similar pattern, but detecting around 80-90% of the total number of species

376 (Figure 5C). In contrast, results concerning PCR replicates of 1L treatments are extremely

377 variable (Figure 5A). Similar to the previous results, PCR replicates R/E plots analyzing ESVs

378 indicate large variation among 1L biological replicates (Figure 5D), and moderate/low variation

379 among 5L and 10L biological replicates. In these cases, 10 PCR replicates detected between 20-

380 60% of the ESVs sampled in this study (Figure 5E-F). One of the biological replicates in the 1L

381 treatment did not have enough DNA for running all PCR replicates.

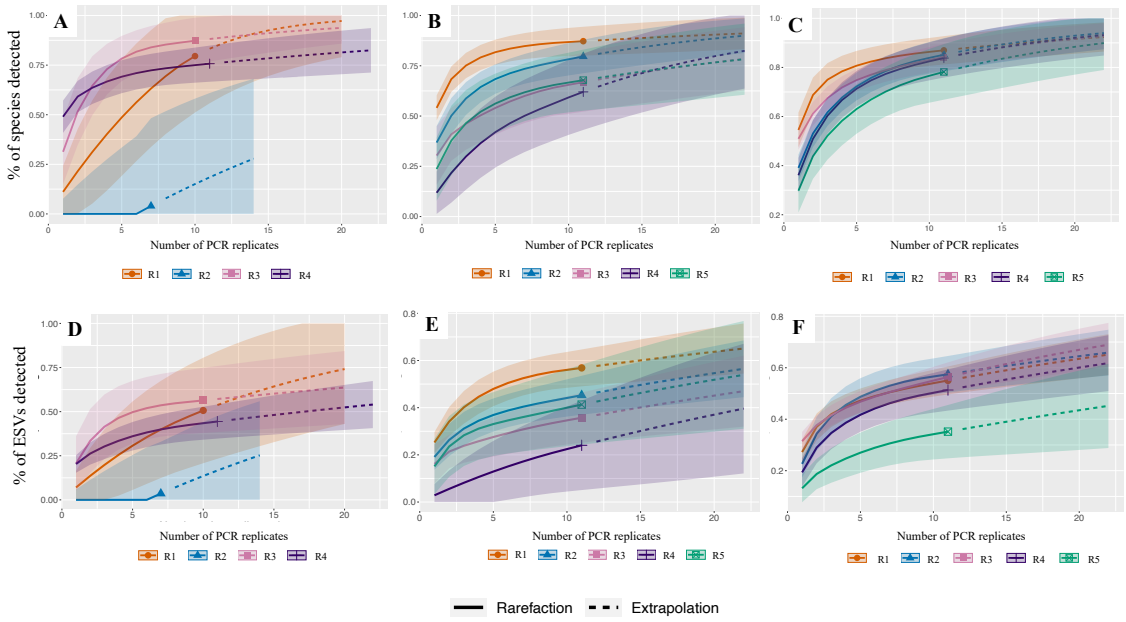


Figure 5. Rarefaction and extrapolation plot depicting the increase of species/ESVs detected as more PCR replicates were added biological replicate per treatment (A: species-1L; B: species-5L; C: species10L; D: ESVs-1L; E: ESVs-5L; F: ESVs-10L).

DISCUSSION

Experimental design is a crucial part of any ecological study, but there are no golden rules when implementing eDNA approaches (Deiner et al., 2017; Dickie et al., 2018; Ruppert et al., 2019; Cote et al., 2023). Sampling design considerations must be carefully considered before going to sea, especially for deep-pelagic eDNA investigations. As far as we know, previous investigations have not tested how different sampling schemes (water volume and replication) using common

