scientific reports

Check for updates

OPEN

An optimized eDNA protocol for fish tracking in estuarine environments

Fouad El Baidouri[1✉], Alison W. Watts[1✉], Jeffrey T. Miller[3], Muriel Kelly[2], Joseph L. Sevigny[2], Heather Gilbert[1] & W. Kelley Thomas[2]

Environmental DNA (eDNA) is revolutionizing how we investigate biodiversity in aquatic and terrestrial environments. It is increasingly used for detecting rare and invasive species, assessing biodiversity loss and monitoring fish communities, as it is considered a cost-effective and noninvasive approach. Some environments, however, can be challenging for eDNA analyses. Estuarine systems are highly productive, complex environments, but samples collected from these settings may exhibit PCR inhibition and a low fish read recovery. Here we present an approach for detecting fish in turbid, highly productive estuarine systems. The workflow includes bead-based extraction, inhibition removal, high fidelity and specificity DNA polymerase (Platinum SuperFi II) and multiplexing the universal MiFish primers. By applying this hybrid method to a variety of complex estuarine samples with known inhibition, we have more than doubled the number of recovered fish species while removing most of the off-target amplification.

Keywords eDNA, MiFish, Platinum SuperFi II, PCR inhibition, Estuarine, Fish species, Multiplex PCR

The application of environmental DNA (eDNA) metabarcoding has significantly reshaped our approach for evaluating biodiversity in aquatic and terrestrial ecosystems. In aquatic systems, this technique has become a useful tool for assessing and monitoring fish communities[1,2]. The method involves PCR amplification of DNA extracted from water samples using primers designed to specifically target universal gene markers such as mitochondrial or nuclear ribosomal genes to identify a wide array of species. A variety of fish primers has been developed in the past years[3]. Among them, the universal primer pairs developed by Miya et al.[4] for broad-range fish eDNA metabarcoding are particularly notable and have become some of the most widely used in the field. This approach supports the detection of diverse fish species from environmental samples. Kawato et al.[5] built on this by introducing a method for detecting deep-sea fish species by multiplexing a set of MiFish primers, originally developed by Miya et al. (MiFish-U and MiFish-E), using a new high-fidelity enzyme. These primers, optimized for enhanced specificity and coverage, demonstrate the potential for precise eDNA analysis in deep aquatic ecosystems. This primer set is widely used and is complemented by a curated MitoFish database[6] that contains mitochondrial genomes and aids in accurate species identification. However, co-amplification of non-target reads, particularly bacteria, can be problematic. These non-target reads arise from the unintended amplification of non-fish DNA, such as bacterial sequences, during PCR. This reduces the proportion of target fish DNA, potentially leading to inaccurate species identification, diminished detection sensitivity, and wasted sequencing effort on non-relevant sequences. Various solutions, including using a touchdown PCR protocol[7], and primer design modifications e.g.[8] have been applied with varying degrees of success.

Despite these efforts, PCR in environmental samples from turbid environments remains challenging[9–11]. Turbidity refers to suspended particles that reduce light transmissivity in water[9]. Particles may be from inorganic sediments or organic material such as plankton or other biomass suspended in the water column. These particles can complicate eDNA analyses in several ways: filter clogging can reduce the amount of water filtered or increase the number of filters to be processed, and high organic material may contain inhibitory compounds[12]. In areas with high biological productivity, fish DNA may comprise a relatively small portion of total DNA recovered from a sample. In addition to low target eDNA concentrations, these samples often exhibit PCR inhibition where the presence of humic acids, and other organic and inorganic compounds can significantly reduce the efficiency and accuracy of the PCR[9].

[1]Department of Civil & Environmental Engineering, University of New Hampshire, Durham, NH 03824, USA. [2]Hubbard Center for Genome Studies, University of New Hampshire, Durham, NH 03824, USA. [3]Department of Molecular, Cellular, and Biomedical Sciences, University of New Hampshire, Durham, NH 03824, USA. ✉email: fouad.elbaidouri@unh.edu; alison.watts@unh.edu

In response to these challenges, we have developed an eDNA workflow with two main protocols, optimized for detecting fish in turbid estuarine samples. We evaluated each step in these protocols to identify a workflow that is adaptable to robotic platforms, cost-efficient, and optimizes the number of fish species and the number of generated fish Amplicon Sequence Variants (ASVs). Our final selected workflow included an automated KingFisher DNA extraction system, which utilizes magnetic beads for high-throughput DNA isolation, and an optimized PCR protocol. This protocol incorporates the Zymo OneStep™ PCR Inhibitor Removal Kit, to reduce inhibitors commonly present in environmental samples, and the Platinum SuperFi II polymerase, recognized for its high fidelity and specificity in amplifying low-concentration DNA. Additionally, we implemented a touchdown program, which progressively lowers the annealing temperature during the initial cycles to enhance specificity and amplification efficiency. This approach is particularly effective when Multiplexing Mifish primers, which are designed to detect a broad range of fish species.

The protocols in this workflow were tested on 48 samples collected from four estuaries in the United States' National Estuarine Research Reserve System (NERRS) a nationwide network of protected estuarine areas dedicated to research, education, and stewardship to inform coastal management. Sites included San Francisco Bay CA, Mission-Aransas TX, Apalachicola FL and Jacques Cousteau NJ. This geographical diversity in site selection was intended to ensure a broad spectrum of environmental conditions, sample variability and species diversity.

San Francisco Bay is a highly urbanized estuary characterized by high turbidity levels resulting from suspended inorganic sediments and organic matter—a legacy of historical activities like including hydraulic gold mining and ongoing urban runoff[13]. Despite reductions in sediment supply due to damming and sediment retention upstream, turbidity remains a defining feature of the estuary. The NERRS System-Wide Monitoring Program (SWMP) collects long-term water quality data at all NERR sites, including this one (see Table S2 in the Supplementary Information). Elevated turbidity at this site presents challenges for eDNA sampling, including filter clogging and PCR inhibition caused by organic and inorganic particulates. Additionally, the estuary sampling sites encompasses a substantial portion of the salinity gradient, ranging from salt marsh conditions at China Camp to brackish marsh at Rush Ranch, resulting in rich biodiversity, including rare and endangered species.

Mission-Aransas is a dynamic estuarine system that offers a contrasting environmental landscape to San Francisco Bay. It is characterized by a complex of wetland, terrestrial, and marine environments—including coastal prairies, freshwater and saltwater marshes, seagrass meadows, mangroves, and oyster reefs[14]. This estuary also experiences large salinity gradients due to variable freshwater inflows from the Mission and Aransas rivers, tidal exchange, and high evaporation rates. During periods of drought, low river flows and high evaporation can lead to hypersaline conditions in shallow bays, while freshwater pulses from rainfall events can significantly lower salinity levels for extended periods. This variability influences nutrient dynamics and biological productivity, resulting in fluctuations of organic matter concentrations.

Apalachicola Bay introduces another dimension of environmental complexity within our study. It encompasses diverse habitats, including barrier islands, riverine systems, and estuarine environments[15]. It features a warm, humid subtropical climate with significant seasonal temperature variations. The estuary receives substantial freshwater input from the Apalachicola River, creating dynamic salinity gradients influenced by seasonal river flows, local rainfall, and tidal exchanges with the Gulf of Mexico. This variability leads to fluctuations in nutrient levels and high biological productivity, supporting a rich biodiversity that includes important commercial fisheries and rare species. The varying freshwater inflows result in changes in turbidity and organic matter concentrations.

Finally, the Jacques Cousteau National Estuarine Research Reserve in New Jersey is one of the most pristine estuarine systems along the densely populated northeastern United States. It encompasses diverse habitats, including upland pine-oak forests, hardwood swamps, tidal marshes, barrier islands, and open estuarine and coastal ocean waters[16]. The reserve features temperate conditions with well-defined seasons and experiences significant seasonal temperature variations. The Mullica River-Great Bay Estuary within the reserve is characterized by low-gradient, southeast-flowing streams that originate from groundwater inflows. These streams have high concentrations of humic acids and organic matter from decaying vegetation, leading to brown-colored waters rich in dissolved organic carbon. Seasonal salinity gradients occur due to variable freshwater inputs, tides, and evaporation rates.

By selecting these four diverse estuaries, we can test our eDNA workflow across a range of challenging environmental conditions, including high turbidity, variable salinity, and high organic matter content. This approach ensures that the protocols we developed are robust and adaptable for detecting fish eDNA in various turbid estuarine environments.

In addition to PCR optimization, we tested two DNA extraction methods to evaluate their effectiveness for fish eDNA detection in estuarine environments: the KingFisher automated bead-based extraction and the Qiagen-based extraction. The KingFisher method, optimized for high-throughput applications, uses paramagnetic beads to automate and streamline DNA isolation, while the Qiagen-based extraction relies on a silica column, a trusted method widely used in eDNA studies. We compared these two approaches to determine which method delivers the highest DNA yield and quality, particularly when working with challenging turbid samples, to improve the overall efficiency and reliability of eDNA-based fish species detection.

