Environmental DNA for detecting marine fish species: testing the volume of sampled water, markers and primers in a public oceanarium

André O. Ferreira[a,b,c,*,†], Cláudia Machado[a,b,†], Cristina Barroso[d,e], Núria Baylina[f], Conceição Egas[d,e], João T. Fontes[a,b], Margarida Fernandes[f], Claudio Oliveira[g], A. Miguel Piecho-Santos[h,i], Sofia Duarte[a,b] & Filipe O. Costa[a,b]

[a] Centre of Molecular and Environmental Biology (CBMA) and ARNET- Aquatic Research Network Associated Laboratory, Department of Biology - University of Minho - Campus de Gualtar, 4710-057 Braga, Portugal
[b] Institute of Science and Innovation for Bio-Sustainability (IB-S), University of Minho - Campus de Gualtar, 4710-057 Braga, Portugal
[c] CIIMAR/CIMAR LA, Interdisciplinary Centre of Marine and Environmental Research, University of Porto, Terminal de Cruzeiros do Porto de Leixões, 4450-208 Matosinhos, Portugal
[d] Center for Neuroscience and Cell Biology (CNC), University of Coimbra, 3004-504 Coimbra, Portugal
[e] Biocant - Transfer Technology Association, 3060-197 Cantanhede, Portugal
[f] Oceanário de Lisboa, 1990-005 Lisboa, Portugal
[g] Institute of Biosciences (IB), São Paulo State University, 18618-698 São Paulo, Brazil
[h] Centre of Marine Sciences (CCMAR), University of Algarve, 8005-139 Faro, Portugal
[i] Portuguese Institute for the Sea and the Atmosphere (IPMA), 1749-077 Lisboa, Portugal

[*]Corresponding author: alof446@gmail.com
[†]André O. Ferreira and Cláudia Machado should be considered joint first authors.

Abstract

To maximize the efficiency of marine fish detection using environmental DNA (eDNA) metabarcoding, methodological factors such as PCR amplification bias and sampled water volume must be carefully considered. In this study, we evaluated the influence of different genetic markers (COI, 12S rRNA, and 16S rRNA) and water sample volumes (0.5, 1.0, 2.5, 5.0, and 10 L) on eDNA-based detection of 60 marine fish species maintained at the *Oceanário de Lisboa*. Five primer sets were used to generate amplicon libraries from triplicate water samples per tested volume. Overall, the combined use of all markers enabled the detection of 81.7% of the fish species present in the tank (n=49), as well as 37 additional exogenous taxa. Detection success varied among primer sets, ranging from 31.7% for COI FishATL_Cocktail1 to 73.3% for Fish16S. When accounting for reference sequence availability, relative detection success reached 82.4% for 12S and

77.2% for 16S. Increasing water volume consistently led to higher species detection across all markers, with detection percentages approaching saturation at the highest volumes tested. When pooling all markers, up to 76.7% of the species present were detected using a water volume of 5 L. These results highlight the limitations of relying on a single genetic marker and demonstrate the strong influence of sampling volume on detection efficiency, a factor likely to be even more critical in open marine environments.

Keywords: DNA metabarcoding, environmental DNA, biomonitoring, fish detection, controlled system

Introduction

Environmental DNA (eDNA) metabarcoding has emerged as a powerful and non-invasive approach for species detection and biodiversity assessment in marine ecosystems (Taberlet *et al.*, 2012; Yamanaka and Minamoto, 2016). Environmental DNA refers to genetic material released by organisms into their surroundings, such as water, sediment, or air, through processes including excretion, reproduction, and tissue degradation (Bruce *et al.*, 2021). By coupling eDNA sampling with high-throughput sequencing (HTS) of short genetic markers, metabarcoding enables the simultaneous detection of multiple taxa from a single environmental sample (Hajibabaei *et al.*, 2011; Taberlet *et al.*, 2012). This represents a substantial advance over traditional morphology-based survey methods (e.g., trawling or net sampling), which are often invasive, labor-intensive, and limited by taxonomic expertise (Valentini *et al.*, 2016; Fraija-Fernández *et al.*, 2020; Gold *et al.*, 2021a). Accordingly, eDNA metabarcoding has been increasingly applied across a wide range of marine environments, demonstrating considerable potential for biodiversity monitoring and conservation (Bucklin *et al.*, 2021; Fediajevaite *et al.*, 2021; Garcia-Vazquez *et al.*, 2021; Gold *et al.*, 2021a; Carvalho *et al.*, 2024).

Among genetic markers, the mitochondrial cytochrome c oxidase subunit I (COI) gene has become the standard for animal DNA barcoding due to its high interspecific variability and extensive taxonomic coverage (Hebert *et al.*, 2003; Ratnasingham and Hebert, 2013; Zhang *et al.*, 2020). COI has been widely used for marine fish identification and benefits from comparatively comprehensive reference databases that are crucial for precise taxonomic assignments (Ward *et al.*, 2005; Costa *et al.*, 2012; Weigand *et al.*, 2019). However, COI primers often suffer from non-specific amplification, particularly of bacterial and other non-target DNA, which can substantially reduce the proportion of

informative fish reads in eDNA metabarcoding datasets (Deagle *et al.*, 2014; Collins *et al.*, 2019). Recent efforts have focused on improving COI primer specificity and performance, yet amplification bias remains a major limitation (Weigand *et al.*, 2019; Schroeder *et al.*, 2020; Zhang *et al.*, 2020; Costa *et al.*, 2025).

To address these limitations, mitochondrial ribosomal RNA markers, particularly 12S rRNA and 16S rRNA, have been increasingly adopted for fish eDNA metabarcoding (Miya *et al.*, 2015; Berry *et al.*, 2017; DiBattista *et al.*, 2017; Ratcliffe *et al.*, 2021). These markers stand out due to their conserved nature, which allows an effective amplification and species-level identification (Valentini *et al.*, 2016). Nevertheless, their application is constrained by comparatively incomplete reference libraries and, in some cases, limited species-level resolution, especially among closely related taxa (Garcia-Vazquez *et al.*, 2021; Kumar *et al.*, 2022; Fontes *et al.*, 2024). Ongoing improvements in sequencing technologies and bioinformatic pipelines are gradually enhancing reference coverage and taxonomic assignment accuracy for these markers (Ushio et al., 2018).

Beyond marker choice, the performance of eDNA metabarcoding is influenced by multiple methodological and environmental factors, including DNA degradation and hydrodynamic transport (Thomsen *et al.*, 2012; Gold *et al.*, 2021a). Sampling design is particularly critical, as water volume directly affects eDNA concentration and detection probability (Bessey *et al.*, 2020). While larger volumes generally increase species detection, this relationship often reaches a point of diminishing returns, resulting in marker and system-specific saturation thresholds (Sigsgaard *et al.*, 2015; Evans *et al.*, 2017). Identifying optimal sampling volumes for different genetic markers is therefore essential for maximizing detection efficiency while maintaining feasibility and cost-effectiveness (Deiner *et al.*, 2017; Bruce *et al.*, 2021). Controlled environments such as aquaria provide an ideal framework for disentangling the effects of marker choice, primer performance, and sampling volume on eDNA metabarcoding outcomes. These systems minimize environmental variability, allow precise knowledge of species composition, and enable robust evaluation of methodological performance under replicable conditions (Kelly *et al.*, 2014; Evans *et al.*, 2017; Morey *et al.*, 2020).

In this study, we evaluate and optimize eDNA metabarcoding strategies for marine fish detection using the *Oceanário de Lisboa* as a controlled experimental system. Specifically, we assess (i) the relative performance of three commonly used mitochondrial markers (COI, 12S rRNA, and 16S rRNA) and their associated primer sets, and (ii) the influence of increasing water sampling volumes (0.5, 1.0, 2.5, 5.0, and 10.0 L) on species

detection efficiency. By comparing marker-specific detection success, primer bias, and volume-dependent saturation patterns, we aim to identify optimal methodological combinations for reliable marine fish monitoring. Ultimately, this work seeks to provide empirically grounded recommendations for eDNA metabarcoding survey design, highlighting the benefits of a multi-marker approach and appropriate sampling volumes to improve species detection in marine environments.

Methods

Study location

The study was conducted at the *Oceanário de Lisboa* (Portugal), which houses a 5-million-liter main aquarium inhabited by 60 marine fish species from 45 genera, including 44 Actinopterygii and 16 Elasmobranchii (***Supplementary Material, Table S1***). The seawater used in the aquarium is artificially generated on-site, minimizing the risk of external eDNA inputs. Tap water is first filtered through activated carbon to remove impurities, after which marine salt is added to produce artificial seawater. Once target physicochemical parameters are reached, the water is transferred to a dedicated reservoir and subsequently supplied to the aquarium, each of which operates with an independent filtration system.

Prior to entering the tank, seawater undergoes both mechanical filtration through sand filters and biological treatment using a degas tower. Water disinfection is performed using ozone, and the system is chlorine-free. The aquarium is maintained at a target temperature of 22 °C (± 0.3 °C), with salinity ranging from 32 to 33 and pH values between 8.10 and 8.20. Complete water turnover occurs approximately every two hours, ensuring continuous filtration of the entire volume. Because the seawater is artificially prepared, sterilized, and continuously filtered prior to entering the aquarium, the introduction of exogenous eDNA from natural marine environments is not expected. Consequently, the detected eDNA signal can be confidently attributed to the species present within the aquarium, including those used for feeding, making this system particularly suitable for evaluating marker performance and sampling strategies under controlled conditions.

Sample collection

Three water samples were collected in June 2021 from the center of the main tank at approximately 3.5 m depth, using a 20 L plastic container that had been sterilized with a 10% bleach solution and thoroughly rinsed with ultrapure water prior to sampling. Before

each aliquot was withdrawn, the water in the container was manually agitated to homogenize the sample and promote the even suspension of eDNA-containing particles. From each 20 L sample, aliquots were immediately filtered at different volumes (0.5, 1.0, 2.5, 5.0, and 10.0 L) through 0.45-µm glass microfiber filters (47 mm diameter; Merck Millipore) using an EZ-Fit™ Manifold filtration system equipped with three Microfil® funnels connected to an EZ-Stream vacuum filtration pump (Merck Millipore, Darmstadt, Germany). Each filtration unit corresponded to a single biological replicate.

All filtration equipment was decontaminated before use by soaking in a 10% bleach solution for at least 30 min. A 0.5 L ultrapure water sample was included as a negative control to monitor potential cross-contamination during sample processing. Filtration proceeded sequentially from the negative control to increasing sample volumes. Although the filtration manifold itself was not cleaned between volumes, the funnels and all surfaces that came into contact with the filters or unfiltered sample water were decontaminated with bleach followed by ethanol before filtering each subsequent volume. Following filtration, filters were placed into sterile plastic tubes and stored at −20 °C until DNA extraction.