396 equipment (e.g., Niskin Bottles coupled with CTD) can impact deep-sea biodiversity
397 assessments. Our study shows that common practices adopted in coastal waters (filtering 1L of
398 water) are not ideal for characterizing deep-pelagic fish biodiversity. Filtering 5L or 10L did not
399 result in statistically different species compositions but detected different ESV compositions. In
400 terms of replication, 5 or more biological replicates should be the standard for similar
401 experimental designs using Niskin bottles, and the number of PCR replicates should be higher
402 when the volume of water filtered is lower. Overall, results are similar when investigating
403 species and ESVs. We recommend 5-10 PCR replicates per biological replicate to maximize the
404 detection of taxa and different ESVs. Our results have implications for future eDNA studies
405 targeting deep-pelagic communities and for future eDNA studies comparison.
406
407 Water volume
408
409 In aquatic eDNA studies, a standing question is: how much water should be filtered? Data from
410 studies conducted in shallow systems indicate that filtered water volume is known to impact taxa
411 detection (Muha et al., 2019; Kumar et al., 2022). In these cases, filtered water volume usually
412 varies from 0.5 mL to 2L (Rees et al., 2014; Bessey et al., 2020; Kumar et al., 2022). Directly
413 transposing shallow water standards to mid- and deep-pelagic ecosystems eDNA investigations
414 might not be ideal considering the differences between environments. The open ocean is the
415 largest environment on Earth, covering 70% of our planet, and the deep sea represents 80-90% of
416 our oceans, making it the largest habitat on Earth (Ramirez-Llodra et al., 2010). Therefore,
417 species abundance and distribution are expected to differ substantially across spatial scales when
418 comparing shallow- and deep-pelagic ecosystems. Few studies have investigated the impacts of

419 different filtering protocols, resulting in similar water volumes being filtered between shallow
420 water and deep-pelagic eDNA studies (McClenaghan et al., 2020; Diao et al., 2023). Our results
421 indicate that 1L is far from ideal for characterizing deep-pelagic fish communities. Although
422 some comparisons are not statistically significant, NMDS plots and rarefactions curves suggest
423 large variation among 1L biological replicates in comparison to 5L and 10L treatments. This can
424 be explained by the low eDNA abundance per volume of water in large habitats, leading to
425 stochastic detection of eDNA molecules in 1L samples. Interestingly, species composition
426 differences between 1L vs. 5L and 10L can be explained by the omission of rare species, but also
427 species known to be very common in deep-pelagic habitats within the Gulf of Mexico including
428 species from Gonostomatidae (*Cyclothone pallida*, *Gonostoma elongatum*) and Myctophidae
429 (*Lampanyctus alatus*) (Ross et al., 2010). Therefore, we conclude that 1L does not represent the
430 eDNA available nor describe the local fish diversity, indicating that more water must be filtered
431 when investigating deep-pelagic communities.
432
433 Surprisingly, 5L and 10L detected similar numbers of fish species per biological replicate and
434 showed the same species composition, even though the mean number of reads in the 10L
435 treatment was almost 50% larger than in the 5L treatment. A comparison of the number of
436 operational taxonomic units (OTUs) found in different filtered water volumes (~5-60L) also
437 indicated no changes in total OTUs, although this varied according to location (Yoshida et al.,
438 2023). It should be noted that species detection based on eDNA is constrained by the persistence
439 of DNA molecules in the water, correct amplification via PCR, and reference libraries that
440 contain the necessary information to identify species correctly. Our results are relevant when
441 considering the effort-time-cost balance related to eDNA fieldwork that does not involve an

442 autonomous water sampler and specialized filtering apparatus. Collecting and filtering water can
443 be laborious and time-consuming and involve multiple researchers when using commonly
444 available gear, like Niskin bottles and vacuum pumps. Here, we suggest that, when limitations to
445 water collection exist, efforts toward increasing filtered water volume above 10L may not impact
446 the recovery of deep-pelagic species. A similar conclusion was shown when comparing 250mL
447 and 1.5L collected using a CTD (McClenaghan et al., 2020), but also when comparing CTD
448 samples (~2L) and a recently developed large-volume autonomous eDNA sampler with in-situ
449 filtration mounted on the midwater robot Mesobot (~40-60L) (Govindarajan et al., 2022). In the
450 latter study, the community composition detected was not different when comparing the CTD
451 samples and the Mesobot, but the number of sequence reads and ESV was higher when using the
452 autonomous samples (Govindarajan et al., 2022). This is similar to our findings, as the 10L
453 treatment detected more ESVs and a different ESVs community composition than the 1L and 5L
454 treatments. A striking conclusion is that 10L did not detect more species, but detected more
455 intraspecific variation (i.e., more ESVs per species). We believe both 5L and 10L treatments
456 reflected the natural proportion of fish eDNA that could be detected because filtering more water
457 did not detect more species. Therefore, choosing how many liters of water will be filtered should
458 also consider whether the detection of more ESVs is part of the goal of the study. For instance,
459 recent analytical developments have suggested that the number of intraspecific ESVs can be used
460 to estimate species abundance (Yoshitake et al., 2019). The understanding of the impacts of
461 filtered volumes on the characterization of deep-pelagic communities is still under development,
462 and more studies will illuminate whether there is an ideal volume that can be filtered. We should
463 also note that Govindarajan et al. (2022) targeted the 18S rRNA region and many metazoan
464 groups at lower taxonomic resolution, while we targeted 12S rRNA and fish species at a higher