Results and discussion

DNA extraction

As interest in molecular monitoring increases, the use of robotic systems in eDNA extraction and handling is becoming more common, especially for processing larger numbers of samples. Robotic systems offer advantages such as increased efficiency and consistency, though they also have potential disadvantages, including the risk

of cross-contamination and higher initial costs. Our goal was to validate an automated bead-based extraction method with comparable performance to commonly used column-based methods like QIAGEN kit-based extractions.

Samples were extracted with the automated Kingfisher protocol using magnetic beads, and DNA concentrations ranged between 1.84 ng/µl and 25.8 ng/µl (**Fig. S1,** Supplementary Information). As expected, we observed variability among the sites with Mission-Aransas displaying the highest yields (**Fig. S1**). Variability within each site reflects conditions at each of the four sampling locations (**Table S1**, Supplementary information). We did not observe any cross-contamination effects, and we found no significant differences between DNA concentrations from samples obtained with the bead-based Kingfisher extraction and the QIAGEN protocol (*p-value* = 0.7; **Fig. S2**, Supplementary Information). This suggests that our automated DNA extraction using magnetic beads is equally effective under various conditions.

PCR inhibition removal and optimization

Initial evaluations of PCR amplification were conducted through visual inspection of E-Gels. The absence of amplification in samples, where target species were expected, indicates the presence of inhibitory compounds as seen in samples from San Francisco Bay[9] (Fig. 1a. KAPA.CTRL). When an inhibition removal step was introduced to the same protocol using Zymo, amplification was observed, but primarily of off-target sequences, as evidenced by bands appearing higher than the expected ~ 400 bp with adapters (Fig. 1b. KAPA.ZYMO). Further improvements were achieved by utilizing Platinum SuperFi II polymerase (Fig. 1c. PLATI.ZYMO), which resulted in amplification within the expected size range for fish (~ 300 bp with adapters). This underscores the effectiveness of a tailored polymerase in enhancing target-specific amplification while reducing off-target effects. The improved specificity can be attributed to the hot-start mechanism of the Platinum SuperFi II polymerase, which activates the enzyme only at high temperatures, thereby preventing the extension of misprimed targets and primer-dimers. Additionally, the high processivity and fidelity of this polymerase help minimize the occurrence of nonspecific products. Our experience illustrates that while adding an inhibition removal step incurs additional costs, it substantially enhances species detection by mitigating the impact of PCR inhibitors present in estuarine samples.

Balancing amplification fidelity and costs

Inhibition is prevalent across many estuarine samples processed in our lab, often muting or completely preventing PCR amplification. Laboratory managers might opt to apply inhibition removal selectively based on initial PCR results or incorporate it routinely into the workflow to consistently improve outcomes. The Zymo OneStep™ PCR Inhibitor Removal Kit costs approximately $2.07 per sample and using Platinum SuperFi II PCR Master Mix adds around $2.78 per sample, bringing the total increase to $4.85 per sample excluding labor costs. These steps significantly improve the specificity and success of amplification, especially in challenging environments. There is a risk however, that an overly specific polymerase will result in non-detection of some target species.

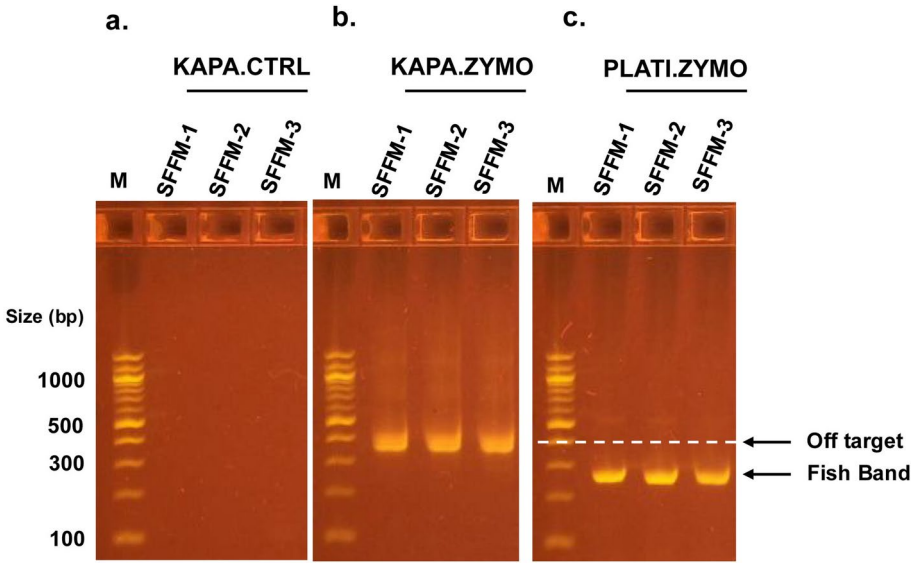

Fig. 1. PCR inhibition removal using Zymo OneStep. Panel 1a: Samples from the San Francisco NERRS site show no detectable PCR amplification due to inhibition when using the KAPA protocol without inhibition removal (KAPA.CTRL). Panel 1b: The introduction of the Zymo OneStep™ PCR Inhibitor Removal Kit (KAPA.ZYMO) results in visible amplification but also generates off-target DNA fragments, as indicated by additional bands above the expected range. Panel 1c: Using the Zymo kit in combination with the Platinum SuperFi II polymerase and a touchdown PCR program, Platinum.ZYMO (PLATI.ZYMO), amplification is restored, and the resulting DNA fragments fall within the expected range of 200–300 bp, as indicated by cleaner bands. M: DNA ladder (marker) indicates the fragment size.

Optimizing fish detection

To enhance the accuracy and efficiency of fish species detection in estuarine environments, we implemented several methodological adjustments (see Methods). When developing this study, we selected sites known from our previous eDNA analyses to be problematic due to turbid water and common inhibitors (**Table S2,** Supplementary Information). Even relatively small increases in the number of species detected through eDNA significantly improve the method's utility in these challenging conditions. Initially, KAPA.CTRL, our standard lab protocol used without Zymo cleanup, established a baseline with a total of 59 species detected across all sites. Upon introducing an inhibition removal step with Zymo (KAPA.ZYMO), we observed an increase in the mean number of fish species from 3 to 5 per sample, culminating in a total of 78 species. This shift highlights a significant enhancement, illustrating the effectiveness of Zymo cleanup in our settings. Further trials with Platinum SuperFi II were initially conducted without Zymo cleanup. However, due to poor amplification performance illustrating its high sensitivity to inhibitors, these trials did not progress to sequencing (data not shown). This necessitated the integration of Zymo cleanup steps in subsequent experiments. With the combined adjustments of Zymo cleanup and a refined PCR touchdown program using Platinum SuperFi II (Platinum.ZYMO), the mean number of fish species elevated to 7 per sample, with a total of 98 species detected across all sites—an overall increase of approximately 25.64% from the KAPA.ZYMO setup (Fig. 2a).

The impact of inhibition removal on ASV numbers was minimal (mean of 13 ASVs for KAPA.CTRL and 14 with KAPA.ZYMO); however, the comprehensive integration of Zymo with the Platinum protocol markedly increased the mean number of ASVs per sample to 46 (Fig. 2b).

When considering individual sites, each of the four locations showed an increase in fish detection, regardless of different environmental conditions in each site (Fig. 3). Specifically, the Jacques Cousteau (JC) site demonstrated a significant rise in the number of species from 29 with KAPA.ZYMO to 42 with Platinum.ZYMO, an increase of 44.83%. Similarly, the number of unique species at Padilla Bay (PB) increased from 17 to 26, marking a 52.94% rise and Mission-Aransas (MA) experienced a 34.78% increase, with species counts rising from 23 to 31. The San Francisco (SF) site saw a more modest increase of 25.00%, moving from 20 to 25 species. These enhancements can be attributed to the combined efficacy of Zymo cleanup and Platinum amplification, integrated with a modified touchdown PCR program, which collectively improved the detection of fish across these diverse estuarine environments.

The Mish-E primer was added in response to feedback from resource managers that elasmobranchs (rays and sharks) were not being detected in areas where they were known to be present. Although we did detect skate (family Rajidae) with the MiFish-U alone, the combined primers detected four additional elasmobranchii belonging to three different families: Dasyatidae (Dasyatidae spp., *Hypanus sabinus*), Rajidae (Rajidae spp.), and Rhinopteridae (Rhinoptera spp.).