DNA extraction

DNA was extracted from filters using the DNeasy® PowerSoil® Kit (QIAGEN), which has been shown to efficiently recover fish eDNA from filter paper (Eichmiller, Miller and Sorensen, 2016). Extractions were performed within a UV-sterilized DNA extraction workstation (30 min exposure) and all surfaces were cleaned with a 10% bleach solution prior to use to minimize contamination. To monitor potential contamination during extraction, negative controls consisting of new filters were included to check for DNA in kit reagents or laboratory consumables. Each filter was halved, and each half was extracted separately, yielding a total of 34 extractions from aquarium and control samples. Extractions followed the manufacturer's protocol, with DNA eluted to a final volume of 60 µL. The two technical replicates per sample were subsequently pooled (120 µL total), and DNA concentrations were assessed using a Thermo Scientific™ NanoDrop™ One/OneC spectrophotometer (Thermo Fisher Scientific, Waltham, MA, USA) solely to provide operational estimates for standardizing the total DNA input into the target-specific PCR, rather than as absolute measurements of amplifiable double-stranded DNA. All DNA extracts were stored at -20 °C until PCR amplification and high-throughput sequencing.

PCR Amplification and High-Throughput Sequencing

Samples were prepared for Illumina MiSeq sequencing at Genoinseq (Biocant, Cantanhede, Portugal) using previously validated PCR primers for COI, 12S rRNA, and 16S rRNA genes, which have been shown to successfully amplify marine fish DNA (***Table 1***). The selection of the three genetic markers and their respective primers was based on their specificity, taxonomic breadth, fragment size, and documented amplification success: 2 fish specific primer cocktails (FishATL_Cocktail1 and FishATL_Cocktail2) (Costa *et al.*, 2025) and the generalist primer pair mlCOIintF/LoboR1 for COI (Leray *et al.*, 2013; Lobo *et al.*, 2013); the cocktail MiFish-U/E for 12S rRNA (Miya *et al.*, 2015); and Fish16sFD/16s2R for 16S rRNA (Berry *et al.*, 2017).

For each marker, two PCR reaction replicates were performed for each sample, using the KAPA HiFi HotStart PCR Kit (Kapa Biosystems, Cape Town, South Africa) according to the manufacturer's suggestions. To amplify the COI region, three primer sets were used: mlCOIintF and reverse primer LoboR1 (Leray *et al.*, 2013; Lobo *et al.*, 2013) and two primer cocktails, FishATL_Cocktail1 and FishATL_Cocktail2 (Costa *et al.*, 2025). In the former, 0.3 µM of each PCR primer and 5 µL of template DNA, were used in a total volume of 25 µL. The PCR conditions involved a 3 min denaturation at 95 ºC, followed by 35 cycles of 98 ºC for 20 s, 60 ºC for 30 s, and 72 ºC for 30 s, and a final extension at 72 ºC for 5 min. For the primer cocktails, each contained 3 µM of the following primers: FishATL_C1: FishF1, FishF2, FishF6, and FishATL_C2: 219, 217F, 214F, for the forward, and FishATL_C1: 433R, 432R, 434R, and FishATL_C2: FishR1, FishR2, FishR7, for the reverse primers, respectively, and 1 µL of template DNA in a total volume of 25 µL. The PCR conditions involved a 3 min denaturation at 95 ºC, followed by 5 cycles of 95 ºC for 30 s, 56 ºC for 60 s and 72 ºC for 60 s, 30 cycles of 95 ºC for 30 s, 60 ºC for 60 s and 72 ºC for 60s, and a final extension at 72 ºC for 10min.

The 12S rRNA gene used 0.6 µM of each PCR primer: a pool of forward primers MiFish-U-F and MiFish-E-F and a pool of reverse primers MiFish-U-R and MiFish-E-R (Miya *et al.*, 2015) and 2.5 µL of template DNA in a total volume of 25 µL. The PCR conditions involved a 3 min denaturation at 95 ºC, followed by 38 cycles of 98 ºC for 20 s, 60 ºC for 30 s, and 72 ºC for 30 s, and a final extension at 72 ºC for 5 min.

The fish 16S mitochondrial gene was amplified using 0.3 µM of each PCR primer: forward primer 16SF/D and reverse primer 16s2R-deg (Berry *et al.*, 2017) and 5 µL of template DNA in a total volume of 25 µL. The PCR conditions involved a 3 min

denaturation at 95 ºC, followed by 45 cycles of 98 ºC for 20 s, 54 ºC for 30 s, and 72 ºC for 30 s, and a final extension at 72 ºC for 5 min.

Replicates for each sample and marker were pooled prior to library preparation. Negative controls showed no amplification for any primer set. Libraries were then reamplified in a limited-cycle PCR to add Illumina adapters and dual indexes. The PCR conditions consisted of an initial denaturation at 95 ºC for 3 min, followed by 8 cycles of 95 ºC for 30 s, 55 ºC for 30 s and 72 ºC for 30 s, and a final extension at 72 ºC for 5 min. The resulting products were then purified, normalized using the SequalPrep 96-well kit (Thermo Fisher Scientific), pooled, and sequenced using 250 bp paired-end reads on the Illumina MiSeq platform (MiSeq Reagent Kit v3, 600 cycles, ~100,000 reads per sample), following the manufacturer's protocols.

Table 1. Primer sequences used for the amplification of genetic markers COI, 12S rRNA, and 16S rRNA. (F) - forward; (R) - reverse; bp - base pairs.

Primer set (Abbr.) and length	Direction (5'-3')	Reference
mlCOIintF/LoboR1 (Leray-Lobo) COI (~313 bp)	**(F) mlCOIintF**: GGWACWGGWTGAACWGTWTAYCCYCC	(Leray *et al.*, 2013)
	(R) LoboR1: TAAACYTCWGGRTGWCCRAARAAYCA	(Lobo *et al.*, 2013)
FishATL_Cocktail1 (FishATL_C1) COI (~400 bp)	**(F) FishF1:** TCAACCAACCACAAAGACATTGGCAC	(Ward *et al.*, 2005)
	(F) FishF2: TCGACTAATCATAAAGATATCGGCAC	(Ward *et al.*, 2005)
	(F) FishF6: ACYAAYCACAAAGAYATTGGCA	(Jennings *et al.*, 2019)
	(R) 433R: AGCCCCTAGAATTGAGGARACHC	(Costa *et al.*, 2025)
	(R) 432R: CCCGACATAGYATTCCCHC	(Costa *et al.*, 2025)
	(R) 434R: CCCCTAGAATTGAGRAAACHCC	(Costa *et al.*, 2025)
FishATL_Cocktail2 (FishATL_C2) COI (~400 bp)	**(F) 219F:** CCCGACATRGCATTCCCYC	(Costa *et al.*, 2025)
	(F) 217F: CCCCCGACATRGCMTTCCC	(Costa *et al.*, 2025)
	(F) 214F: GDGCCCCCGACATAGCATTY	(Costa *et al.*, 2025)
	(R) Fish R1: TAGACTTCTGGGTGGCCAAAGAATCA	(Ward *et al.*, 2005)
	(R) Fish R2: ACTTCAGGGTGACCGAAGAATCAGAA	(Ward *et al.*, 2005)
	(R) Fish R7: TARACTTCTGGRTGDCCRAAGAAYCA	(Jennings *et al.*, 2019)
MiFish-U/E (MiFish) 12S rRNA (~170 bp)	**(F) MiFish-U-F:** GTCGGTAAAACTCGTGCCAGC	(Miya *et al.*, 2015)
	(F) MiFish-E-F: GTTGGTAAATCTCGTGCCAGC	
	(R) MiFish-U-R: CATAGTGGGGTATCTAATCCCAGTTTG	
	(R) MiFish-E-R: CATAGTGGGGTATCTAATCCTAGTTTG	
Fish 16S (Fish16S) 16S rRNA (~200 bp)	**(F) 16SF/D:** GACCCTATGGAGCTTTAGAC	(Berry *et al.*, 2017)
	(R) 16s2R-deg: CGCTGTTATCCCTADRGTAACT	(Deagle *et al.*, 2007)

Bioinformatic analysis

Raw reads generated through high-throughput sequencing were compiled in fastq files and processed using PMiFish version 2.4.1 (available at

https://github.com/rogotoh/PMiFish.git; Miya *et al.*, 2020). This pipeline was chosen given its ability to analyze data from different genetic markers with a standardized workflow, while allowing adjustments to several settings at different steps of the analysis. It follows the structure of conventional metabarcoding data processing tools, performing sequential tasks using USEARCH version 10.0.240 (Edgar, 2010) for read processing. The workflow includes paired-end read merging, primer removal, quality filtering, dereplication, denoising, and taxonomic assignment. The specific parameters established allowed 3 mismatches within the alignment region during the merging step; quality filtering discarded sequences shorter than 50 bp compared with the expected amplicon length for the respective primer set; dereplication combined identical sequences, while the denoising step generated amplicon sequence variants (ASVs) by eliminating chimeras and erroneous sequences. All ASVs were taxonomically assigned using a similarity threshold equal to or greater than 97% against the respective reference databases, with sequences returning no match being excluded from downstream analyses. Finally, the pipeline creates molecular operational taxonomic units (MOTUs) that are represented by the ASV with the highest read count within the grouped sequences.

Reference databases for the three genetic markers were compiled specifically for this case study and focused on fish species, prioritizing assignment accuracy despite potentially increasing the proportion of unassigned sequences (Gold *et al.*, 2021b; Mugnai *et al.*, 2023). For COI, records were retrieved from the BOLD database (Ratnasingham *et al.*, 2024) using the BAGS software (https://github.com/tadeu95/BAGS; Fontes *et al.*, 2021), which also provides the Barcode Index Number (BIN; Ratnasingham and Hebert, 2013) associated with each sequence. For the 12S and 16S markers, reference sequences were obtained from NCBI GenBank (Sayers *et al.*, 2025), including both marker-specific sequences and complete mitochondrial genomes. For both BOLD and GenBank, all available records assigned to Actinopterygii, Elasmobranchii, Myxini, Holocephali and Petromyzontiformes were mined, and GenBank entries lacking accuracy verification by GenBank staff or without species-level identification were excluded. In both cases, species names were cross-referenced with the World Register of Marine Species (WoRMS Editorial Board, 2025) database to retain only marine or brackish taxa.