465 taxonomic resolution. The taxonomic resolution and efficiency vary according to the primer used
466 and the region targeted (Ruppert et al., 2019; Zhang et al., 2020; Polanco et al., 2021; Kumar et
467 al., 2022), being an important step during experimental design. We used a taxon-specific primer
468 (MiFish) to better understand the nuances of sampling schemes at the highest taxonomic
469 resolution (i.e., species). It is not possible to say if the results of Govindarajan et al. (2022)
470 would be different whether they had targeted a different region (i.e., used a different primer) or
471 investigated higher taxonomic resolution. It is important to notice that we are discussing these
472 data under a community composition framework, with the goal of comparing samples across
473 time and space. When the main goal of an investigation is to detect one or a few specific and/or
474 rare taxa, higher volumes should be aimed as DNA detection likely increases (Sepulveda et al.,
475 2019; Bessey et al., 2020; Govindarajan et al., 2022). Our study also reports evidence for this
476 latter statement: the 10L detected an absolute number of species higher than 1L and 5L
477 treatments (but the number of species per biological replicate is not different from the 5L
478 treatment). These rare species have a minor role when discussing communities'
479 similarities/differences. We should notice that this is related to the eDNA signal detected, but
480 some species might not be rare in terms of abundance at the organismal level. As mentioned
481 before, this could be explained by species shedding less DNA or primer incompatibilities.
482
483 Biological and PCR replicates
484
485 Another crucial factor to consider when performing eDNA investigations is the number of
486 biological and PCR replicates. Our work is one of the first to address the variation in these
487 factors when characterizing deep-sea systems. Variation among 1L biological and PCR replicates

488 was higher than other treatments for both species and ESVs datasets. Five biological and 10 PCR
489 replicates of the 1L treatment were insufficient to assess the local fish biodiversity (i.e., the
490 species and ESVs detected in this study). In fact, 1L treatments detected a different species/ESV
491 composition and fewer species/ESV than the other treatments, and even our extrapolation
492 analyses could not reach levels compared to the other treatments. This means that not only the
493 volume of water but also the number of replicates is important when considering eDNA applied
494 to mid- and deep-water pelagic systems. Our results indicate that, under certain circumstances, it
495 could be better to favor less volume and more replicates if one is limited on the amount of water
496 that they can sample using their available gear (e.g., Niskin bottles coupled with CTD can only
497 sample a specific total volume of water). We provide empirical data to the literature reinforcing
498 previous suggestions for other systems recommending more replicates to improve the detection
499 probabilities and to estimate fish biodiversity reliably (Bessey et al., 2020; Stauffer et al., 2021;
500 Cote et al., 2023; Yoshida et al., 2023). Considering species detection, five biological replicates
501 of the 5L and 10L treatments did not yield different results, but biodiversity sampling was
502 maximized by adding more biological replicates in our study. Therefore, an optimal deep-sea
503 eDNA experimental design could aim for a minimum of five biological replicates of a minimum
504 of 5L, which is doable considering most traditional water samplers (e.g., McClenaghan et al.,
505 2020; Feng et al., 2022; Govindarajan et al., 2023). However, future studies should address this
506 topic because the trade-offs between water filtered and the number of biological replicates using
507 autonomous samplers have yet to explored.
508
509 Similarly, the investigation of the ideal number of PCR replicates is still scarce in the literature.
510 Amplification is a stochastic process during PCR reactions, and more reactions will probably