Polymerase optimization and enhanced read counts

The Platinum.ZYMO method increased the detection of fish reads (see Methods *Fish Reads and ASVs*) by reducing off-target amplification, allowing more sequencing capacity to focus on target species. Figure 4a shows a strong correlation between raw reads and fish reads using the Platinum.ZYMO approach ($\beta = 0.620$, $p < 0.001$,

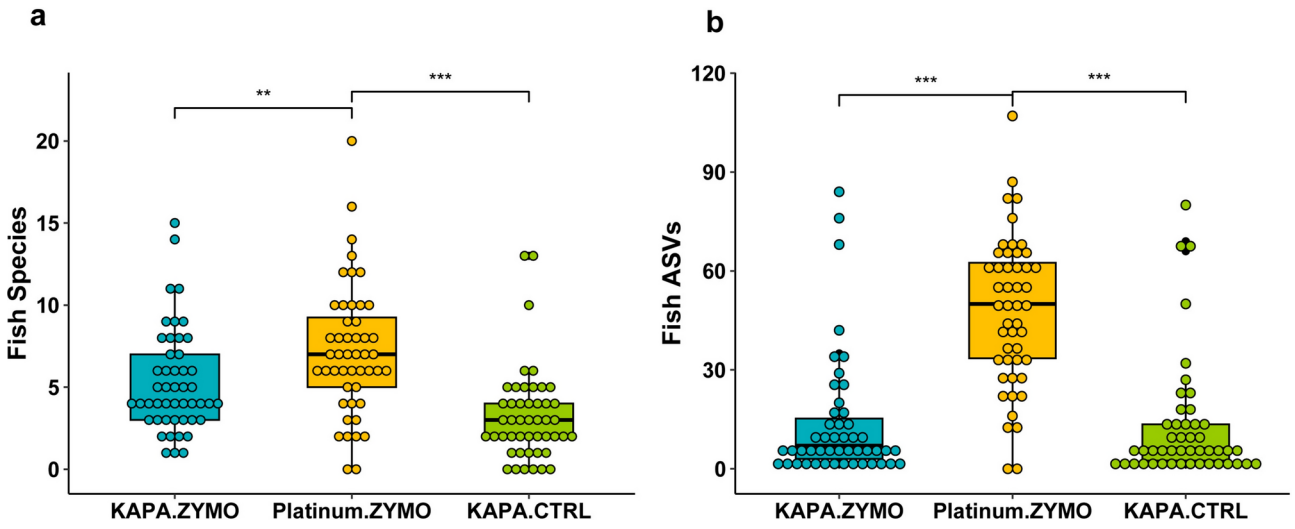

Fig. 2. The number of fish species and ASVs increases with Zymo cleanup and the Platinum PCR protocol. Inhibition removal increased the mean number of fish species, and the addition of both inhibition removal and the Platinum PCR protocol significantly increased both number of species and ASVs (p-value < 0.01 and < 0.001, respectively). The difference is also significant between Platinum.ZYMO and KAPA.CTRL. Asterisks in the figure indicate statistical significance, with ** for $p < 0.01$, and *** for $p < 0.001$. Figure 1. was generated using data from three separate E-gels, as indicated by the boundaries visible in the figure. The absence of a single full-length gel image is due to the original experimental setup involving separate gel runs.

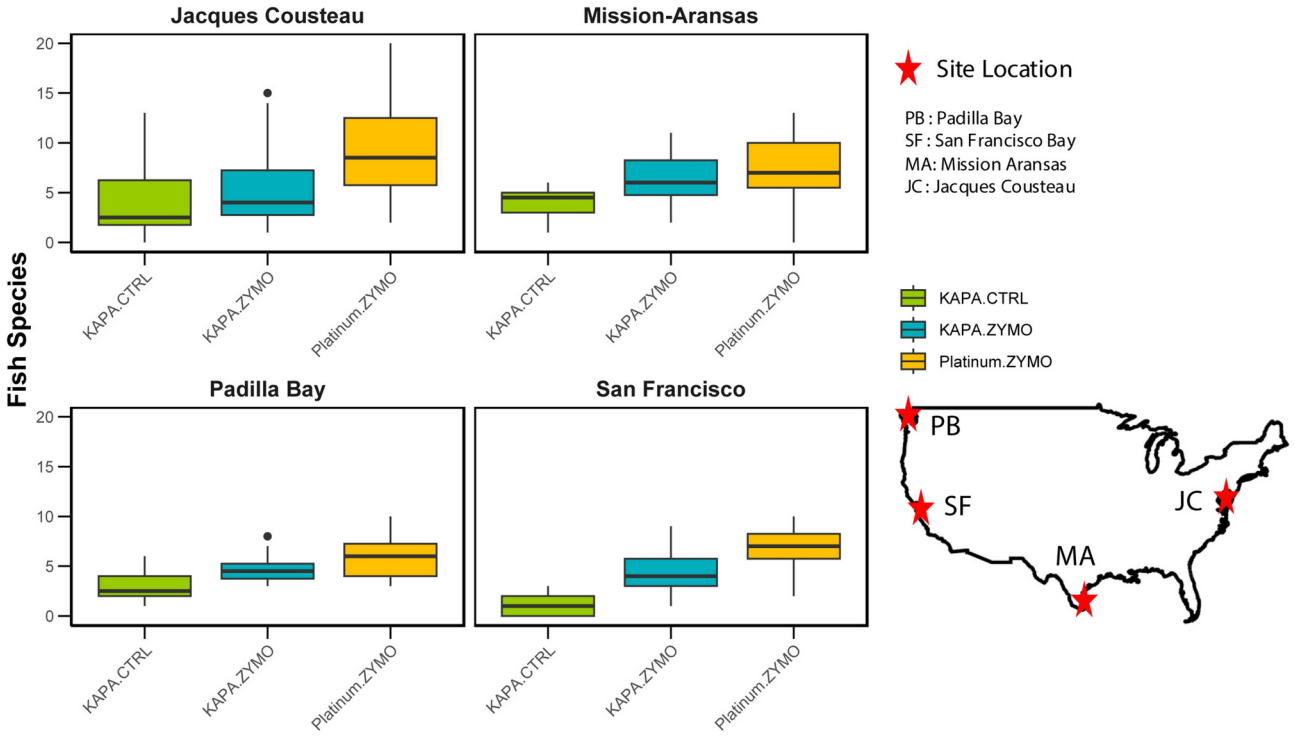

Fig. 3. Enhanced fish detection across individual Sites. Box plots comparing the distribution of fish species between the three tested protocols at four different estuaries (lower panel map), each location has different environmental conditions and expected species. Consistently, the use of platinum with Zymo cleanup resulted in a higher detection rate of fish species at every site. Statistical significance ($p < 0.05$) between treatments was assessed using Tukey HSD tests. Specifically, at Jacques Cousteau, Platinum.ZYMO was significantly different from both KAPA.CTRL and KAPA.ZYMO. At Mission-Aransas, KAPA.ZYMO and Platinum.ZYMO were significantly different from KAPA.CTRL. At Padilla Bay and San Francisco, all three treatments showed significant differences from one another.

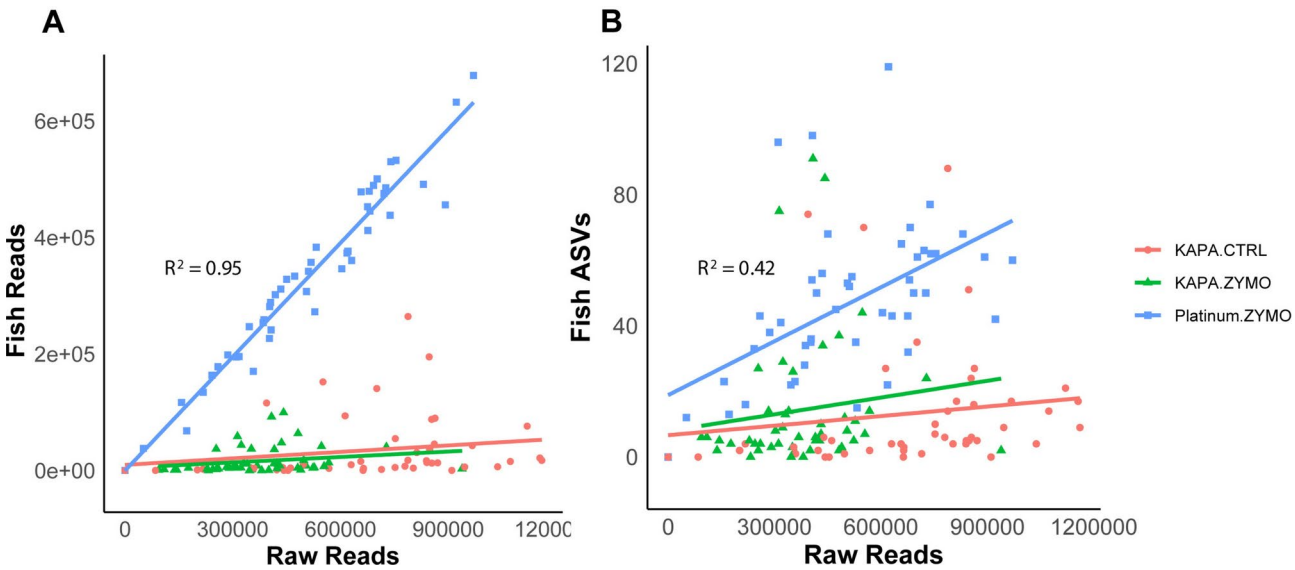

Fig. 4. Comparison Between Fish Reads and Raw reads. A linear regression analysis shows an increase in sequencing depth correlates with an increased detection of fish ASVs but not species. The gray shaded area represents the 95% Confidence Interval.