Several curation procedures were undertaken to achieve the most precise species-level identifications and increase the reliability of the employed reference databases. All taxonomic names were verified against WoRMS to rectify synonyms and unaccepted names both during the compilation of the reference libraries and after the final taxonomic

assignment. In addition, since every COI record had an associated BIN, this information was double-checked with BOLD to detect potential misidentifications or cases of poor species discrimination where multiple taxa share the same BIN. Finally, when the taxonomic assignment matched more than one closely related species, the most probable species and consequent correct species name were selected based on the official species list of the aquarium, when applicable. In cases where ambiguity prevailed, additional genetic information from the other genetic markers was used to refine the identifications, applying a cross-validation among metabarcoding datasets.

Based on documented data from previous water eDNA studies and our personal experience, no minimum-read threshold was applied in the main analysis to consider a record valid, due to the generally low read counts associated with multiple valid detections obtained using environmental DNA. In addition to the main analysis, a sensitivity analysis was conducted by applying minimum thresholds of five and ten reads. Species richness obtained under each criterion was compared separately for Oceanarium-listed and exogenous taxa to assess the trade-off between reducing potentially spurious low-read detections and retaining genuine detections of species known to be present in the Oceanarium.

Data analysis

Venn diagrams were generated using the InteractiVenn web tool (https://interactivenn.net/; Heberle *et al.*, 2015) to evaluate the overlap and exclusivity of fish species detections among the different primer pairs, for the total of fish species and for the total of fish species listed in Oceanarium. The resulting diagrams were refined and edited using Inkscape 1.2 (Inkscape Project, 2020). Bar plots illustrating absolute and relative detections were produced using the ggplot2 package (Wickham, 2016) in RStudio (version 4.0.4; RStudio Team, 2020). Scatter plots depicting the relationship between sample volume and the number of detected fish species were generated within the same environment. Data manipulation and visualization workflows were supported by the "tidyr" (Wickham *et al.*, 2024) and "dplyr" (Wickham *et al.*, 2023) packages, enabling efficient data processing and preparation for graphical representation.

Results

Initial HTS results

The total number of initial raw reads obtained across all samples ranged from 804,379 to 1,998,935 for the Fish16S and the mlCOIintF/LoboR1 primer sets, corresponding to 86.1% and 83.8% of usable reads, respectively (***Table 2, Supplementary Material Table S2***). The lowest proportion of usable reads were observed for FishATL_C1 and FishATL_C2, with only 24.5% of and 39.2% of the initial reads being retained. Overall, the proportion of reads assigned to fish species (>97% identity) was very low compared to the original raw and usable reads, reaching its maximum for the 12S primer set at 6.2% of usable reads, while the two COI fish-specific cocktails registered percentages below 1%. Interestingly, the generalist COI primer set had the highest number of fish-matching reads (65,196), followed by the 12S and 16S primer sets, while the COI fish-specific cocktails (FishATL_C2 and FishATL_C1) produced the lowest, matching their lower performance in the previous metric (***Table 2***). Exploratory BLASTn screening showed that most off-target matches were bacterial, particularly Pseudomonadota, with only one human-associated sequence detected across all primer sets. mlCOIintF/LoboR1 detected a wider variety of non-target bacterial and eukaryotic groups, while the COI cocktails exhibited higher percentages of sequences without GenBank matches than the other primer sets, particularly FishATL_C1, which may be related to their degenerate and multiplexed design.

Table 2. Total number of reads generated in Illumina MiSeq high-throughput sequencing (raw reads), the reads retained after all the processing steps of the bioinformatics pipeline (usable reads), and the reads assigned to fish species, for each primer used. * reads submitted to taxonomic assignment; **reads with taxonomic classification at the species level.

	COI			**12S**	**16S**
	mlCOIintF /LoboR1	FishATL _C1	FishATL _C2	MiFish-U/E	Fish16S
Raw reads	1,998,935 (100%)	1,308,603 (100%)	1,512,115 (100%)	917,663 (100%)	804,379 (100%)
Usable reads* (% vs. Raw reads)	1,675,764 (83.83%)	320,386 (24.48%)	592,697 (39.20%)	781,929 (85.21%)	692,312 (86.07%)
Fish reads** (% vs. Raw reads) (% vs. Usable reads)	65,196 (3.26%) (3.89%)	580 (0.04%) (0.18%)	2,869 (0.19%) (0.48%)	48,678 (5.30%) (6.23%)	21,186 (2.63%) (3.06%)

Oceanarium Fish Species Detection

A total of 86 fish species were detected in the Oceanarium, when considering results from all genetic markers (***Supplementary Material, Table S3***). From the official list of 60 fish species provided by the *Oceanário de Lisboa* , 49 species were identified via eDNA metabarcoding, corresponding to 81.7% of the total diversity. At higher taxonomic levels, eDNA metabarcoding detected 39 of 45 genera, 24 of 26 families, and 14 of 15 orders present in the official species list, demonstrating the method's broad taxonomic coverage. An assessment of the efficiency of each primer set in detecting the species present in the aquarium (***Figure 1***), showed that Fish16S was the most effective primer pair, detecting 44 species. This was closely followed by MiFish-U/E and mlCOIintF/LoboR1 primer pairs which detected 42 and 41 species, respectively. In contrast, the COI fish-specific cocktails showed lower performance: FishATL_C2 detected 32 species, whereas FishATL_C1 detected only 19. Notably, three species were uniquely detected by individual primer pairs: *Diplodus cervinus* with Fish16S, *Gymnothorax funebris* with 12S MiFish-U/E, and *Pomacanthus paru* with FishATL_Cocktail1.

Considering the primer set with the highest detection rate, Fish16S, five species remained undetected beyond the eleven species not recovered by any primer set, corresponding to a detection deficit of approximately 10% of the total recovered diversity. For the 12S MiFish-U/E primers, 7 additional species would be missed (14% deficit), followed by COI Leray-Lobo, which would fail to detect eight species (16% deficit). In contrast, the COI fish-specific cocktails showed substantially poorer performance: FishATL_Cocktail2 would miss 17 additional species (35% deficit) while FishATL_Cocktail1 would fail to identify 30 species (61% deficit), in a highly incomplete representation of the true fish diversity present in the aquarium. Overall, these results highlight that no single primer set fully captures species diversity, and that the combined use of COI, 12S, and 16S markers provides a more comprehensive and reliable assessment of aquarium biodiversity.

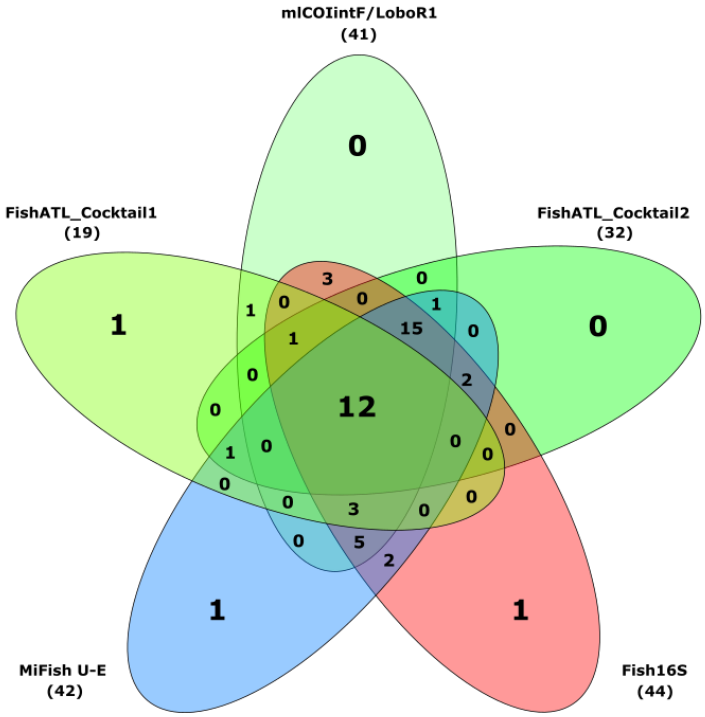

Figure 1. Partitioning of fish species detected by each primer set belonging to the official list of the *Oceanário de Lisboa* (49 species in total). The number in parentheses corresponds to the total number of fish species from the official list identified by each primer set.

The results presented above reflect the absolute detection success of fish species present in the Oceanarium, assuming that all taxa are represented in the reference databases used for taxonomic assignment of the quality-filtered reads. For the COI marker, this assumption holds true, as all 60 fish species from the Oceanarium inventory were represented in the reference database. In contrast, examination of the reference libraries for the remaining genetic markers revealed gaps in taxonomic coverage. Specifically, the 12S database lacked reference sequences for nine species present in the Oceanarium (*Aetomylaeus bovinus, Anisotremus virginicus, Argyrosomus regius, Caranx ruber, Diplodus vulgaris, Himantura undulata, Myliobatis aquila, Rhinobatos rhinobatos,* and *Taeniurops grabatus*) while the 16S database was missing sequences for three species (*Aetomylaeus bovinus*, *Himantura undulata*, and *Rhinobatos rhinobatos*). Consequently, detection of these species using the corresponding markers would be impossible, regardless of primer performance (***Supplementary Material, Table S4***).

Accounting for these differences in reference database completeness allows the estimation of relative detection success, defined as the proportion of detectable species given the available reference sequences. Under this framework, the detection rate of the 12S MiFish-U/E primers increased to 82.4% of the available diversity (42 species of the 51 from the Oceanarium list in the 12S database), slightly surpassing Fish16S, which reached a relative detection rate of 77.2% (44 of the 57 species represented in the 16S database) (***Figure 2***). In contrast, both absolute and relative detection percentages with the COI primer sets remained unchanged, reflecting the completeness of the COI reference database for the fish species inhabiting the Oceanarium.