511 result in the detection of more taxa by chance. The exact number needed will vary depending on
512 the probability of DNA detection, which varies according to species abundance, rate of shedding
513 DNA, environmental conditions, and other factors (Ruppert et al., 2019). A simulation study
514 showed that four PCR replicates might result in high levels of false negatives, and six to eight
515 PCR replicates should be performed depending on the probability of DNA detection (Ficetola et
516 al., 2015). As discussed before, it is feasible to assume that DNA is sparsely distributed in the
517 deep sea, considering this is the largest habitat on Earth and individuals' density might not be
518 high. Still, PCR replicates are often low in studies investigating deep waters (2-4 PCR replicates;
519 e.g., McClenaghan et al., 2020; Canals et al., 2021; Govindarajan et al., 2022; Cote et al. 2023;
520 Yoshida et al., 2023). Our results show that all samples benefited from more PCR replicates
521 when considering species and ESVs detection. Besides biological replicates, the ideal sampling
522 scheme should balance the volume of water filtered with the number of PCR replicates. Our
523 results indicate that at least 5-10 PCR replicates should be the goal, considering the water
524 volume filtered. That is, if limited by the water volume, the best strategy is to increase the
525 number of PCR replicates; if not limited by water volume, fewer PCR replicates can be
526 performed. Ideally, future studies should first find the optimal design to maximize water volume
527 and PCR replicates accordingly.
528
529 Limitations
530
531 It is important to mention the limitations of using common equipment in deep-sea eDNA studies.
532 The maximum water volume is limited by the capacity and number of Niskin Bottles attached to
533 the CTD and the number of deployments within a scientific mission. Our study leverages

534 common gear found on research vessels and could not test if filtering more than 10L would
535 change the biodiversity characterization, especially at the species level, as ESV diversity is
536 favored by filtering more water. Previous studies did not find different species composition when
537 comparing filtering 2.5L using Niskin Bottles and 40-60L using an autonomous vehicle in the
538 Gulf of Mexico and targeting the 18S region (Govindarajan et al., 2022). However, community
539 characterization changed in some locations near Japan after filtering >25L of water per
540 biological replicate and targeting the 12S region (Yoshida et al., 2023). A potential future
541 standard in this type of study could be first to investigate the ideal water volume needed to
542 characterize the community at the specific ocean basin using a specific primer. Considering our
543 results and Govindarajan et al. (2022), 5L-10L might be sufficient for the northern Gulf of
544 Mexico when targeting the 18S and 12S rRNA markers. We should note that some taxa might
545 shed less DNA than others (Andruszkiewicz Allan et al., 2021), which should also be considered
546 when planning an eDNA investigation. For instance, we could not detect substantial decapod
547 eDNA using the MiDecap primer set (Komai et al., 2019) using the volume of water and
548 replicates used in this study (data and results not shown). This is surprising, considering the
549 diversity and abundance of pelagic decapods in the Gulf of Mexico collected using nets and with
550 available reference sequences (Burdett et al., 2017; Varela et al., 2021). We predict that primer
551 compatibility and shedding may have been involved in the lack of recovery, but further studies
552 should explore the impact of water volume and reference libraries on the recovery of deep-sea
553 crustaceans and other taxa. Simulation analysis indicates that increasing the water volume
554 filtered could lead to similar species being detected using eDNA and traditional methods,
555 reinforcing this idea (Cote et al., 2023). Additionally, if detecting specific species (instead of
556 community composition) is the goal, one might consider maximizing the water volume collected,