$R^2 = 0.950$). The correlation is less apparent, although still significant ($\beta = 0.000$, $p = 0.008$, R^2 of 0.420) for ASVs (Fig. 4b).

Reducing off target amplification and increasing fish reads has several positive implications for resource management; low abundance sequences are more likely to be amplified, which supports identification of rare species and increases detection sensitivity. Read counts for individual species are higher, allowing greater confidence in detections. Higher numbers of more diverse ASVs also raises the possibility of finer scale resolution of genetic diversity within species. Maximum Likelihood trees of unique fish ASVs detected in samples from San Francisco Bay estuary, suggests the highest diversity of unique ASVs was identified with the Platinum.ZYMO method (Fig. 5).

Challenges of ASV interpretation

The interpretation of ASVs presents unique challenges, particularly in distinguishing true biological signals from technical noise. The abundance of reads can reflect the nuances of extraction, PCR, and sequencing processes rather than actual biological abundance. Rare ASVs, which may appear only once across datasets, can represent

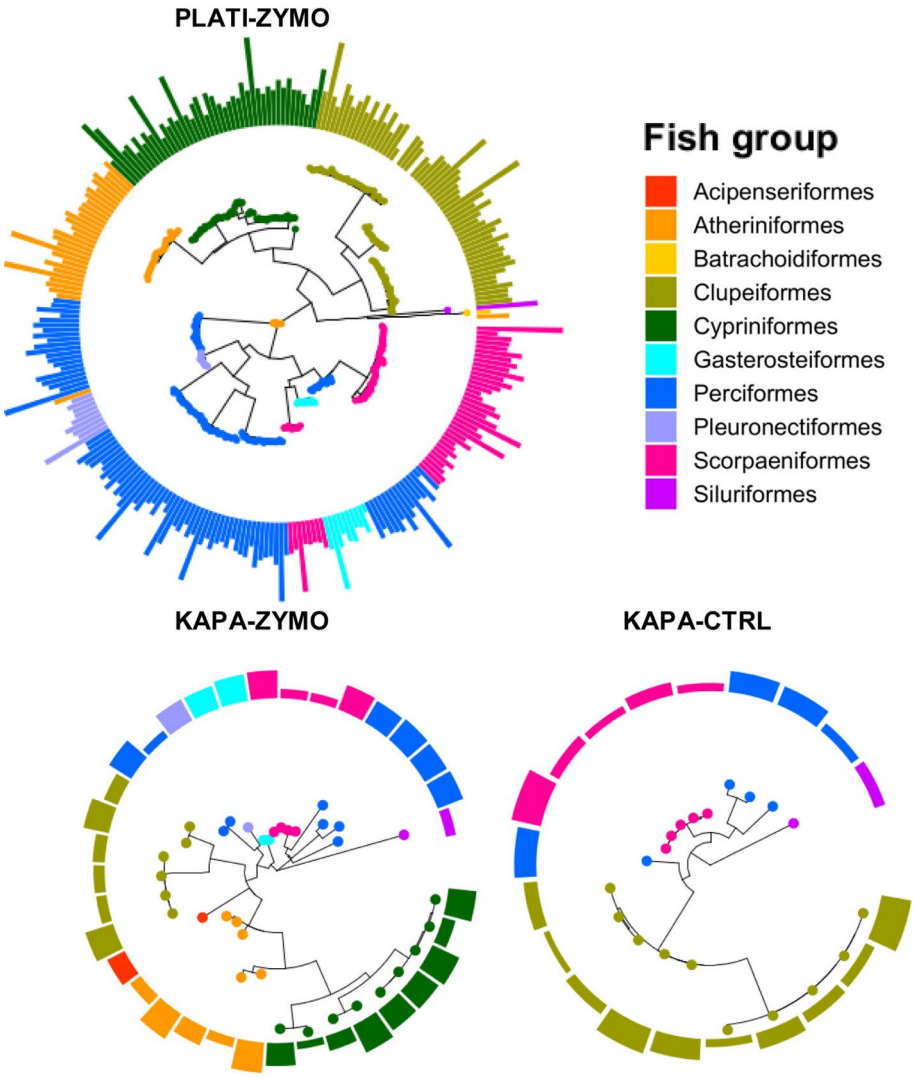

Fig. 5. Diversity of Fish ASVs Across San Francisco Bay. Maximum Likelihood (ML) trees depicting the diversity of unique fish ASVs detected across three different treatments: KAPA.CTRL (control without inhibitor removal), KAPA.ZYMO (with inhibitor removal), and Platinum.ZYMO (with Platinum SuperFi II polymerase and Zymo inhibitor removal). The tree tips represent individual ASVs, which are color-coded based on their taxonomic fish group. The trees are built using the ASVs detected from San Francisco Bay samples, with each treatment showcasing different ASV richness and distribution. Surrounding each tree, an outer ring displays bar plots of log10-transformed read counts, where each bar represents the abundance of a specific ASV. Bar colors correspond to the fish groups of the ASVs, visually illustrating the variation in read abundance and taxonomic representation between treatments. The Platinum.ZYMO protocol shows the highest diversity and read abundance of fish ASVs, reflecting the enhanced sensitivity and accuracy of this protocol in detecting fish eDNA in complex environments like San Francisco Bay.

genuinely rare biological entities, heteroplasmy[17], or PCR artifacts such as errors or stochastic variations in sample processing.

In our analysis, the high sensitivity of the Platinum SuperFi II protocol combined with Zymo cleanup (Platinum.ZYMO) enabled the detection of a broader array of fish ASVs, capturing true biodiversity more effectively (Fig. 5). For example, in San Francisco Bay, which was the most refractory site due to increased PCR inhibition, KAPA.ZYMO detected 40 ASVs assigned to 20 species, whereas Platinum.ZYMO yielded 284 unique ASVs assigned to 25 species (Fig. 5, ML Tree). Applying a cutoff threshold of 1% of fish reads to remove rare ASVs resulted in 26 ASVs for KAPA.ZYMO assigned to 19 species and 25 ASVs for Platinum.ZYMO assigned to 20 species. The 1% cutoff, however, represented only 40 reads for KAPA.ZYMO but 3,010 reads for Platinum.ZYMO.

By setting a minimum cutoff of 50 reads per ASV for Platinum.ZYMO, we successfully recovered all 25 fish species and the majority of ASVs (223). This demonstrates the method's robustness in detecting species while managing potential artifacts. Denoising methods like DADA2, which is an algorithm for resolving true biological sequences from noisy data, enhance the reliability of detected ASVs by removing errors and chimeras, thereby reflecting true biological variation.

While rare ASVs can represent genuinely rare taxa or intraspecific variations (such as haplotypes), it is also important to recognize that some may result from sequences not yet present in reference databases, as highlighted by Devloo-Delva et al.[18] and Zaiko et al.[19]. Recent updates to platforms like MitoFish and MiFish Pipeline have expanded reference datasets, improving species identification for many taxa, but database gaps remain (Zhu et al. 2023). These limitations underscore the need for careful interpretation of rare ASVs to avoid distorting community composition analyses. Bioinformatic processing using denoising methods, clustering, and establishing appropriate cutoff thresholds remain critical steps in validating the authenticity of rare ASVs.

Enhancing taxonomic resolution in eDNA studies through ASVs

Recent studies in microbial ecology comparing ASVs and OTUs, such as those by García-López et al.[20] and Chiarello et al.[21], highlight the value of ASVs in providing higher taxonomic resolution and accuracy, especially when appropriate filtering steps are applied. These studies indicate that while OTUs can artificially inflate diversity by including spurious low-abundance sequences, ASVs offer a more accurate and reliable depiction of community composition, provided that denoising and abundance filtering are rigorously applied.