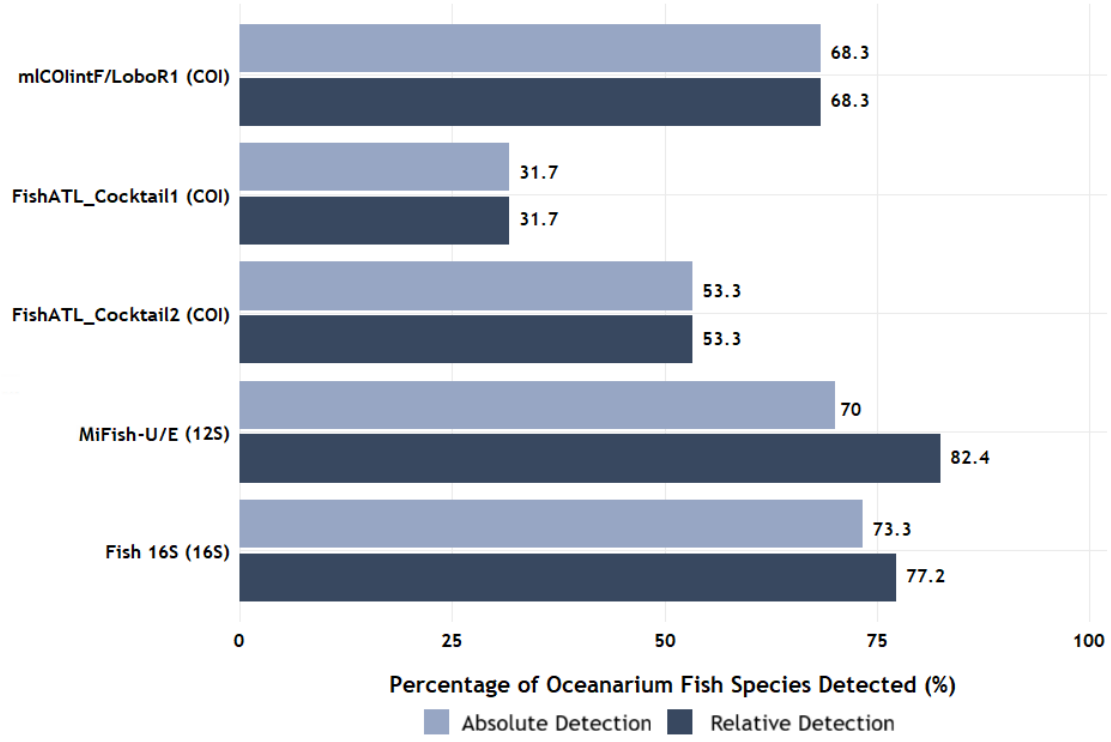

Figure 2. Percentage of species detected by each primer pair, shown as absolute values and as relative values considering only species represented in the corresponding genetic databases.

Detection of Exogenous Taxa

In addition to the 49 fish species from the Oceanarium's official inventory that were detected, eDNA metabarcoding revealed a further 37 fish species presumed to be absent from the closed system. Overall, detections comprised 24 orders, 40 families, and 65 genera. Among the COI primer sets, the Leray-Lobo primer pair detected the highest number of species (58 species from 47 genera), followed by FishATL_C2 (38 species

from 31 genera) and FishATL_C1 (23 species from 20 genera) (***Figure 3; Supplementary Material Table S5***).

When results from the three COI primer sets were combined, a total of 63 fish species belonging to 51 genera were detected, although only 16 species were shared across all three primer sets (***Supplementary Material, Figure S1***). The generalist COI primer combination mlCOIintF/LoboR1 outperformed the fish-specific cocktails, uniquely identifying 19 fish taxa and failing to detect only five species (*Argyrosomus regius*, *Carcharias taurus*, *Epinephelus lanceolatus*, *Pomacanthus paru*, and *Sebastes mentella*). FishATL_C1 and FishATL_C2 each detected two exclusive species and shared one additional exclusive detection. Interestingly, the primer sets targeting the 12S and 16S markers each detected 63 species, matching the combined performance of the COI primer sets (***Supplementary Material, Figure S2***), and representing 51 and 49 genera, respectively, with MiFish-U/E and Fish16S yielding nine and seven exclusive species detections (***Figure 3***). Despite incompleteness in the 12S and 16S reference libraries, both markers provided unique and complementary information.

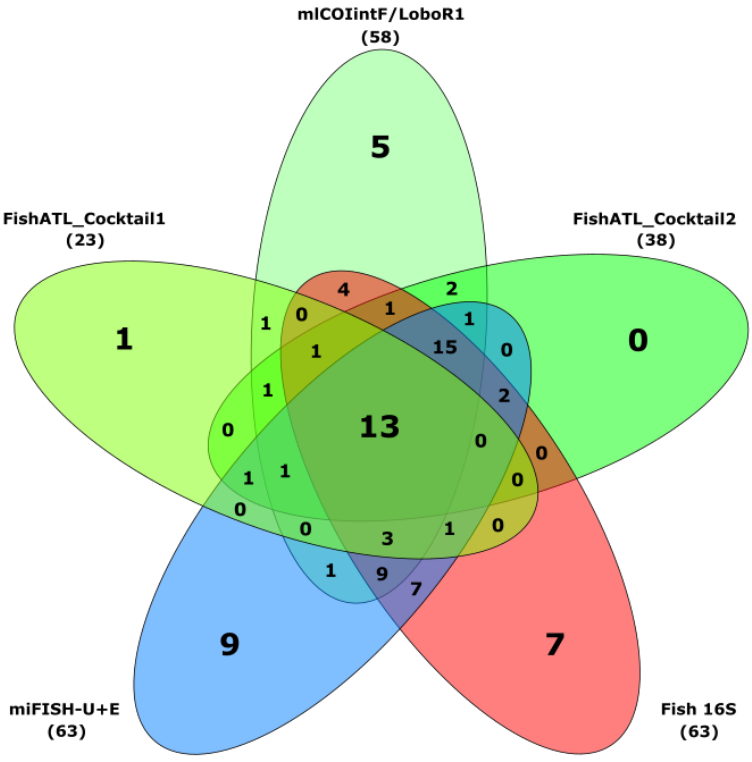

Figure 3. Partitioning of fish species detected by each primer set across all collected samples (86 species in total). The number in parentheses corresponds to the total number of fish species identified by each primer set.

Species Detection under Different Minimum-Read Thresholds

After conducting our main analysis without a minimum-read threshold, given the typically low read counts associated with species detections in eDNA studies and the potential non-detection of Oceanarium-listed species, a sensitivity analysis was performed to assess the trade-off resulting from this decision and its impact on retaining genuine detections of species known to be present in the Oceanarium while reducing potentially spurious low-read detections. Increasing the minimum-read threshold reduced species richness for all primer sets, although its effect varied among primers and markers (***Table 3;*** detailed results provided in ***Supplementary Material, Table S6)***.

When detections from all primer sets were combined, the minimum 5-read threshold reduced cumulative richness from 86 to 79 species, excluding two of the 49 Oceanarium-listed species (4.1%) and five of the 37 exogenous taxa (13.5%). The minimum 10-read threshold retained 74 species, excluding six listed species (12.2%) and six exogenous taxa (16.2%). Although the ≥5-read threshold removed proportionally more exogenous than listed taxa, increasing the threshold from five to ten reads excluded four additional Oceanarium-listed species but only one additional exogenous taxon, resulting in the exclusion of the same number of taxa (n = 6) from each category compared with the main analysis without a minimum-read criterion.

Table 3. Effect of minimum-read thresholds on fish species richness detected by each primer set and by their combined results. Values in parentheses indicate Oceanarium-listed species (Oce.) and exogenous taxa (Exo.), respectively. Species losses were calculated relative to the main analysis conducted without a minimum-read threshold.

Primer set	No Min. Reads	≥5 Reads		≥10 Reads	
	Total *(Oce./Exo.)*	**Total** *(Oce./Exo.)*	**Excluded** *(Oce./Exo.)*	**Total** *(Oce./Exo.)*	**Excluded** *(Oce./Exo.)*
All primer sets	**86** *(49/37)*	**79** *(47/32)*	**7** *(2/5)*	**74** *(43/31)*	**12** *(6/6)*
Cumulative COI	**63** *(45/18)*	**56** *(42/14)*	**7** *(3/4)*	**48** *(34/14)*	**15** *(11/4)*
mlCOIintF/LoboR1	**58** *(41/17)*	**51** *(37/14)*	**7** *(4/3)*	**43** *(29/14)*	**15** *(12/3)*
FishATL_C1	**23** *(19/4)*	**6** *(5/1)*	**17** *(14/3)*	**3** *(3/0)*	**20** *(16/4)*
FishATL_C2	**38** *(32/6)*	**30** *(26/4)*	**8** *(6/2)*	**21** *(17/4)*	**17** *(15/2)*
12S MiFish-U/E	**63** *(42/21)*	**56** *(37/19)*	**7** *(5/2)*	**54** *(35/19)*	**9** *(7/2)*
Fish16S	**63** *(44/19)*	**49** *(39/10)*	**14** *(5/9)*	**37** *(33/4)*	**26** *(11/15)*

The impact of these thresholds was more pronounced when considering the results of individual primer sets. Total species losses ranged from 7 to 17 under the five-read minimum criterion and from 9 to 26 under the ten-read minimum criterion.

FishATL_Cocktail1 was particularly sensitive to this filtering, with richness decreasing from 23 species without a threshold to 6 and 3 species under the 5 and 10 read minimum criteria, respectively. In contrast, Fish16S showed the greatest reduction in exogenous detections, which decreased from 19 without filtering to 10 with a minimum of five reads and 4 with a minimum of ten reads, while the corresponding number of Oceanarium-listed species decreased from 44 to 39 and 33. Nevertheless, most exogenous taxa not included in the official *Oceanário de Lisboa* species list remained detectable after the cumulative integration of all primer sets, with 32 of 37 retained using the 5-read filter and 31 retained using the 10-read filter (***Supplementary Material, Table S6***).

Influence of Water Volume in Species Detection

To evaluate the effect of sampled water volume on fish species detection within the Oceanarium, identifications from all primer sets were pooled for each volume category. As expected, increasing water volume resulted in a higher overall number of detected fish species (***Figure 4A***). Samples with the lowest volume (0.5 L) yielded 44 species detections, of which 24 corresponded to species listed in the Oceanarium inventory. Progressive increases in recovered diversity were observed with larger volumes: 1 L samples identified 45 species (29 listed species), 2.5 L identified 53 species (39 listed species), 5 L identified 64 species (46 listed species), and the highest volume, 10 L, identified 68 species (47 listed species).

Considering only Oceanarium-listed species, 25 additional species were recovered beyond the 0.5 L volume: eight were first detected at 1 L, nine at 2.5 L, seven at 5 L, and only one at 10 L (***Supplementary Material, Table S7***). Species first detected at 2.5 and 5 L were predominantly reef-associated fishes, although the additional detections at 2.5 L also included coastal pelagic species. At 10 L, the only additional species detected was the sparid *Diplodus cervinus*. Curiously, detection patterns were not entirely cumulative: three species detected at 0.5 L were not recovered at 1 L, while at each subsequent volume, two species previously detected at lower volumes were not recovered (***Supplementary Material, Table S7***).