557 as mentioned before (Sepulveda et al., 2019; Bessey et al., 2020; Govindarajan et al., 2022).
558 Still, the exact volume might be taxa-specific and should also be a first step in this type of
559 investigation. Another factor to be considered is that eDNA might degrade while samples are
560 waiting to be or while they are being processed (Holman et al., 2022). In our case, all replicates
561 were filtered in a randomized order to avoid this pitfall. However, the decay rate of deep-sea
562 eDNA samples is still unknown for many taxa, and future studies should address this knowledge
563 gap (but see McCartin et al., 2022 for insights).
564
565 Besides the technical limitations mentioned, the primary pitfall of all deep-water eDNA
566 investigations is the catastrophic lack of reliable reference libraries. eDNA approaches are 100%
567 reliant on the completeness and quality of available reference databases for properly identifying
568 taxa (Bylemans et al., 2018; Collins et al., 2019; Schenekar et al., 2020). In our study, ~50% of
569 our reads could not be confidently assigned to a fish species. Another situation that highlights the
570 importance of reliable reference libraries is that we also detected a few Indo-Pacific species in
571 our final dataset. We cannot confirm if they represent real occurrences or problems in reference
572 sequences available in public databases. This study, among others, highlights the urgent need for
573 deep-pelagic reference libraries that are linked to curated, voucher specimens deposited in
574 museums that can be verified by taxonomic experts. Current efforts, including our own, are
575 expanding deep-sea reference libraries (e.g., Bucklin et al., 2010; Weigand et al., 2019; Varela et
576 al., 2021; Govindarajan et al., 2022), which need to be standardized, shared and extended to as
577 many taxa and locations as possible.
578
579 Conclusions

580 The application of eDNA to mid- and deep-water ecosystems is gaining popularity, however few
581 studies have experimentally tested the effects of different sampling protocols on the recovery of
582 species biodiversity (Canals et al., 2021; Easson et al., 2020; Govindarajan et al., 2021; Laroche
583 et al., 2020; Merten et al., 2021). Our study provides empirical results on how experimental
584 design can impact the characterization of deep-pelagic communities. Specifically, we indicate
585 that a minimum of 5L-10L of water should be filtered when targeting the deep-scattering layer
586 and paired with a minimum of 5 or more biological replicates for deep-sea eDNA studies when
587 investigating species composition. Additionally, we show the importance of more PCR replicates
588 when analyzing deep-sea samples. Technological advances in eDNA samplers are crucial and are
589 expanding our understanding of deep-sea ecosystems. But, until high-powered technology
590 becomes assessable for all, using Niskin bottles coupled with a CTD can be a way to
591 democratize the field and increase our local and global-scale observations. Most importantly,
592 because deep-ocean explorations are constrained by logistics and budget, we believe our work
593 can help in providing some guidance on potential standards and how eDNA data generated by
594 different research groups across the globe could be comparable when the focus is mid- or deep-
595 water community characterization using common gear.
596
597 REFERENCES
598
599 Billings, A., Kaiser, C., Young, C.M., Hiebert, L.S., Cole, E., Wagner, J.K.S., Van Dover, C.
600 L., 2017. SyPRID sampler: a large-volume, high-resolution, autonomous, deep-ocean
601 precision plankton sampling system. Deep Sea Research Part II: topical Studies in
602 Oceanography. Advances in deep-sea biology: biodiversity, ecosystem functioning

and conservation 137, 297–306. https://doi.org/10.1016/j.dsr2.2016.05.007.

Yamahara, K.M., Preston, C.M., Birch, J., Walz, K., Marin, R., Jensen, S., Pargett, D., Roman, B., Ussler, W., Zhang, Y., Ryan, J., Hobson, B., Kieft, B., Raanan, B., Goodwin, K.D., Chavez, F.P., Scholin, C., 2019. In situ autonomous acquisition and preservation of marine environmental DNA using an autonomous underwater vehicle. Front. Mar. Sci. 6, 373. https://doi.org/10.3389/fmars.2019.00373.

Yoerger, D.R., Govindarajan, A.F., Howland, J.C., Llopiz, J.K., Wiebe, P.H., Curran, M., Fujii, J., Gomez-Ibanez, D., Katija, K., Robison, B.H., Hobson, B.W., Risi, M., Rock, S. M., 2021. A hybrid underwater robot for multidisciplinary investigation of the ocean twilight zone. Sci. Robot. 6 (55), eabe1901 https://doi.org/10.1126/scirobotics. abe1901.