Stoeckle et al.[22] further validate the use of ASVs by demonstrating that reproducible detection of fish species in eDNA samples can be achieved with higher sensitivity, especially for rare species, when using an ASV approach. They found that rare eDNA, representing species with low abundance, can often fall below detection thresholds when using traditional methods. However, ASVs, particularly when coupled with internal standards or deeper sequencing, allow for the recovery of these rare species. The study emphasizes the importance of amplifying a larger proportion of DNA samples to overcome issues of low-copy-number templates, highlighting the value of ASVs in ensuring a more comprehensive and accurate representation of biodiversity. Moreover, Stoeckle et al.[22] underscore the challenge of non-target DNA (e.g., bacterial sequences) distorting species detection. They demonstrate that using ASVs, along with filtering and denoising approaches, can mitigate these effects by focusing on high-quality, species-specific sequences, thereby improving species-level identification and quantifying eDNA abundance with higher precision.

Furthermore, the application of ASVs allows us to capture intraspecific variation, such as haplotypes, which can be overlooked when using OTU clustering based on fixed identity thresholds (e.g., 97%). This increased resolution is critical for understanding fine-scale genetic diversity within species, adding another layer of biological insight that OTUs often miss—a point underscored by both Miya et al.[23] and our findings. Thus, while OTUs have traditionally been used in fish eDNA metabarcoding, the growing body of evidence supports a shift toward ASV-based approaches. ASVs not only provide more accurate species-level identifications but also offer greater insights into genetic diversity, which are essential for monitoring biodiversity in dynamic ecosystems like estuaries. By adopting ASVs in our study, we enhance the reliability and resolution of species detection, ultimately contributing to more effective conservation and management strategies.

PCR replicates

Replicate samples may be collected or generated at any point in the eDNA workflow, from field sampling to sequencing replicates. Multiple studies have highlighted the benefits of replicates in enhancing the reliability of biological data while also recognizing that additional samples increase the cost of analyses[24–26]. Balancing the value of the additional information against budget and time constraints is crucial for researchers and managers. In this protocol, we recommend conducting triplicate PCRs for each sample, which are then pooled back into one sample for sequencing. Our testing of PCR replicates (see Methods section) in this sample set confirms earlier findings[27]: increasing PCR replicates substantially improves the detection of fish species, supporting data robustness and reducing the likelihood of spurious results due to PCR variability (**Fig. S3**, Supplementary Information).

Application to other water systems

Although this workflow was designed to improve fish detection in estuarine environments, we also evaluated its utility in different water systems. For this purpose, we tested an additional data set collected from freshwater streams in New England. These samples were collected and processed using similar methodologies to our estuarine samples, ensuring consistency in testing. In these 32 samples, where turbidity and inhibition have not been noted, the new approach detected a significantly higher number of fish ASVs (Fig. 6, $p\text{-value} < 0.001$) with an average of 16 for the test control (KAPA.CTRL) and 34 with the new protocol (Platinum.ZYMO). Those ASVs

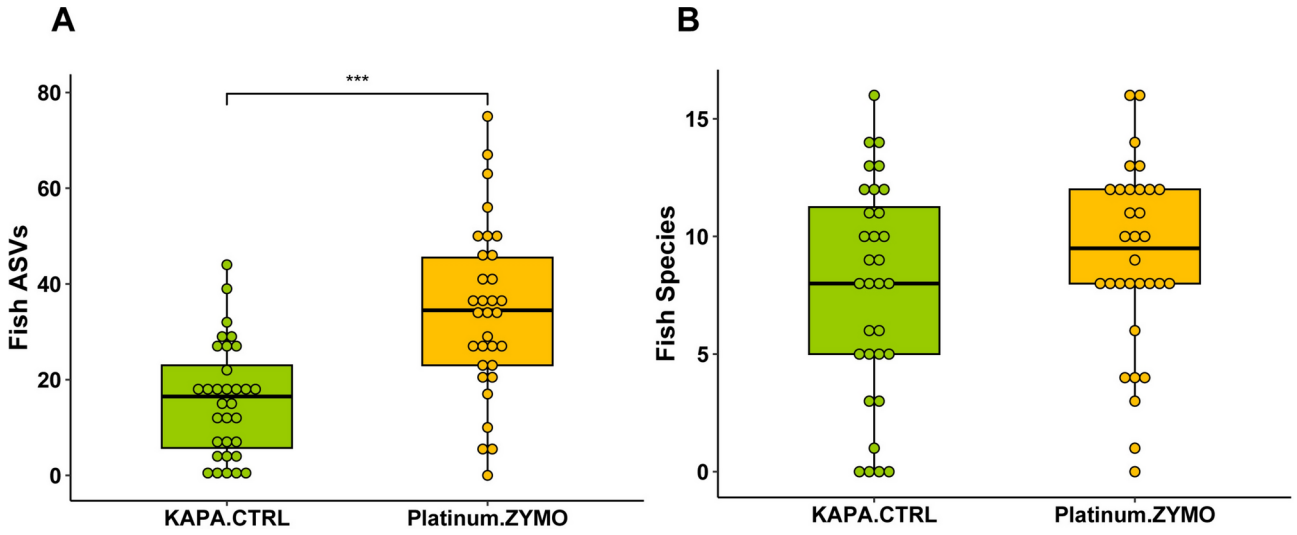

Fig. 6. Enhanced fish detection in stream waters. Box plots comparing the distribution of fish ASVs and species between the control (KAPA.CTRL) and the new protocol (Platinum.ZYMO) on stream water samples. Significant p-values are indicated by asterisks with *** for $p < 0.001$.

were assigned to 41 species for KAPA.CTRL and 46 for Platinum.ZYMO. This may be useful when surveying rare species or other organisms where amplifying read counts could enhance detection.

Enhancing fish eDNA efficiency in complex environments

Our study extends the foundational work of Miya et al.[4] by addressing the ongoing challenge of non-target amplification in eDNA metabarcoding, particularly in estuarine environments characterized by high turbidity and PCR inhibitors. While Miya et al.'s MiFish primers have been crucial for broad fish species detection, their work also highlighted the frequent amplification of non-target bacterial sequences, which can reduce detection efficiency. Our findings demonstrate that the use of Platinum SuperFi II polymerase, as similarly shown by Kawato et al.[5] in deep-sea environments, significantly enhances target fish DNA amplification while minimizing off-target reads. This reduction in non-target amplification is particularly important for estuarine systems, where high levels of bacterial and organic matter often complicate eDNA analysis.

In addition to Platinum SuperFi II polymerase, we employed a touchdown PCR approach, and a dual strategy of DNA dilution combined with an inhibitor removal kit (Zymo OneStep™). The touchdown PCR technique, which gradually reduces the annealing temperature, improved the specificity of amplification and reduced non-specific amplifications, allowing for more efficient targeting of low-concentration fish DNA in the presence of inhibitors. Meanwhile, the combination of DNA dilution with the Zymo inhibitor removal kit was instrumental in mitigating the effects of PCR inhibitors that are prevalent in highly turbid environments, like San Francisco Bay. This combined approach not only prevented PCR inhibition but also optimized amplification efficiency, reducing interference from environmental contaminants and improving the quality of detected fish DNA.

Our results support findings by Kawato et al.[5] and Miya et al.[23] who emphasized the necessity of robust protocols that address environmental inhibitors in eDNA analysis. The use of both inhibitor removal and touchdown PCR in our workflow increased species detection by over 25% compared to traditional methods. This improvement underscores the effectiveness of integrating advanced techniques—such as inhibitor removal, high-fidelity polymerases, and optimized PCR strategies—into eDNA workflows for challenging environments. Our findings highlight the importance of refining each step of the eDNA process, from extraction to amplification, to enhance the sensitivity and reliability of fish species detection, particularly in complex and inhibitor-rich environments like estuaries.

Conclusions and recommended workflow

The final recommended workflow (Fig. 7) includes a bead-based extraction performed on a robotic system, inhibition removal, multiplexed primers with Platinum HiFi polymerase and triplicate PCR (**Fig. S3**). The addition of the Zymo cleanup, a more expensive polymerase and triplicate PCR increases the overall costs; researchers and managers should consider project goals and site conditions when selecting a laboratory protocol. Our goal was to combine existing, proven methods into a protocol that optimizes detection of fish species in estuaries. All the components of this workflow have been used in other studies, and the steps are expected to be familiar and readily adapted by labs conducting eDNA work. The PCR steps can be completed by hand or on a robotic system, allowing high throughput processing of large sample numbers.

This workflow optimizes fish detection with the MiFish 12S primer. Adding additional primers targeting other mitochondrial regions[28,29] would likely increase the number of species detected.

