

## SPECIAL ISSUE ARTICLE

# 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

Mark Y. Stoeckle  | Jesse H. Ausubel | Michael Coogan

Program for the Human Environment,  
The Rockefeller University, New York,  
New York, USA

**Correspondence**

Mark Y. Stoeckle, Program for the Human  
Environment, The Rockefeller University,  
1230 York AVE, New York, NY 10065,  
USA.

Email: [mark.stoeckle@rockefeller.edu](mailto:mark.stoeckle@rockefeller.edu)

**Funding information**

National Oceanic and Atmospheric  
Administration, Grant/Award Number:  
NA20OAR01104XX; US Fish and  
Wildlife Service Wildlife and Sport Fish  
Restoration Program

**Abstract**

Single-species PCR assays accurately measure eDNA concentration. Here we test whether multi-species PCR, that is, metabarcoding, with an internal standard can quantify eDNA of marine bony fish. Replicate amplifications with Riaz 12S gene primers were spiked with known amounts of a non-fish vertebrate DNA standard, indexed separately, and sequenced on an Illumina MiSeq. Fish eDNA copies were calculated by comparing fish and standard reads. Relative reads were directly proportional to relative DNA copies, with average and maximum variance between replicates of about 1.3- and 2-fold, respectively. There was an apparent minimum threshold for consistent amplification of about 10 eDNA copies per PCR reaction. The internal DNA standard corrected for suppression of fish read counts due to amplification of non-fish DNA. To assess potential PCR bias among species, we compared reads obtained with Riaz 12S primers to those with modified MiFish primers, which target a different 12S gene segment. Our results provide evidence that Riaz 12S gene metabarcoding with an internal DNA standard quantifies marine bony fish eDNA over a range of about 10–5000 copies per reaction, without indication of significant PCR bias among teleost species. In mid-Atlantic coastal samples, eDNA rarity was the main limitation to reproducible detection and quantification, and this was partly overcome by increasing the amount of a DNA sample amplified. Our findings support incorporating a DNA standard in 12S metabarcoding to help quantify eDNA abundance for marine bony fish species.

**KEYWORDS**

bottom trawl survey, coastal ecosystem, eDNA, environmental DNA, marine fish, metabarcoding, quantitative metabarcoding

This is an open access article under the terms of the [Creative Commons Attribution-NonCommercial](https://creativecommons.org/licenses/by-nc/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.

© 2022 The Authors. *Environmental DNA* published by John Wiley & Sons Ltd.

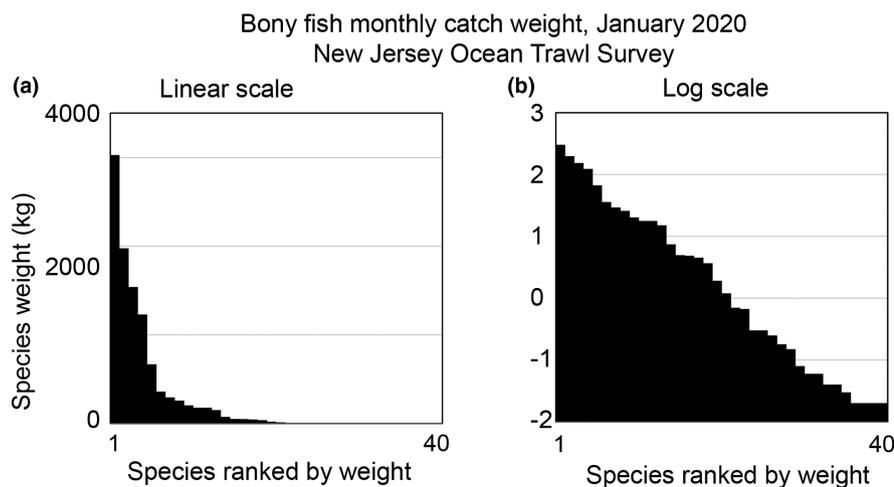
## 1 | INTRODUCTION

Addressing human impacts on the ocean calls for regular monitoring of marine life (Halpern et al., 2008; Lubchenco et al., 2020). Accurate assessment of marine fish populations, for example, enables effective fisheries management, helps gauge value of Marine Protected Areas (MPAs), and helps reveal ecological footprints of maritime industries including aquaculture and wind farms (Ahmadia et al., 2015; Fernandes et al., 2001; Hollingworth, 2000). Traditional survey methods for monitoring marine fish—capture, sonar, and visual—are relatively expensive and need specialized equipment and trained personnel. In some settings, traditional surveys may be harmful or otherwise unsuitable. For instance, bottom trawls may damage sea floor habitat and cannot be deployed at rocky sites (de Groot, 1984). Environmental DNA offers an informative addition to current survey techniques. Advantages include relatively low cost, modest field equipment, performance by a wide variety of personnel, harmlessness to environment, and applicability in difficult environments (Bourlat et al., 2013; Hansen et al., 2018). Inherent disadvantages are absent information on organism size, life stage, sex, and health. Despite limitations, eDNA appears poised to become a routine monitoring tool for marine life (Gilbey et al., 2021; Hinz et al., 2022). Current challenges to wider adoption include lack of standardized methods, incomplete DNA reference libraries, and, most importantly, limited understanding of how eDNA abundance relates to organism abundance under various environmental conditions (Allan et al., 2020; Andruszkiewicz et al., 2019; Collins et al., 2018; Jeunen et al., 2019).