Cook, A. B., Bernard, A. M., Boswell, K. M., Bracken-Grissom, H., D'Elia, M., DeRada, S., ... & Sutton, T. T. (2020). A Multidisciplinary approach to investigate deep-pelagic ecosystem dynamics in the Gulf of Mexico following Deepwater Horizon. Frontiers in Marine Science, 7, 548880.

Galanti, Shasha and Gunsalus (2021). Pheniqs 2.0: accurate, high-performance Bayesian decoding and confidence estimation for combinatorial barcode indexing. BMC Bioinformatics 22.

Martin (2011), Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet. journal 17.

Torbjørn et al. VSEARCH: a versatile open source tool for metagenomics. PeerJ 4 e2584 (2016).

Edgar and Flyvbjerg (2015), Error filtering, pair assembly and error correction for next-generation sequencing reads. Bioinformatics 31.

Edgar (2016), UNOISE2: improved error-correction for Illumina 16S and ITS amplicon sequencing. doi: https://doi.org/10.1101/081257.

Benson et al. (2005), GenBank. Nucleic Acids Res 33.

McClenaghan, B., Fahner, N., Cote, D., Chawarski, J., McCarthy, A., Rajabi, H., Singer, G. and Hajibabaei, M., 2020. Harnessing the power of eDNA metabarcoding for the detection of deep-sea fishes. PLoS One, 15(11), p.e0236540.

Canals, O., Mendibil, I., Santos, M., Irigoien, X., & Rodríguez-Ezpeleta, N. (2021). Vertical stratification of environmental DNA in the open ocean captures ecological patterns and behavior of deep-sea fishes. Limnology and Oceanography Letters, 6(6), 339-347.

Polanco F, A., Richards, E., Flück, B., Valentini, A., Altermatt, F., Brosse, S., Walser, J.C., Eme, D., Marques, V., Manel, S., & Albouy, C. (2021). Comparing the performance of 12S

mitochondrial primers for fish environmental DNA across ecosystems. Environmental DNA, 3(6), 1113-1127.

Kumar, G., Farrell, E., Reaume, A. M., Eble, J. A., & Gaither, M. R. (2022). One size does not fit all: Tuning eDNA protocols for high-and low-turbidity water sampling. Environmental DNA, 4(1), 167-180.

Ruppert, K. M., Kline, R. J., & Rahman, M. 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, e00547.

Govindarajan, A. F., McCartin, L., Adams, A., Allan, E., Belani, A., Francolini, R., ... & Herrera, S. (2022). Improved biodiversity detection using a large-volume environmental DNA sampler with in situ filtration and implications for marine eDNA sampling strategies. Deep Sea Research Part I: Oceanographic Research Papers, 189, 103871.

McClenaghan, B., Fahner, N., Cote, D., Chawarski, J., McCarthy, A., Rajabi, H., ... & Hajibabaei, M. (2020). Harnessing the power of eDNA metabarcoding for the detection of deep-sea fishes. PLoS One, 15(11), e0236540.

Diao, C., Wang, M., Zhong, Z., Li, Y., Xian, W., & Zhang, H. (2023). Biodiversity exploration of Formosa Ridge cold seep in the South China Sea using an eDNA metabarcoding approach. Marine Environmental Research, 190, 106109.

Yoshida, T., Kawato, M., Fujiwara, Y., Nagano, Y., Tsuchida, S., & Yabuki, A. (2023). Optimization of environmental DNA analysis using pumped deep-sea water for the monitoring of fish biodiversity. Frontiers in Marine Science, 9, 965800.

Bessey, C., Jarman, S. N., Berry, O., Olsen, Y. S., Bunce, M., Simpson, T., ... & Keesing, J. (2020). Maximizing fish detection with eDNA metabarcoding. Environmental DNA, 2(4), 493-504.

Collins, R. A., Bakker, J., Wangensteen, O. S., Soto, A. Z., Corrigan, L., Sims, D. W., ... & Mariani, S. (2019). Non-specific amplification compromises environmental DNA metabarcoding with COI. Methods in Ecology and Evolution, 10(11), 1985-2001.