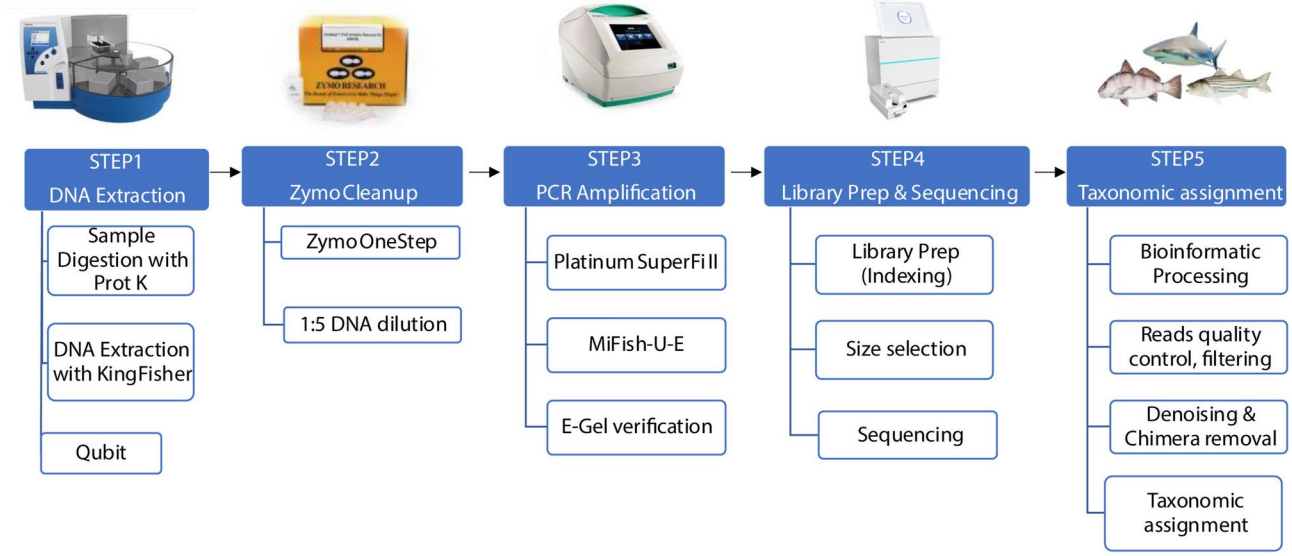

Fig. 7. Recommended workflow. An overview of this workflow is discussed in the methods section, and detailed extraction and PCR protocols can be accessed through protocols.io[32,34].

Protocol	Inhibition removal	DNA Dilution	PCR Touchdown	Primers
KAPA.CTRL	None	1:5	2nd step: 55C	MifFish-U
KAPA.ZYMO	Zymo	1:5	2nd step: 55C	MiFish-U
Platinum.ZYMO	Zymo	1:5	2nd step: 60C	MiFish-U MiFish-E

Table 1. PCR workflows evaluated in this study.

Methods

Method validation

To validate our workflow, we tested the automated DNA extraction and the three PCR protocols in Table 1 on samples collected from four estuaries. Forty-eight samples were collected during the same period (May 2023) from four different locations within each estuary (12 from each site) to ensure that seasonal or time-related variables do not confound across-site comparisons.

We then reviewed the results to evaluate the steps that provide the greatest benefit in terms of number of fish species and fish ASVs. For clarity and consistency, we use the term species to refer to both cases where ASVs were matched to the species level and instances where ASVs could not be matched to the species level but were assigned to a higher taxonomic rank, such as genus (e.g. the genus *Fundulus*). All taxonomic assignments are based on a lowest common ancestor approach, where sequences are matched to the most specific taxonomic level possible using the available reference database.

Field collection

The samples used in this study were collected from a range of NERRS sites (see above) as part of a pilot project incorporating eDNA into a long-term multi-site monitoring network. All the samples were 1-L water samples, collected from the surface, then transported to a nearby lab for filtering. Each sample was filtered through 1.5um porosity glass fiber filters and filters were replaced if clogging occurred. The filters were stored in pre-prepared tubes containing 4 ml Longmire's buffer (Longmire et al. 1997) to preserve the samples. Up to 3 filters from the sample were included in one tube.

Laboratory analyses and sequencing

All lab work was conducted in dedicated Biosafety Level 2 (BSL2) molecular laboratory facilities at the University of New Hampshire. All steps from sample extraction through sequencing were performed in separate, dedicated spaces within the building. We adhered strictly to standard laboratory cleaning practices for BSL2 including the separation of Pre- and Post-PCR laboratory spaces. PCR Mix preparation was done in one designated area, while DNA template addition was performed in a physically separate space to minimize the risk of contamination. All experiments were conducted under laminar flow hoods and Biosafety cabinets for sample preparation, DNA extraction, purification and amplification. All surfaces were treated with 70% ETOH, 10% bleach and dH2O and were exposed to UV light before and after experimentation.

DNA extraction

Extraction Method comparison

Bead-based extraction methods have been successfully applied in many eDNA studies[30,31]. We adapted this technique for use on a KingFisher Flex System (Thermo Fisher, Waltham, MA). We compared the Kingfisher protocol[32] to a Qiagen-based extraction which we have used in previous studies[33] to evaluate effectiveness.

KingFisher automated extraction: Sample vials containing filters in 400 ml Longmire's buffer were received and entered into the lab tracking system. 10 µl of Proteinase K (0.02X) was mixed with 90 µl of Longmire buffer (LM) and added to each sample vial bringing the total volume to 500 ml. The mixture was incubated at 56 °C for 90 min. For DNA isolation, 320 µl of Illumina SPRI paramagnetic beads (CAT: 20,060,057) were added to the lysed samples to ensure a 1:8 sample-to-beads ratio. DNA extraction was performed using The KingFisher Flex System (Thermo Fisher, Waltham, MA), which included two 80% ethanol washes, followed by elution in 100 µl of 10 mM Tris–HCl. Our detailed protocol is available on protocol.io[32].

Qiagen-based extraction: Filters were placed in a lyse and spin basket with 400 ml of buffer ATL and 20 µl of proteinase K and incubated at 56 ∘C for one hour, then centrifuged. The remainder of the filter extraction was performed on a QIAcube Connect system (QIAGEN®, Hilden, Germany) following the QIAamp DNA Mini protocol (Qiagen Cat. 51,326). This kit is designed for extraction of DNA from tissue and is readily adapted to the QIAcube automated system. Although the reagents are proprietary the manufacturer notes that the method removes inhibition and contaminants.

Samples from both methods were eluted to 100 µl. DNA concentration was determined using the Qubit dsDNA HS Assay kit (Thermo Fisher, Waltham, MA), per the manufacturer's instructions. To monitor contamination, extraction blanks, consisting of H2O and plain buffer without filter, were extracted with each sample set. Once extracted, 50 µl of the sample was diluted 1:5 and stored in a -20C freezer for use in this project. The remaining sample was archived in a -80C freezer for potential future use.

Inhibition Removal. ZYMO OneStep-96 PCR Inhibitor Removal Kit (CAT: D6035) was used to remove inhibition from all samples following the manufacturer's instructions. Briefly, for 96 well plates, 150 µl of Prep Solution were added to each Silicon-A™-HRC Plate well, incubated for 5 min, and then centrifuged at 3,500 × g for 5 min. Subsequently, 50 µl of DNA was added to the prepared plate, mounted on a new Elution Plate, and centrifuged at 3,500 × g for 3 min. The filtered DNA was used for subsequent PCR reactions.

PCR-optimization

Initially we tested a variety of published PCR protocols as well as several polymerases and inhibitors on samples from these sites. We then selected the most promising alternatives for more detailed testing. The different PCR workflows we evaluated are summarized in Table 1 below. All the PCR comparisons were performed on samples extracted using the automated Kingfisher extraction method.

This approach builds on the foundations laid by Kawato and Miya[4,5], and others to incorporate automated DNA extraction, and advanced PCR design to significantly reduce off-target amplification, which consists mainly of bacterial sequences[5,8], and enhance PCR efficiency and reliability in the presence of inhibitors. The overall goal of this work is to improve eDNA-based fish species detection from problematic turbid samples collected in estuarine environments.

PCR Verification Amplification of target regions was verified on 2% E-Gel electrophoresis (Thermo Fisher, Waltham, MA, CAT: G820802). Absent or muted bands in samples with verified DNA concentrations indicate that the samples are likely to contain inhibiting agents. E-Gel images are also used as an initial screen for successful amplification of positive controls, and absence of contamination in negative controls.