Notably, samples of 5.0 L recovered up to 74.4% of the total species richness, a result that closely approaches the 79.1% recovery achieved with 10 L samples. A similar pattern was observed when considering only Oceanarium-listed species detected across different volumes, with 5.0 L samples recovering 76.7% of the aquarium diversity compared to 78.3% with 10 L (***Figure 4B***). Overall, increased sample volume led to higher species

richness for all primer sets. The only exception to this general trend was observed for FishATL_Cocktail1, with 10 Oceanarium-listed species detected at 2.5 L and 9 at 5 L.

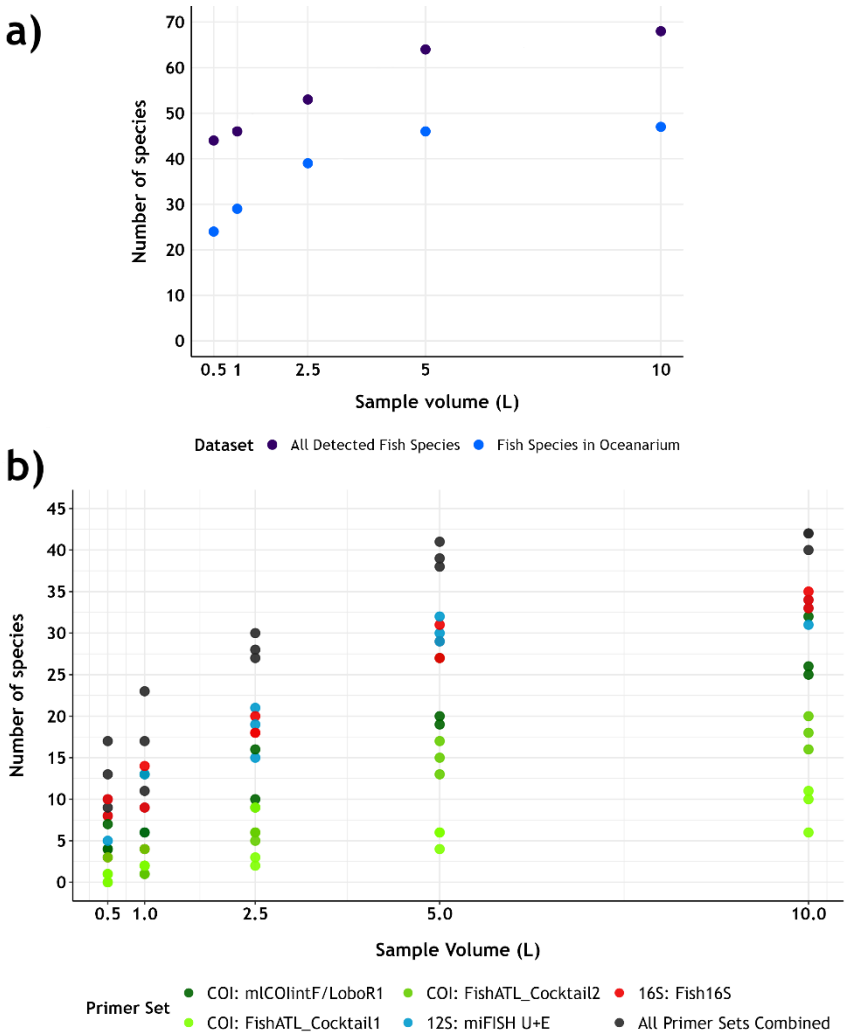

Figure 4. Effect of sampled water volume on fish species detection. (A) Cumulative number of unique fish species detected at each filtration volume after pooling detections across all primer sets and the three independent water samples, considering both all detected species and species listed in the Oceanarium's official inventory. (B) Number of Oceanarium-listed species detected by each primer set at each filtration volume. Each point represents one of the three independently collected water samples and its corresponding filtered aliquot.

Discussion

This study provides valuable insights into the optimization of eDNA metabarcoding protocols for marine fish detection by evaluating how genetic markers and water volumes influence species recovery in a controlled but phylogenetically diverse setting. Our results show that combining multiple markers and increasing water volume improves detection percentages, but practical and methodological constraints still limit full species recovery.

These findings support ongoing efforts toward the improvement of eDNA-based fish monitoring, particularly in complex open marine environments.

Marker choice strongly influences fish detection success

Our results clearly demonstrate that marker choice strongly influences fish species detection in eDNA metabarcoding surveys. Among individual primer sets, Fish16S showed the highest absolute detection success, recovering 73.3% of the fish diversity present in the Oceanarium, followed closely by the 12S MiFish-U/E primers (70%). When COI primer sets were considered individually, their performance was generally lower; however, when the three COI primer sets were combined, COI yielded the highest overall number of detected species (45 species, 75%). This result reflects the cumulative effect of integrating multiple primer combinations, rather than superior performance of a single COI primer pair. This pattern is consistent with previous studies showing that COI offers broad taxonomic resolution but is prone to amplification bias due to high sequence variability and primer mismatches across taxa (Collins *et al.*, 2019). Thus, while COI can substantially contribute to overall diversity recovery, its effectiveness depends strongly on primer design and on combining multiple primer sets to widen amplification breadth. Across all water volumes tested, no single primer pair or marker recovered more than ~82% of the fish species listed in the Oceanarium, even when accounting for reference database completeness. Although this represents a relatively high detection success in a controlled system, detection efficiency is likely to be lower in open natural environments, where eDNA dilution, transport, and degradation further constrain detectability (Barnes *et al.*, 2014; Deiner *et al.*, 2016; Goldberg *et al.*, 2016; Collins *et al.*, 2018).

Complementarity among markers supports a multi-marker strategy

The 12S MiFish-U/E and 16S Fish16S primers showed similarly high detection success and outperformed the best-performing COI primer (Leray-Lobo) when considered individually. However, they did not detect identical species assemblages. Each primer set detected at least one unique species from the Oceanarium inventory, and this complementarity became more pronounced when unlisted species were included, with 9 and 7 exclusive detections for 12S and 16S, respectively. The MiFish 12S primers are among the most widely used markers in fish eDNA studies and are frequently applied as a single-marker approach (Miya *et al.*, 2015; Port *et al.*, 2016; Valentini *et al.*, 2016). Our results confirm their strong performance but also demonstrate that 16S provides

comparable detection success while capturing partially distinct species sets. Importantly, using any single primer set would have resulted in approximately 20% of species remaining undetected, even under optimal conditions.

Failed detections may reflect limited availability of eDNA shed by certain species, but can also result from PCR amplification bias driven by primer-template mismatches and unequal amplification efficiencies (Elbrecht and Leese, 2015; Collins *et al.*, 2019; Shaffer *et al.*, 2025). Such consistent, species-specific amplification biases are particularly problematic for routine biodiversity monitoring, as they may lead to recurrent non-detection of certain taxa regardless of their actual abundance in the ecosystem.

Overall, our findings reinforce that a multi-marker approach provides the most reliable strategy for comprehensive fish biodiversity assessment. The combined use of COI, 12S, and 16S markers helps offset individual marker limitations: 12S and 16S offer higher specificity for fish DNA and robust amplification, while COI provides superior taxonomic resolution. In addition, until reference libraries are complete, using multiple markers reduces the risk of systematic detection failures due to missing reference sequences. Despite its superior taxonomic resolution and near-complete reference library coverage, the use of COI in eDNA metabarcoding also entails clear cost-benefit trade-offs. COI amplification is particularly prone to PCR bias due to high primer degeneracy, variable binding efficiency across taxa, and the co-amplification of non-target metazoans, which can reduce effective sequencing depth for fish and increase analytical costs (Deagle *et al.*, 2014; Elbrecht and Leese, 2015; Collins *et al.*, 2019). These limitations help explain its lower detection performance relative to 12S and 16S when applied as a single marker, despite its extensive reference coverage. Within a multiple-marker framework, however, COI retains clear relevance by compensating for the lower taxonomic resolution of mitochondrial ribosomal markers and enabling confident species-level discrimination where 12S and 16S assignments remain ambiguous (Deiner *et al.*, 2017). Consequently, a multi-marker strategy represents not only a biological necessity to circumvent primer-specific biases and marker-dependent detection gaps, but also a pragmatic cost-benefit compromise that balances amplification robustness, taxonomic resolution, and reference library limitations (Thomsen and Willerslev, 2015; Weigand *et al.*, 2019).

Reference library completeness critically shapes apparent marker performance

Marker performance is intrinsically linked to the completeness and accuracy of reference databases. After adjusting detection success to account for species represented in each

reference library, the relative detection rate of 12S increased to 82.4%, slightly surpassing that of 16S (77.2%), while COI detection percentages remained unchanged due to its complete coverage of the Oceanarium species list. These results highlight the critical importance of curated, comprehensive, and regionally relevant barcode libraries for reliable eDNA-based monitoring (Weigand *et al.*, 2019). Incomplete or biased reference libraries can reduce apparent detection success and lead to ambiguous or incorrect assignments, particularly for markers such as 12S and 16S that target shorter and more conserved regions (Miya *et al.*, 2015). Despite these limitations, 12S and 16S remain highly valuable because of their amplification robustness and suitability for degraded eDNA. Qualitative differences among markers are also evident in their taxonomic resolution. COI generally provides greater discriminatory power among closely related taxa, whereas 12S and 16S more frequently produces ambiguous matches among congeners, as expected from the comparatively higher interspecific divergence of COI (Kumar *et al.*, 2022; Fontes *et al.*, 2024). This overall pattern was reflected in our results, although ambiguous assignments occurred for all markers and required comparison with the official Oceanarium species list to resolve species-level identifications: 7 of the 45 taxa cumulatively detected with COI (15.6%), 15 of the 42 taxa detected with 12S (35.7%), and 8 of the 44 taxa detected with 16S (18.2%).

Additionally, approximately 6% of all detected species were uniquely identified by a single marker, further confirming that no individual marker can capture the full extent of fish diversity. A multi-marker strategy therefore represents the best compromise between detection breadth and taxonomic resolution (Thomsen and Willerslev, 2015; Jiang *et al.*, 2022; Teixeira *et al.*, 2023; Ferreira *et al.*, 2024).

Volume of sampled water had a positive effect on species detection

Sampled water volume had a clear positive effect on species detection, with larger volumes generally resulting in higher species richness across primer sets. The only exception was observed for FishATL_Cocktail1, which detected ten Oceanarium-listed species at 2.5 L and nine at 5 L. This small difference may reflect stochastic variation associated with PCR amplification and sequencing, including occasional taxon dropout, rather than an effect of filtration volume (Ficetola *et al.*, 2015). However, because no minimum-read criterion was applied in the main analysis, the possible contribution of low-frequency PCR or sequencing artefacts cannot be excluded. Variability in species richness was also observed among the three replicate samples across filtration volumes,

with standard deviations ranging from 0 to 7.5 species depending on the primer set and filtration volume (***Supplementary Material, Table S5***). Because these replicates originated from independently collected water samples, this variability likely reflects the combined effects of heterogeneous eDNA distribution, variation arising during sampling and filtration, and stochasticity during PCR amplification and sequencing (Ficetola *et al.*, 2015; Macher *et al.*, 2021; Parsley *et al.*, 2025). As demonstrated by our results, integrating multiple replicates with multi-primer and multi-marker strategies can reduce taxon non-detection and the risk of false negatives by accounting for sample-level variability and primer-specific differences, thereby providing a more comprehensive and representative estimate of species richness (Goldberg *et al.*, 2016; Alberdi *et al.*, 2018).