Schenekar, T., Schletterer, M., Lecaudey, L. A., & Weiss, S. J. (2020). Reference databases, primer choice, and assay sensitivity for environmental metabarcoding: Lessons learnt from a re-evaluation of an eDNA fish assessment in the Volga headwaters. River Research and Applications, 36(7), 1004-1013.

Stauffer, S., Jucker, M., Keggin, T., Marques, V., Andrello, M., Bessudo, S., ... & Waldock, C. (2021). How many replicates to accurately estimate fish biodiversity using environmental DNA on coral reefs?. Ecology and Evolution, 11(21), 14630-14643.

Canals, O., Mendibil, I., Santos, M., Irigoien, X., & Rodríguez-Ezpeleta, N. (2021). Vertical stratification of environmental DNA in the open ocean captures ecological patterns and behavior of deep-sea fishes. Limnology and Oceanography Letters, 6(6), 339-347.

Varela, C., Golightly, C., Timm, L. E., Wilkins, B., Frank, T., Fenolio, D., ... & Bracken-Grissom, H. D. (2021). DNA barcoding enhances large-scale biodiversity initiatives for deep-pelagic crustaceans within the Gulf of Mexico and adjacent waters. The Journal of Crustacean Biology, 41(1), ruab005.

Komai, T., Gotoh, R. O., Sado, T., & Miya, M. (2019). Development of a new set of PCR primers for eDNA metabarcoding decapod crustaceans. Metabarcoding and Metagenomics, 3, e33835.

McCartin, L. J., Vohsen, S. A., Ambrose, S. W., Layden, M., McFadden, C. S., Cordes, E. E., ... & Herrera, S. (2022). Temperature controls eDNA persistence across physicochemical conditions in seawater. Environmental Science & Technology, 56(12), 8629-8639.

Andruszkiewicz Allan, E., 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.

Holman, L. E., Chng, Y., & Rius, M. (2022). How does eDNA decay affect metabarcoding experiments?. Environmental DNA, 4(1), 108-116.

Kumar, G., Reaume, A. M., Farrell, E., & Gaither, M. R. (2022). Comparing eDNA metabarcoding primers for assessing fish communities in a biodiverse estuary. PloS one, 17(6), e0266720.

Zhang, S., Zhao, J., & Yao, M. (2020). A comprehensive and comparative evaluation of primers for metabarcoding eDNA from fish. Methods in Ecology and Evolution, 11(12), 1609-1625.

Muha, T. P., Robinson, C. V., Garcia de Leaniz, C., & Consuegra, S. (2019). An optimised eDNA protocol for detecting fish in lentic and lotic freshwaters using a small water volume. PLoS One, 14(7), e0219218.

Feng, Y., Sun, D., Shao, Q., Fang, C., & Wang, C. (2022). Mesozooplankton biodiversity, vertical assemblages, and diel migration in the western tropical Pacific Ocean revealed by eDNA metabarcoding and morphological methods. Frontiers in Marine Science, 9, 1004410.

Govindarajan, A. F., Llopiz, J. K., Caiger, P. E., Jech, J. M., Lavery, A. C., McMonagle, H., ... & Zhang, W. G. (2023). Assessing mesopelagic fish diversity and diel vertical migration with environmental DNA.

Ramirez-Llodra, E., Brandt, A., Danovaro, R., De Mol, B., Escobar, E., German, C. R., ... & Vecchione, M. (2010). Deep, diverse and definitely different: unique attributes of the world's largest ecosystem. Biogeosciences, 7(9), 2851-2899.

Rees, H. C., Maddison, B. C., Middleditch, D. J., Patmore, J. R., & Gough, K. C. (2014). The detection of aquatic animal species using environmental DNA–a review of eDNA as a survey tool in ecology. Journal of applied ecology, 51(5), 1450-1459.

Weigand, H., Beermann, A. J., Čiampor, F., Costa, F. O., Csabai, Z., Duarte, S., ... & Ekrem, T. (2019). DNA barcode reference libraries for the monitoring of aquatic biota in Europe: Gap-analysis and recommendations for future work. Science of the Total Environment, 678, 499-524.