PCR Amplification—Platinum. We found that the touchdown profile reduced off target amplification, but for many sites the target band was still effectively hidden by amplification of bacteria. To increase target amplification, we used Platinum SuperFi II polymerase. The Platinum enzyme was designed for higher fidelity and specificity and increased resistance to PCR inhibitors and to have a universal primer annealing temperature of 60 °C which allows multiplexing PCR. However, we found that some samples amplified using the Platinum polymerase still exhibited inhibition. In these cases, adding an additional inhibition removal step improved amplification. Amplification Protocol: The PCR reaction was established in a total volume of 20 µl, incorporating 10 µl of Platinum SuperFi II Master Mix (2X), 2 µl of each equimolar primer mix (MiFish-U and MiFish-E, 5 uM each), 4 µl of genomic DNA, and 2 µl of PCR-grade H2O. Nextera adapters were added to each primer. The amplification followed a Touchdown (TD) PCR strategy to reduce non-specific amplification. The protocol began with a lid temperature set at 105 °C, maintaining a reaction volume of 20 µl. The initial denaturation step was at 95 °C for 3 min, followed by a cycling phase starting at 94 °C for 30 s, an annealing step at 69.5 °C decreasing by 1.5 °C per cycle for 30 s, and an extension at 72 °C for 1 min and 30 s for 13 cycles. This was succeeded by an additional cycling phase: 94 °C for 30 s, 60 °C for 30 s, and 72 °C for 45 s for 25 cycles, concluding with a final extension at 72 °C for 10 min and a hold at 4 °C indefinitely. A step-by-step protocol in available on protocols.io[32,34].

PCR replicates

The value of PCR replicates on fish species detection, was tested with the optimized protocol incorporating Platinum and Zymo clean up. Samples from each of the four sites were tested. For each sample, three replicate PCRs were performed. These replicates were then sequenced as separate samples.

PCR Amplification KAPA. Our initial protocol used KAPA HiFi HotStart (KAPA) and a touchdown (TD) thermal profile modified from Pitz et al. (2020). The cycling conditions were the same as described for Platinum above except that the annealing temperature of the second cycling phase was set to 55 °C. The PCR reaction was established in a total volume of 12 µl, incorporating 6 µl of KAPA Master Mix (2X), 0.7 µl of the forward and 0.7 µl of the reverse MiFish-U primers, 2 µl of genomic DNA, and 2.6 µl of PCR-grade H2O.

Primers. The MiFish-U primer was designed to amplify bony fish species while the MiFish-E[4] targets elasmobranch (sharks and rays). Both types are important in estuarine systems, and multiplexing these primers expands the list of potential target species.

The primer and adapter sequences used in this study are listed below.

Primer & adapter sequences used in this study

MiFish-MIX-F (forward)

MiFish-U-F: 5′- TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGGTCGGTAAAACTCGTGCCAGC-3′.

MiFish-E–F: 5′- TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGGTTGGTAAATCTCGTGCCAGC-3′.

MiFish-MIX-R (reverse)

MiFish-U-R: 5′- GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGCATAGTGGGGTATCTAATCCCAGTTTG-3′.

MiFish-E-R: 5′- GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAG-CATAGTGGGGTATCTAATCCTAGTTTG-3′.

Library preparation and sequencing

PCR products were submitted to the Hubbard Center for Genome Studies (HCGS) for library preparation and sequencing, on an Illumina Novaseq 6000 instrument. Briefly, the amplicon products from the 1st PCR were prepared by incorporating dual-index barcodes and Illumina sequencing adapters (P5 and P7) into the DNA fragments. The PCR amplification conditions were set as follows: initial denaturation at 94 °C for 3 min, followed by 15 cycles of denaturation at 94 °C for 20 s, annealing at 55 °C for 30 s, and extension at 72 °C for 15 s. A final extension phase was conducted at 72 °C for 7 min to ensure complete amplification of target sequences. The amplified libraries were purified using BluePippin to selectively remove short fragments and primer dimers.

Sequence Processing and Taxonomic Assignment. Fastp v 0.23.0 was used to trim the poly-g tails and low-quality sequences (mean quality < 25, window size 2 bp) from the raw, demultiplexed fastq files[35]. The demultiplexed fastqs were then imported to qiime2 (qiime2-2022.8) for adapter trimming, denoising, and taxonomic classification[36]. The cutadapt plugin in qiime2 was used to trim primer sequences and filter out read-pairs that do not contain the priming sequences[37]. Read-pairs were retained if the MiFish primer sequences matched the allowed error rate of 0.1. Retained read-pairs were then denoised, merged, and checked for chimeras with the DADA2 qiime2 plugin[38]. Forward and reverse reads were truncated to 110 and 105 base-pairs respectively and were required to have at least 12 base-pairs (qiime2 default) of overlap to merge the forward and reverse reads. To reduce the number of false positive unique ASVs, we used the 'pseudo' pooling method where samples are initially denoised independently, and then denoised a second time with prior knowledge of the ASVs that are found in at least two samples in the initial denoising step. We then classified the ASVs with the 'classify-consensus-vsearch' and the 'sci-kit learn' feature classifiers in qiime2 using the MitoFish database[6]. We ran feature classification at default parameters except we set the number of accepted reference sequences (percent-identity) to 90% (default is 80%), and the number of accepted sequences to 'all' instead of the top 10 (maxaccepts). The BLAST search results were retained for confirming classifications for each ASV.

Fish reads and ASVs

In this study, "fish reads" refer to the total number of sequencing reads assigned to fish species after the raw data are processed and filtered. Fish reads provide a measure of how much DNA from fish species is present in the sample. In contrast, fish ASVs represent unique genetic sequences identified in the sample. These ASVs allow us to differentiate between individual species and, in some cases, variations within species.

Phylogenetic analysis

ASVs identified from each treatment were aligned using the MAFFT software[39]. The alignment was optimized under the E-INS-i algorithm using a gap open penalty of 1.5 and an offset value of 0.1, with the scoring matrix fine-tuned using a kappa value of 50. These settings were chosen to enhance the alignment accuracy of closely related fish ASVs. The aligned ASVs were used to reconstruct phylogenetic trees through the maximum likelihood method implemented in RAxML[40]. The trees were generated under the General Time Reversible (GTR) model with a gamma distribution to accommodate rate variation across sites. This modeling choice is particularly effective for sequences exhibiting high variability, as is common with environmental DNA samples from diverse fish populations.

Statistical analyses

A one-way Analysis of Variance (ANOVA) was performed separately for Fish species and Fish ASVs using the `aov` function in R to determine if the observed differences using our optimized protocol were statistically significant from the other treatments. The null hypothesis for each ANOVA was that the means of Fish species and Fish ASVs for all treatment groups were equal.The primary objective of the statistical analysis was to evaluate the effects of different treatments on both Fish species and Fish ASVs. The treatments included were KAPA.CTRL (control), KAPA.ZYMO (with Zymo clean-up), and Platinum.ZYMO (with Zymo clean-up). The p-values for the treatment comparisons were extracted from the ANOVA results. If the p-value was less than 0.05, indicating significant differences among the treatment groups, a post-hoc Tukey's Honest Significant Difference (HSD) test was conducted using the 'TukeyHSD' function in R to identify which specific treatment groups differed from each other. All statistical analyses were conducted using R software (version 4.3.2) with the 'aov' function for ANOVA and the 'TukeyHSD' function for post-hoc comparisons. To further validate our optimized protocol on data collected from stream waters in New England. 1000 ml of water per sample were filtered using a 0.45 µm

Whatman cellulose nitrate filters and were processed in the same way for DNA extraction and PCR as our four NERRS sites. For these samples we used a T-test to determine if there was a statistically significant difference between the control treatment "KAPA.CTRL" and the optimized protocol "Platinum.ZYMO".

Additionally, we conducted a linear regression analysis to evaluate the relationship between sequencing depth (total raw reads) and the detection of fish ASVs and species. This analysis was performed using the lm function in R (version 4.3.2), with fish ASVs and species regressed against the total number of raw reads. To estimate the precision of the regression line, a 95% Confidence Interval (CI) was calculated using the predict() function in R, and the confidence interval was visualized with the ggplot2 package. This approach allowed us to assess whether increased sequencing depth is associated with improved detection of fish ASVs or species.

Data availability

The raw datasets used and analyzed in this study are available on the NCBI Sequence Read Archive (SRA) under the project number PRJNA1136877. Supplementary data (Version v3), including ASVs and comprehensive lists of detected fish species, can be accessed on Zenodo at https://doi.org/https://doi.org/10.5281/zenodo.12753119.