When all markers were combined, up to ~74% of the total species richness and ~77% of the Oceanarium-listed species were recovered with just 5.0 L of filtered water, with only marginal gains observed at higher volumes. This pattern suggests an approaching saturation point, in line with previous studies reporting diminishing returns beyond intermediate sampling volumes (Cantera *et al.*, 2019). Nevertheless, technological advances in filtration systems are rapidly changing practical constraints. High-capacity, portable filtration devices now enable the processing of more than 30 L of water within short timeframes, potentially improving the detection of rare, low-biomass, or weakly shedding species (Spens *et al.*, 2017; Li *et al.*, 2018; Capo *et al.*, 2020). In such contexts, filtering larger volumes may be advantageous, although trade-offs between detection efficiency, processing time, and logistical feasibility must be carefully balanced in field applications. Overall, our results indicate that moderate water volumes can capture most local fish diversity under favorable conditions, while increased sampling effort may be best targeted strategically, for instance in low-density systems, early detection of rare species, or when monitoring objectives prioritize completeness over efficiency.

Detection of Exogenous Taxa and Minimum-Read Thresholds Analysis

Comparison with the Oceanarium's official inventory confirmed the overall reliability of eDNA metabarcoding, while also revealing mismatches in the form of both apparent false positives and false negatives. Among the 86 fish taxa detected, 37 were not listed in the official inventory and therefore represent potential false positives.

External contamination during water renewal is unlikely, as incoming water was quarantined and ultra-filtered prior to entering the system. A plausible explanation for several of these detections is feed-derived DNA, as many aquarium species are fed whole

625 fish or fish-based diets, introducing exogenous DNA into the system, as observed at the
626 *Oceanário de Lisboa* . For instance, *Clupea harengus*, *Engraulis encrasicolus*, *Sardina*
627 *pilchardus*, *Merluccius merluccius*, *Mallotus villosus*, *Oncorhynchus kisutch*, and
628 *Osmerus eperlanus* are commonly used in aquaculture or as frozen feed, and their DNA
629 likely originated from dietary sources.
630 Taxonomic misassignments arising from incomplete or poorly resolved reference
631 databases may also explain some detections of unlisted species, particularly for the 12S
632 and 16S markers (Goldberg *et al.*, 2016; Collins *et al.*, 2019; Weigand *et al.*, 2019; Fontes
633 *et al.*, 2024). In several cases, detected taxa belonged to genera known to be present in
634 the aquarium (e.g. *Abudefduf, Argyrosomus, Heniochus, Himantura*), suggesting that the
635 eDNA signal likely originated from a resident congener whose reference sequence was
636 absent, misidentified, or insufficiently discriminated in the database. This was evident,
637 for example, in the case of *Himantura uarnak*, initially thought to be present in the
638 Oceanarium inventory but later revised by staff to *Himantura undulata;* nevertheless,
639 eDNA metabarcoding detected *H. uarnak*. Other detections may reflect low-level
640 environmental contamination or non-target amplification, despite strict filtration and
641 handling procedures. Small pelagic species such as *Aphia minuta*, *Boops boops* or
642 *Atherina presbyter*, which are naturally present in local coastal waters and may have
643 contributed trace DNA even in closed systems (Evans *et al.*, 2017; Collins *et al.*, 2018).
644 Because no minimum-read criterion was applied in the main analysis, we explored
645 whether read filtering would substantially reduce the number of exogenous taxa initially
646 detected. Applying a minimum threshold of five reads reduced the number of exogenous
647 taxa from 37 to 32 while retaining 47 of the 49 Oceanarium-listed species, removing
648 proportionally more exogenous taxa. However, when the minimum threshold for valid
649 records was increased to ten reads, 31 exogenous and 43 listed species were retained,
650 resulting in the exclusion of the same number of taxa from each category (n = 6) compared
651 with the initial analysis. After read filtering, most exogenous taxa remained detectable
652 (32 and 31 taxa under the 5 and 10-read thresholds, respectively), indicating that the
653 absence of a minimum-read criterion can explain only part of the exogenous richness
654 initially recovered.
655 These results illustrate the trade-off between reducing potentially spurious low-read
656 detections and retaining genuine detections of species known to be present in the
657 Oceanarium. Although the effects were more pronounced and variable among individual
658 primer sets, cumulative integration across all primers and markers buffered these losses,

reflecting their complementary detection profiles. While this complementarity mitigated the loss of Oceanarium-listed species by retaining detections missed by individual primers, it also preserved several exogenous taxa, some of which may represent genuine exogenous DNA introduced into the closed system through feeding or other pathways discussed above. Nevertheless, persistence above these thresholds does not confirm the biological origin of these detections or exclude index hopping, contamination, or other technical artefacts associated with high-throughput sequencing.

Together, these results highlight the need for cautious interpretation of presence-absence data in closed systems and for systematic cross-referencing with feeding regimes and reference library completeness.

False negatives in species detection

Conversely, several species known to be present in the Oceanarium were not detected, indicating false negatives. These failed detections are likely attributable to a combination of low biomass, reduced DNA shedding rates, and species-specific ecological or physiological traits (Klymus *et al.*, 2015). Several undetected species were represented by a single individual (e.g., *Haemulon macrostomum*, *Himantura undulata*, *Myripristis jacobus*, *Paranthias furcifer*, *Rhinobatos rhinobatos*, and *Pomacanthus imperator*), whereas others had very low total biomass (e.g.; *Chaetodon rafflesia* and *Similiparma lurida*) (***Supplementary Material, Table S1***), which may explain their lack of detection under continuous water circulation and filtration.

Behavioural and habitat-related traits may also have influenced eDNA availability. Water samples were collected at a depth of 3.5 m, approximately in the middle of the tank, reflecting the approach used in many open-water eDNA studies, in which water is collected only a few meters below the surface. However, because samples were collected exclusively from the water column and away from the bottom, eDNA from benthic species may have remained in the lower layers and thus not been collected. *Rhinobatos rhinobatos* and *Himantura undulata* are benthic elasmobranchs that spend long periods resting on or partially buried within the substrate, potentially releasing eDNA preferentially into sediments rather than the water column. Similarly, *Myripristis jacobus*, *Paranthias furcifer*, and *Ocyurus chrysurus* tend to occupy specific depth strata or refuge areas and may exhibit limited movement within the tank, generating spatially restricted eDNA signals that are less likely to be captured by bulk water sampling. This represents

a sampling design limitation arising from logistical constraints and the inability to sample all strata of the water column.

Species-specific physiology may further contribute to reduced eDNA shedding. For instance, taxa with thick mucus layers, high biomass but low abundance (e.g., elasmobranchs), or lower metabolic rates due to ageing may release proportionally less detectable DNA into the surrounding water (Maruyama *et al.*, 2014; Jo *et al.*, 2019). Herbivorous or grazing species such as *Acanthurus* spp. and *Similiparma lurida* may also contribute less eDNA via feeding-related processes compared to carnivorous species, which can release additional DNA through prey handling and digestion.

Taxonomic resolution issues may additionally explain some false negatives where closely related congeners were present (Wilcox *et al.*, 2014). *Chaetodon rafflesii*, *Haemulon macrostomum*, and *Similiparma lurida* were not detected, whereas other species from the same genera (*Chaetodon auriga*, *C. lunula*; *Haemulon flavolineatum*; *Abudefduf hoefleri*, *A. saxatilis*, *A. vaigiensis*) were successfully identified. This pattern suggests limited discriminatory power within certain genera, potentially due to low interspecific divergence in the targeted marker region and/or incomplete or uneven reference database coverage. Such limitations are particularly relevant for ribosomal markers, which may lack sufficient resolution for closely related taxa (Fontes *et al.*, 2024).

Finally, primer-template mismatches may have contributed to the consistent failed detection of some species, especially if primer binding sites differ from those represented during primer design or validation (Collins *et al.*, 2019). Evaluating primer mismatch profiles for undetected taxa and expanding reference libraries with locally relevant species could help clarify these failures and inform future primer optimisation strategies.

Conclusions

Globally, our results illustrate the challenges of reaching full fish species detection using eDNA metabarcoding, even employing a comprehensive approach combining multiple primers and comparatively large water volumes in a closed artificial system. Although a fair level of detection was achieved, and the merits of eDNA metabarcoding for fish species monitoring still hold, it is likely that its performance will be lower in open natural marine environments, where eDNA may be diluted by orders of magnitude and additional factors may contribute to eDNA loss and lower detection efficiency.

To minimize false negatives in fish species monitoring in natural environments, the combination of multiple markers (e.g. COI, 12S and/or 16S) is recommended. This

strategy will help countering limiting factors such as primer bias, insufficient discrimination power and reference libraries incompleteness. Whereas the usage of multiple markers implies greater sequencing costs, it will also maximize the information generated by unit of sampling effort. With sequencing becoming tendentially less expensive, the cost-benefit balance of using multiple markers will likely increase. Furthermore, the efficiency costs of failed detections when using single-primer approaches must also be accounted for. This is particularly relevant if certain species are systematically overlooked for technical reasons, leading to persistent biases in monitoring data. Finally, our results also indicated that larger volumes of seawater should be filtered (~5 L), compared to those typically used in most studies (usually up to 1-2 L). Although processing these volumes may present logistical constraints and increase the risk of filter clogging, these limitations can be substantially reduced through direct on-site filtration using portable pumps and recently developed large-surface-area encapsulated filters, which eliminate the need to transport bulk water and allow larger volumes to be processed directly in the field (Douchet *et al.*, 2022; Roblet *et al.*, 2024). The wider adoption and continued improvement of these filtration systems will be important for obtaining more representative assessments of aquatic communities using eDNA.

Acknowledgements

We sincerely thank the staff of the *Oceanário de Lisboa* for their collaboration and assistance in the collection of the samples within the aquarium tank, which made this study possible.