Bucklin, A., Ortman, B. D., Jennings, R. M., Nigro, L. M., Sweetman, C. J., Copley, N. J., ... & Wiebe, P. H. (2010). A "Rosetta Stone" for metazoan zooplankton: DNA barcode analysis of species diversity of the Sargasso Sea (Northwest Atlantic Ocean). Deep Sea Research Part II: Topical Studies in Oceanography, 57(24-26), 2234-2247.

Iglesias, I.S., Santora, J.A., Fiechter, J. & Field, J.C. (2023). Mesopelagic fishes are important prey for a diversity of predators. Frontiers in Marine Sciences, 10, 10.3389/fmars.2023.1220088

Irigoien, X., Klevjer, T. A., Røstad, A., Martinez, U., Boyra, G., Acuña, J. L., ... & Kaartvedt, S. (2014). Large mesopelagic fishes biomass and trophic efficiency in the open ocean. Nature communications, 5(1), 3271.

Gibert, C., & Escarguel, G. (2019). PER-SIMPER—A new tool for inferring community assembly processes from taxon occurrences. Global Ecology and Biogeography, 28(3), 374-385.

Hsieh, T. C., Ma, K. H., & Chao, A. (2016). iNEXT: an R package for rarefaction and extrapolation of species diversity (H ill numbers). Methods in Ecology and Evolution, 7(12), 1451-1456.

Govindarajan, A. F., Francolini, R. D., Jech, J. M., Lavery, A. C., Llopiz, J. K., Wiebe, P. H., & Zhang, W. (2021). Exploring the use of environmental DNA (eDNA) to detect animal taxa in the mesopelagic zone. Frontiers in Ecology and Evolution, 9, 574877.

Visser, F., Merten, V. J., Bayer, T., Oudejans, M. G., De Jonge, D. S. W., Puebla, O., ... & Hoving, H. J. T. (2021). Deep-sea predator niche segregation revealed by combined cetacean biologging and eDNA analysis of cephalopod prey. Science Advances, 7(14), eabf5908.

Beng, K. C., & Corlett, R. T. (2020). Applications of environmental DNA (eDNA) in ecology and conservation: opportunities, challenges and prospects. Biodiversity and Conservation, 29, 2089-2121.

Barnes, M. A., & Turner, C. R. (2016). The ecology of environmental DNA and implications for conservation genetics. Conservation genetics, 17(1), 1-17.

Stat, M., Huggett, M. J., Bernasconi, R., DiBattista, J. D., Berry, T. E., Newman, S. J., ... & Bunce, M. (2017). Ecosystem biomonitoring with eDNA: metabarcoding across the tree of life in a tropical marine environment. Scientific reports, 7(1), 12240.

Thurber, A. R., Sweetman, A. K., Narayanaswamy, B. E., Jones, D. O., Ingels, J., & Hansman, R. L. (2014). Ecosystem function and services provided by the deep sea. Biogeosciences, 11(14), 3941-3963.

Ross, S. W., Quattrini, A. M., Roa-Varón, A. Y., & McClain, J. P. (2010). Species composition and distributions of mesopelagic fishes over the slope of the north-central Gulf of Mexico. Deep Sea Research Part II: Topical Studies in Oceanography, 57(21-23), 1926-1956.

Yoshitake, K., Yoshinaga, T., Tanaka, C., Mizusawa, N., Reza, M. S., Tsujimoto, A., Kobayashi, A. & Watabe, S. (2019). HaCeD-Seq: a novel method for reliable and easy estimation about the fish population using haplotype count from eDNA. Marine Biotechnology, 21, 813-820.

Cote, D., McClenaghan, B., Desforges, J., Fahner, N. A., Hajibabaei, M., Chawarski, J., Roul, S., Singer, G., Aubry, C, & Geoffroy, M. (2023). Comparing eDNA metabarcoding and conventional pelagic netting to inform biodiversity monitoring in deep ocean environments. ICES Journal of Marine Science, 80(10), 2545-2562.

Burdett, E. A., Fine, C. D., Sutton, T. T., Cook, A. B., & Frank, T. M. (2017). Geographic and depth distributions, ontogeny, and reproductive seasonality of decapod shrimps (Caridea: Oplophoridae) from the northeastern Gulf of Mexico. Bulletin of Marine Science, 93(3), 743-767.