Received: 15 July 2024; Accepted: 1 January 2025
Published online: 07 January 2025

References

1. Blackman, R. et al. Environmental DNA: The next chapter. *Mol. Ecol.* https://doi.org/10.1111/mec.17355 (2024).
2. Nagarajan, R. P. et al. Environmental DNA methods for ecological monitoring and biodiversity assessment in estuaries. *Estuar. Coast.* **45**, 2254–2273. https://doi.org/10.1007/s12237-022-01080-y (2022).
3. Zhang, S., Zhao, J. & Yao, M. A comprehensive and comparative evaluation of primers for metabarcoding eDNA from fish. *Method. Ecol. Evol.* **11**, 1609–1625 (2020).
4. Miya, M. et al. MiFish, a set of universal PCR primers for metabarcoding environmental DNA from fishes: Detection of more than 230 subtropical marine species. *R. Soc. Open Sci.* **2**(7), 150088 (2015).
5. Kawato, M. et al. Optimization of environmental DNA extraction and amplification methods for metabarcoding of deep-sea fish. *MethodsX* **8**, 101238 (2021).
6. Zhu, T., Sato, Y., Sado, T., Miya, M. & Iwasaki, W. MitoFish, MitoAnnotator, and MiFish Pipeline: Updates in 10 years. *Mol. Biol. Evol.* https://doi.org/10.1093/molbev/msad035 (2023).
7. Pitz, K., Truelove, N., Nye, C., Michisaki, R. P. & Chavez, F. Environmental DNA (eDNA) 12S Metabarcoding Illumina MiSeq NGS PCR Protocol (Touchdown) V.2. https://doi.org/10.17504/protocols.io.bcppivmn (2020).
8. Stoeckle, M. Y. et al. Current laboratory protocols for detecting fish species with environmental DNA optimize sensitivity and reproducibility, especially for more abundant populations. *ICES J. Mar. Sci.* **79**, 403–412 (2022).
9. Holmes, A. E. et al. Evaluating environmental DNA detection of a rare fish in turbid water using field and experimental approaches. *PeerJ* **12**, e16453 (2024).
10. Kumar, G., Farrell, E., Reaume, A. M., Eble, J. A. & Gaither, M. R. One size does not fit all: Tuning eDNA protocols for high- and low-turbidity water sampling. *Environ. DNA* **4**, 167–180 (2022).
11. Hunter, M. E., Ferrante, J. A., Meigs-Friend, G. & Ulmer, A. Improving eDNA yield and inhibitor reduction through increased water volumes and multi-filter isolation techniques. *Sci. Rep.* https://doi.org/10.1038/s41598-019-40977-w (2019).
12. Schrader, C., Schielke, A., Ellerbroek, L. & Johne, R. PCR inhibitors - occurrence, properties and removal. *J. Appl. Microbiol.* **113**, 1014. https://doi.org/10.1111/j.1365-2672.2012.05384.x (2012).
13. Ferner, M. C. (eds). A Profile of the San Francisco Bay National Estuarine Research Reserve (2011).
14. Evans, A., Madden, K. & Morehead Palmer, S. THE ECOLOGY AND SOCIOLOGY OF THE MISSION-ARANSAS ESTUARY AN ESTUARINE AND WATERSHED PROFILE (2012).
15. H. Lee Edmiston. *A River Meets the Bay The Apalachicola Estuarine System* (2008).
16. Kennish, M. J. & Lutz, R. A. *CHARACTERIZATION OF THE JACQUES COUSTEAU NATIONAL ESTUARINE RESEARCH RESERVE* (2004).
17. Rensch, T., Villar, D., Horvath, J., Odom, D. T. & Flicek, P. Mitochondrial heteroplasmy in vertebrates using ChIP-sequencing data. *Genome. Biol.* https://doi.org/10.1186/s13059-016-0996-y (2016).
18. Devloo-Delva, F. et al. How does marker choice affect your diet analysis: comparing genetic markers and digestion levels for diet metabarcoding of tropical-reef piscivores. *Mar. Freshw. Res.* **70**, 8–18 (2019).
19. Zaiko, A. et al. Towards reproducible metabarcoding data: Lessons from an international cross-laboratory experiment. *Mol. Ecol. Resour.* **22**, 519–538 (2022).
20. García-López, R. et al. Otus and asvs produce comparable taxonomic and diversity from shrimp microbiota 16s profiles using tailored abundance filters. *Genes. (Basel)* **12**(4), 564 (2021).
21. Chiarello, M., McCauley, M., Villéger, S. & Jackson, C. R. Ranking the biases: The choice of OTUs vs. ASVs in 16S rRNA amplicon data analysis has stronger effects on diversity measures than rarefaction and OTU identity threshold. *PLoS One* **17**(2), e0264443 (2022).
22. Stoeckle, M. Y., Ausubel, J. H. & Coogan, M. 12S gene metabarcoding with DNA standard quantifies marine bony fish environmental DNA, identifies threshold for reproducible detection, and overcomes distortion due to amplification of non-fish DNA. *Environ. DNA* https://doi.org/10.1002/edn3.376 (2024).
23. Miya, M., Gotoh, R. O. & Sado, T. MiFish metabarcoding: A high-throughput approach for simultaneous detection of multiple fish species from environmental DNA and other samples. *Fish. Sci.* **86**, 939–970. https://doi.org/10.1007/s12562-020-01461-x (2020) (**Preprint at**).
24. Macher, T. H. et al. Beyond fish edna metabarcoding: Field replicates disproportionately improve the detection of stream associated vertebrate species. *Metab. Metag.* **5**, 59–71 (2021).
25. Shirazi, S., Meyer, R. S. & Shapiro, B. Revisiting the effect of PCR replication and sequencing depth on biodiversity metrics in environmental DNA metabarcoding. *Ecol. Evol.* **11**, 15766–15779 (2021).
26. Stauffer, S. et al. How many replicates to accurately estimate fish biodiversity using environmental DNA on coral reefs?. *Ecol. Evol.* **11**, 14630–14643 (2021).
27. Miya, M. Environmental DNA metabarcoding: A novel method for biodiversity monitoring of marine fish communities. *Annu. Rev. Mar. Sci.* https://doi.org/10.1146/annurev-marine-041421 (2021).
28. Riaz, T. et al. EcoPrimers: Inference of new DNA barcode markers from whole genome sequence analysis. *Nucl. Acid. Res.* **39**(21), e145-e145 (2011).
29. Evans, N. T. et al. Fish community assessment with eDNA metabarcoding: Effects of sampling design and bioinformatic filtering. *Canadi. J. Fish. Aquat. Sci.* **74**, 1362–1374 (2017).

30. Guthrie, A. M., Nevill, P., Cooper, C. E., Bateman, P. W. & van der Heyde, M. On a roll: a direct comparison of extraction methods for the recovery of eDNA from roller swabbing of surfaces. *BMC Res. Notes* https://doi.org/10.1186/s13104-023-06669-5 (2023).
31. Sanches, T. M. & Schreier, A. D. Optimizing an eDNA protocol for estuarine environments: Balancing sensitivity, cost and time. *PLoS One* **15**(5), e0233522 (2020).
32. El Baidouri F. et al. Automated eDNA Extraction from Estuarine Samples Using Magnetic Beads. (2024) https://doi.org/10.17504/protocols.io.5jyl82jn9l2w/v1.
33. Crane, L. C., Goldstein, J. S., Thomas, D. W., Rexroth, K. S. & Watts, A. W. Effects of life stage on eDNA detection of the invasive European green crab (Carcinus maenas) in estuarine systems. *Ecol. Indic.* **124**, 107412 (2021).
34. El Baidouri, F., Gilbert, H. L., Watts, A., El, F. & Unh, B. 12S PCR Metabarcoding Protocol for Fish Detection in Estuarine Samples. https://doi.org/10.17504/protocols.io.3byl49wqogo5/v1. (2024)
35. Chen, S., Zhou, Y., Chen, Y. & Gu, J. Fastp. An ultra-fast all-in-one FASTQ preprocessor. In *Bioinformatics,* vol. 34 i884–i890 (Oxford University Press, 2018).
36. Bolyen, E. et al. Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. *Nat. Biotechnol.* https://doi.org/10.1038/s41587-019-0209-9 (2019) (**Preprint at**).
37. Martin, M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet. J.* **17**(1), 10–12 (2011).
38. Callahan, B. J. et al. DADA2: High-resolution sample inference from Illumina amplicon data. *Nat. Method.* **13**(7), 581–583 (2016).
39. Katoh, K. & Standley, D. M. MAFFT multiple sequence alignment software version 7: Improvements in performance and usability. *Mol. Biol. Evol.* **30**, 772 (2013).
40. Stamatakis, A. RAxML version 8: A tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* **30**, 1312 (2014).

Acknowledgements

We thank the NERRS Research Collaborative, NOAA, specifically the following NERRS sites, San Francisco Bay, Padilla Bay, Mission Aransas and Jacques Cousteau. Bioinformatic analyses were supported by New Hampshire-INBRE through an Institutional Development Award (IDeA), P20GM103506, from the National Institute of General Medical Sciences of the NIH.

Author contributions

A.W., F.E.B. conceived the experiment, designed the methodology and led the writing of the manuscript; F.E.B conducted the experiments and optimization steps. M.K. and W.K.T conducted the initial DNA extraction optimization. J.M. and J.S. conducted the bioinformatic analyses. H.G. conducted initial sample processing. A.W., F.E.B, J.M. analyzed data. J.S. and W. K.T analyzed stream data. All authors contributed critically to the drafts and gave their final approval for publication.

Declarations

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-025-85176-y.

Correspondence and requests for materials should be addressed to F.E.B. or A.W.W.