Funding

This work was funded by FCT, I.P. through the project "*A-Fish-DNA-Scan: Cutting-edge DNA-based approaches for improved monitoring and management of fisheries resources along Magellan-Elcano's Atlantic route"* (CIRCNABRB/0156/2019; https://doi.org/10.54499/CIRCNA/BRB/0156/2019). Additional support was provided by the CBMA (UIDB/04050/2020; https://doi.org/10.54499/UIDB/04050/2020) and ARNET (LA/P/0069/2020; https://doi.org/10.54499/LA/P/0069/2020) programme contracts, funded by national funds through FCT, I.P.

Individual support was provided by FCT, I.P. to A.O.F. (DFA/BD/6653/2020; https://doi.org/10.54499/2020.06653.BD), C.M. (UMINHO/BID/2021/36) and S.D. (https://doi.org/10.54499/CEECIND/00667/2017/CP1458/CT0001).

This work also benefited from funding provided by the European Regional Development Fund (ERDF) through COMPETE 2020, as well as national funds from FCT, I.P. under the strategic projects UIDB/04539/2020, UIDP/04539/2020 and LA/P/0058/2020 and contributed to the CCMAR Strategic Research Plan (UIDB/04326/2020, UIDP/04326/2020 and LA/P/0101/2020).

Author contributions

André O. Ferreira - Conceptualization, Data curation, Investigation, Methodology, Visualization, Writing - original draft; Cláudia Machado - Conceptualization, Investigation, Methodology, Writing - original draft; Cristina Barroso - Methodology, Resources; Núria Baylina - Methodology, Resources; Conceição Egas - Methodology, Resources; João T. Fontes - Investigation, Methodology, Writing - original draft; Margarida Fernandes - Methodology, Resources; Cláudio Oliveira - Methodology, Resources; A. Miguel Piecho-Santos - Conceptualization, Methodology; Sofia Duarte - Conceptualization, Investigation, Methodology, Supervision, Writing - original draft; Filipe O. Costa - Conceptualization, Funding acquisition, Supervision, Writing - original draft. All authors - Writing - review & editing.

Data availability

All sequence data generated in this study are publicly available at the NCBI Sequence Read Archive (SRA) under BioProject accession PRJNA1427588. Additional data and supporting materials are provided in the Supplementary Material files.

Competing interests

The authors declare there are no competing interests.

References

Alberdi, A. *et al.* (2018) 'Scrutinizing key steps for reliable metabarcoding of environmental samples', *Methods in Ecology and Evolution*, 9(1), pp. 134–147. doi: 10.1111/2041-210X.12849.

Barnes, M. A. *et al.* (2014) 'Environmental conditions influence eDNA persistence in aquatic systems', *Environmental science & technology*, 48(3), pp. 1819–1827. doi: 10.1021/es404734p.

794 Berry, T. E. *et al.* (2017) 'DNA metabarcoding for diet analysis and biodiversity: A case
795 study using the endangered Australian sea lion (Neophoca cinerea)', *Ecology and*
796 *Evolution*, 7(14), pp. 5435–5453. doi: 10.1002/ece3.3123.
797 Bessey, C. *et al.* (2020) 'Maximizing fish detection with eDNA metabarcoding',
798 *Environmental DNA*, 2(4), pp. 493–504. doi: 10.1002/edn3.74.
799 Bruce, K. *et al.* (2021) *A practical guide to DNA-based methods for biodiversity*
800 *assessment.* Pensoft Advanced Books. doi: 10.3897/ab.e68634.
801 Bucklin, A. *et al.* (2021) 'Toward a global reference database of COI barcodes for marine
802 zooplankton', *Marine Biology*, 168(6), pp. 1–26. doi: 10.1007/s00227-021-03887-y.
803 Cantera, I. *et al.* (2019) 'Optimizing environmental DNA sampling effort for fish
804 inventories in tropical streams and rivers', *Scientific reports*, 9(1), p. 3085. doi:
805 10.1038/s41598-019-39399-5.
806 Capo, E. *et al.* (2020) 'Effects of filtration methods and water volume on the
807 quantification of brown trout (Salmo trutta) and Arctic char (Salvelinus alpinus) eDNA
808 concentrations via droplet digital PCR', *Environmental DNA*, 2(2), pp. 152–160. doi:
809 10.1002/edn3.52.
810 Carvalho, C. O. *et al.* (2024) 'Harnessing eDNA metabarcoding to investigate fish
811 community composition and its seasonal changes in the Oslo fjord', *Scientific Reports*,
812 14(1), p. 10154. doi: 10.1038/s41598-024-60762-8.
813 Collins, R. A. *et al.* (2018) 'Persistence of environmental DNA in marine systems',
814 *Communications Biology*, 1(1), p. 185. doi: 10.1038/s42003-018-0192-6.
815 Collins, R. A. *et al.* (2019) 'Non-specific amplification compromises environmental DNA
816 metabarcoding with COI', *Methods in Ecology and Evolution*, 10(11), pp. 1985–2001.
817 doi: 10.1111/2041-210X.13276.
818 Costa, F. O. *et al.* (2012) 'A Ranking System for Reference Libraries of DNA Barcodes:
819 Application to Marine Fish Species from Portugal', *PLoS ONE*, 7(4), p. e35858. doi:
820 10.1371/journal.pone.0035858.
821 Costa, G. O. *et al.* (2025) 'Identification of fish eggs and larvae using a novel DNA
822 metabarcoding approach', *Neotropical Ichthyology*, 23, p. e250005. doi: 10.1590/1982-
823 0224-2025-0005.
824 Deagle, B. E. *et al.* (2007) 'Studying seabird diet through genetic analysis of faeces: A
825 case study on macaroni penguins (Eudyptes chrysolophus)', *PLoS ONE*, 2(9), p. e831.
826 doi: 10.1371/journal.pone.0000831.
827 Deagle, B. E. *et al.* (2014) 'DNA metabarcoding and the cytochrome c oxidase subunit I

828 marker: Not a perfect match', *Biology Letters*, 10(9), pp. 2–5. doi:
829 10.1098/rsbl.2014.0562.
830 Deiner, K. *et al.* (2016) 'Environmental DNA reveals that rivers are conveyer belts of
831 biodiversity information', *Nature communications*, 7(1), p. 12544. doi:
832 10.1038/ncomms12544.
833 Deiner, K. *et al.* (2017) 'Environmental DNA metabarcoding: Transforming how we
834 survey animal and plant communities', *Molecular Ecology*, 26(21), pp. 5872–5895. doi:
835 10.1111/mec.14350.
836 DiBattista, J. D. *et al.* (2017) 'Assessing the utility of eDNA as a tool to survey reef-fish
837 communities in the Red Sea', *Coral Reefs*, 36(4), pp. 1245–1252. doi: 10.1007/s00338-
838 017-1618-1.
839 Douchet, P. *et al.* (2022) 'Make visible the invisible: Optimized development of an
840 environmental DNA metabarcoding tool for the characterization of trematode parasitic
841 communities', *Environmental DNA*, 4(3), pp. 627–641. doi: 10.1002/EDN3.273.
842 Edgar, R. C. (2010) 'Search and clustering orders of magnitude faster than BLAST',
843 *Bioinformatics*, 26(19), pp. 2460–2461. doi: 10.1093/bioinformatics/btq461.
844 Eichmiller, J. J., Miller, L. M. and Sorensen, P. W. (2016) 'Optimizing techniques to
845 capture and extract environmental DNA for detection and quantification of fish',
846 *Molecular ecology resources*, 16(1), pp. 56–68. doi: 10.1111/1755-0998.12421.
847 Elbrecht, V. and Leese, F. (2015) 'Can DNA-based ecosystem assessments quantify
848 species abundance? Testing primer bias and biomass-sequence relationships with an
849 innovative metabarcoding protocol', *PLoS ONE*, 10(7), pp. 1–16. doi:
850 10.1371/journal.pone.0130324.
851 Evans, N. T. *et al.* (2017) 'Fish community assessment with eDNA metabarcoding:
852 Effects of sampling design and bioinformatic filtering', *Canadian Journal of Fisheries
853 and Aquatic Sciences*, 74(9), pp. 1362–1374. doi: 10.1139/cjfas-2016-0306.
854 Fediajevaite, J. *et al.* (2021) 'Meta-analysis shows that environmental DNA outperforms
855 traditional surveys, but warrants better reporting standards', *Ecology and Evolution*,
856 11(9), pp. 4803–4815. doi: 10.1002/ece3.7382.
857 Ferreira, A. O. *et al.* (2024) 'Multi-marker DNA metabarcoding for precise species
858 identification in ichthyoplankton samples', *Scientific Reports*, 14, p. 19772. doi:
859 10.1038/s41598-024-69963-7.
860 Ficetola, G. F. *et al.* (2015) 'Replication levels, false presences and the estimation of the
861 presence/absence from eDNA metabarcoding data', *Molecular Ecology Resources*, 15(3),

862 pp. 543–556. doi: 10.1111/1755-0998.12338.
863 Fontes, J. T. *et al.* (2021) 'BAGS: An automated Barcode, Audit & Grade System for
864 DNA barcode reference libraries', *Molecular Ecology Resources*, 21(2), pp. 573–583.
865 doi: 10.1111/1755-0998.13262.
866 Fontes, J. T. *et al.* (2024) 'Benchmarking the discrimination power of commonly used
867 markers and amplicons in marine fish (e)DNA (meta)barcoding', *Metabarcoding and*
868 *Metagenomics*, 8, p. e128646. doi: 10.3897/MBMG.8.128646.
869 Fraija-Fernández, N. *et al.* (2020) 'Marine water environmental DNA metabarcoding
870 provides a comprehensive fish diversity assessment and reveals spatial patterns in a large
871 oceanic area', *Ecology and Evolution*, 10(14), pp. 7560–7584. doi: 10.1002/ece3.6482.
872 Garcia-Vazquez, E. *et al.* (2021) 'eDNA metabarcoding of small plankton samples to
873 detect fish larvae and their preys from Atlantic and Pacific waters', *Scientific Reports*,
874 11(1), pp. 1–14. doi: 10.1038/s41598-021-86731-z.
875 Gold, Z., Sprague, J., *et al.* (2021a) 'eDNA metabarcoding as a biomonitoring tool for
876 marine protected areas', *PLoS ONE*, 16(2), p. e0238557. doi:
877 10.1371/journal.pone.0238557.
878 Gold, Z., Curd, E. E., *et al.* (2021b) 'Improving metabarcoding taxonomic assignment: A
879 case study of fishes in a large marine ecosystem', *Molecular Ecology Resources*, 21(7),
880 pp. 2546–2564. doi: 10.1111/1755-0998.13450.
881 Goldberg, C. S. *et al.* (2016) 'Critical considerations for the application of environmental
882 DNA methods to detect aquatic species', *Methods in Ecology and Evolution*, 7(11), pp.
883 1299–1307. doi: 10.1111/2041-210X.12595.
884 Hajibabaei, M. *et al.* (2011) 'Environmental barcoding: A next-generation sequencing
885 approach for biomonitoring applications using river benthos', *PLoS ONE*, 6(4). doi:
886 10.1371/journal.pone.0017497.
887 Heberle, H. *et al.* (2015) 'InteractiVenn: A web-based tool for the analysis of sets through
888 Venn diagrams', *BMC Bioinformatics*, 16(1), pp. 1–7. doi: 10.1186/s12859-015-0611-3.
889 Hebert, P. D. N. *et al.* (2003) 'Biological identifications through DNA barcodes',
890 *Proceedings of the Royal Society B: Biological Sciences*, 270, pp. 313–321. doi:
891 10.1098/rspb.2002.2218.
892 Inkscape Project (2020) 'Inkscape (Version 1.2.2) [Computer software]'. Available at:
893 https://inkscape.org.
894 Jennings, W. B. *et al.* (2019) 'Barcoding the Neotropical freshwater fish fauna using a
895 new pair of universal COI primers with a discussion of primer dimers and M13 primer

tails', *Genome*, 62(2), pp. 77–83. doi: 10.1139/gen-2018-0145.
Jiang, R., Lusana, J. L. and Chen, Y. (2022) 'High-Throughput DNA Metabarcoding as an Approach for Ichthyoplankton Survey in Oujiang River Estuary, China', *Diversity*, 14(12), p. 1111. doi: 10.3390/d14121111.
Jo, T. *et al.* (2019) 'Effect of water temperature and fish biomass on environmental DNA shedding, degradation, and size distribution', *Ecology and Evolution*, 9(3), pp. 1135–1146. doi: 10.1002/ece3.4802.
Kelly, R. P. *et al.* (2014) 'Using environmental DNA to census marine fishes in a large mesocosm', *PloS one*, 9(1), p. e86175. doi: 10.1371/journal.pone.0086175.
Klymus, K. E. *et al.* (2015) 'Quantification of eDNA shedding rates from invasive bighead carp *Hypophthalmichthys nobilis* and silver carp Hypophthalmichthys molitrix', *Biological conservation*, 183, pp. 77–84. doi: 10.1016/j.biocon.2014.11.020.
Kumar, G. *et al.* (2022) 'Comparing eDNA metabarcoding primers for assessing fish communities in a biodiverse estuary', *PLoS ONE*, 17, pp. 1–20. doi: 10.1371/journal.pone.0266720.
Leray, M. *et al.* (2013) 'A new versatile primer set targeting a short fragment of the mitochondrial COI region for metabarcoding metazoan diversity: Application for characterizing coral reef fish gut contents', *Frontiers in Zoology*, 10, p. 34. doi: 10.1186/1742-9994-10-34.
Li, J. *et al.* (2018) 'The effect of filtration method on the efficiency of environmental DNA capture and quantification via metabarcoding', *Molecular Ecology Resources*, 18(5), pp. 1102–1114. doi: 10.1111/1755-0998.12899.
Lobo, J. *et al.* (2013) 'Enhanced primers for amplification of DNA barcodes from a broad range of marine metazoans', *BMC Ecology*, 13, p. 34. doi: 10.1186/1472-6785-13-34.
Macher, T. H. *et al.* (2021) 'Beyond fish eDNA metabarcoding: Field replicates disproportionately improve the detection of stream associated vertebrate species', *Metabarcoding and Metagenomics 5: e66557*, 5, p. e66557. doi: 10.3897/MBMG.5.66557.
Maruyama, A. *et al.* (2014) 'The release rate of environmental DNA from juvenile and adult fish', *PLoS ONE*, 9(12), pp. 1–13. doi: 10.1371/journal.pone.0114639.
Miya, M. *et al.* (2015) 'MiFish, a set of universal PCR primers for metabarcoding environmental DNA from fishes: Detection of more than 230 subtropical marine species', *Royal Society Open Science*, 2(7), p. 150088. doi: 10.1098/rsos.150088.
Miya, M., Gotoh, R. O. and Sado, T. (2020) 'MiFish metabarcoding: a high-throughput

approach for simultaneous detection of multiple fish species from environmental DNA and other samples', *Fisheries Science*, 86(6), pp. 939–970. doi: 10.1007/s12562-020-01461-x.

Morey, K. C., Bartley, T. J. and Hanner, R. H. (2020) 'Validating environmental DNA metabarcoding for marine fishes in diverse ecosystems using a public aquarium', *Environmental DNA*, 2(3), pp. 330–342. doi: 10.1002/edn3.76.

Mugnai, F. *et al.* (2023) 'Be positive: customized reference databases and new, local barcodes balance false taxonomic assignments in metabarcoding studies', *PeerJ*, 11, p. e14616. doi: 10.7717/PEERJ.14616/SUPP-15.

Parsley, M. B. *et al.* (2025) 'Sample-To-Sample Variability Impacts eDNA Quantification, With Implications for Estimating Species Abundance', *Environmental DNA*, 7(5), p. e70195. doi: 10.1002/EDN3.70195.

Port, J. A. *et al.* (2016) 'Assessing vertebrate biodiversity in a kelp forest ecosystem using environmental DNA', *Molecular ecology*, 25(2), pp. 527–541. doi: 10.1111/mec.13481.

Ratcliffe, F. C. *et al.* (2021) 'Quantitative assessment of fish larvae community composition in spawning areas using metabarcoding of bulk samples', *Ecological Applications*, 31(3), pp. 1–10. doi: 10.1002/eap.2284.

Ratnasingham, S. *et al.* (2024) 'BOLD v4: A Centralized Bioinformatics Platform for DNA-Based Biodiversity Data', in *DNA Barcoding: Methods and Protocols*. Springer, pp. 403–441. doi: doi.org/10.1007/978-1-0716-3581-0_26.

Ratnasingham, S. and Hebert, P. D. N. (2013) 'A DNA-Based registry for all animal species: The Barcode Index Number (BIN) system', *PLoS ONE*, 8(7), p. e66213. doi: 10.1371/journal.pone.0066213.

Roblet, S. *et al.* (2024) 'Primer set evaluation and sampling method assessment for the monitoring of fish communities in the North-western part of the Mediterranean Sea through eDNA metabarcoding', *Environmental DNA*, 6(3), p. e554. doi: 10.1002/EDN3.554.

RStudio Team (2020) 'RStudio: Integrated Development for R (Version 4.0.4) [Computer software]'.

Sayers, E. W. *et al.* (2025) 'Database resources of the National Center for Biotechnology Information in 2025', *Nucleic acids research*, 53(D1), pp. D20–D29. doi: 10.1093/nar/gkae979.

Schroeder, A. *et al.* (2020) 'DNA metabarcoding and morphological analysis - Assessment of zooplankton biodiversity in transitional waters', *Marine Environmental*

Research, 160(December 2019). doi: 10.1016/j.marenvres.2020.104946.

Shaffer, M. R. *et al.* (2025) 'Observation bias in metabarcoding', *Molecular Ecology Resources*, p. e14119. doi: 10.1111/1755-0998.14119.

Sigsgaard, E. E. *et al.* (2015) 'Monitoring the near-extinct European weather loach in Denmark based on environmental DNA from water samples', *Biological Conservation*, 183, pp. 46–52. doi: 10.1016/j.biocon.2014.11.023.

Spens, J. *et al.* (2017) 'Comparison of capture and storage methods for aqueous macrobial eDNA using an optimized extraction protocol: advantage of enclosed filter', *Methods in Ecology and Evolution*, 8(5), pp. 635–645. doi: 10.1111/2041-210X.12683.

Taberlet, P. *et al.* (2012) 'Environmental DNA', *Molecular Ecology*, 21(8), pp. 1789–1793. doi: 10.1111/j.1365-294X.2012.05542.x.

Teixeira, D. F. *et al.* (2023) 'DNA metabarcoding assessment of Neotropical ichthyoplankton communities is marker-dependent', *Ecology and Evolution*, 13(10), pp. 1–9. doi: 10.1002/ece3.10649.

Thomsen, P. F. *et al.* (2012) 'Detection of a diverse marine fish fauna using environmental DNA from seawater samples', *PLoS ONE*, 7(8), pp. 1–9. doi: 10.1371/journal.pone.0041732.

Thomsen, P. F. and Willerslev, E. (2015) 'Environmental DNA–An emerging tool in conservation for monitoring past and present biodiversity', *Biological conservation*, 183, pp. 4–18. doi: https://doi.org/10.1016/j.biocon.2014.11.019.

Valentini, A. *et al.* (2016) 'Next-generation monitoring of aquatic biodiversity using environmental DNA metabarcoding', *Molecular Ecology*, 25(4), pp. 929–942. doi: 10.1111/mec.13428.

Ward, R. D. *et al.* (2005) 'DNA barcoding Australia's fish species', *Philosophical Transactions of the Royal Society B: Biological Sciences*, 360(1462), pp. 1847–1857. doi: 10.1098/rstb.2005.1716.

Weigand, H. *et al.* (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, pp. 499–524. doi: 10.1016/j.scitotenv.2019.04.247.

Wickham, H. (2016) *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York. Available at: https://ggplot2.tidyverse.org.

Wickham, H. *et al.* (2023) 'dplyr: A grammar of data manipulation (Version 1.1.4) [R package]'. Available at: https://github.com/tidyverse/dplyr.

Wickham, H., Vaughan, D. and Girlich, M. (2024) 'tidyr: Tidy messy data (Version 1.3.1)

998 [R package]'. Available at: https://cran.r-project.org/package=tidyr.
999 Wilcox, T. M. *et al.* (2014) 'A blocking primer increases specificity in environmental
1000 DNA detection of bull trout (*Salvelinus confluentus*)', *Conservation Genetics Resources*,
1001 6(2), pp. 283–284. doi: 10.1007/s12686-013-0113-4.
1002 WoRMS Editorial Board (2025) *World Register of Marine Species*. doi: 10.14284/170.
1003 Yamanaka, H. and Minamoto, T. (2016) 'The use of environmental DNA of fishes as an
1004 efficient method of determining habitat connectivity', *Ecological Indicators*, 62, pp. 147–
1005 153. doi: 10.1016/j.ecolind.2015.11.022.
1006 Zhang, S., Zhao, J. and Yao, M. (2020) 'A comprehensive and comparative evaluation of
1007 primers for metabarcoding eDNA from fish', *Methods in Ecology and Evolution*, 11(12),
1008 pp. 1609–1625. doi: 10.1111/2041-210X.13485.
1009