Advancing knowledge of how fish species eDNA abundance informs on fish species abundance has two essential components. The first is accurately measuring the concentration of a species eDNA in collected water samples. The second is analyzing how that concentration relates to the species local abundance. This latter component can be challenging regarding marine fish, as there is no gold standard for abundance—all survey methods have biases (Arreguín-Sánchez, 1996; Fraser et al., 2007). In this report, we focus on the first step, accurate measurement, and test whether 12S gene metabarcoding incorporating a DNA standard can be a quantitative

tool for marine bony fish eDNA. Metabarcoding uses broad-range primers that amplify the target gene of multiple species in a taxonomic group, such as vertebrates, and applies Next Generation Sequencing (NGS) to sequence the resulting mixture of amplified DNA (Taberlet et al., 2018). Bioinformatic processing bins NGS output, called reads, into amplicon sequence variants (ASVs), and identifies ASVs by matching to a genetic reference library (Callahan et al., 2016). Metabarcoding is often considered a qualitative tool best suited for presence/absence surveys (Bush et al., 2020; Leray & Knowlton, 2015; Turunen et al., 2021). Hindrances to quantitative metabarcoding include differing PCR bias among species, suppression of target species reads due to amplification of nontarget DNA, and the fact that metabarcoding PCR tends to generate the same number of reads unrelated to the amount of eDNA added (Kelly et al., 2019; Krehenwinkel et al., 2017). In contrast, qPCR and other single-species techniques provide reliable measures of eDNA concentration (Andruszkiewicz et al., 2020; Doi et al., 2015; Shelton, Gold, et al., 2022). However, it may be unwieldy to develop and apply single species assays covering the multitude of taxa that are actively managed or otherwise of interest. For example, in our coastal mid-Atlantic study area, about 30 fish species are subject to catch quotas or are monitored due to threatened status (Atlantic States Marine Fisheries Commission, 2022). For more general ecosystem assessment such as for MPAs, taxonomically broad-range monitoring with metabarcoding would be helpful. Regional targets might include the approximately 100 fish species captured annually in bottom trawl surveys, or the over 400 species recorded in coastal waters (Able, 1992).

In this report, we analyze bony fish eDNA in water samples collected during New Jersey Ocean Trawl Survey (NJOTS) and from an ongoing environmental survey of shoreline sites in coastal New Jersey, New York, and Massachusetts (Hinks & Barry, 2020; Stoeckle et al., 2022, 2021). We test whether adding a DNA standard to metabarcoding PCRs measures absolute concentration of marine bony fish eDNA. In support of this strategy, spike-in standards quantify absolute microbial abundance in 16S rRNA gene metabarcoding (Zemb et al., 2019). Quantifying marine fish eDNA by spiking metabarcoding PCRs with internal standards was first described by



**FIGURE 1** Mid-Atlantic bony fish abundance differs over multiple orders of magnitude, shown in (a) linear and (b) logarithmic scale. Source data are in Table S1.

Ushio et al. (2018). To date, this methodology has not been widely applied to vertebrate eDNA, possibly due to limited information about accuracy and reproducibility. Here, we test the performance of a metabarcoding spike-in protocol to quantify marine fish eDNA. Advantages of our regional focus include long-term data on fish abundance obtained with NJOTS (Levesque, 2019), an excellent DNA reference library for commonly encountered fish species, and, among bony fish, highly conserved binding sites for mitochondrial 12S rRNA gene metabarcoding primers. Marine fish differ widely in abundance. For example, in NJOTS, monthly biomass per species ranges over four or five orders of magnitude (e.g., Figure 1, Table S1). With this in mind, we aimed to evaluate samples covering a wide range of eDNA copies per species. To better understand sources of variation and the value of replicate PCR, all replicates were analyzed individually. We focused on four related goals: first, establishing that MiSeq reads are proportional to eDNA copies; second, defining the lower limit to reproducible detection; third, overcoming distortion due to amplification of nontarget DNA; and fourth, evaluating PCR bias.

## 2 | MATERIALS AND METHODS

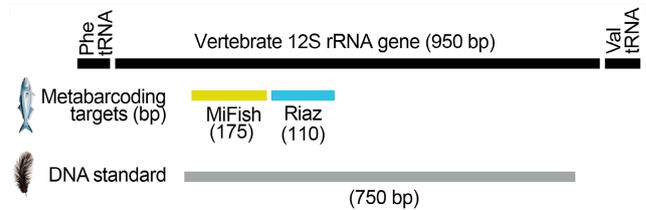
Most of the methods employed in this study have previously been described in detail (Stoeckle et al., 2022, 2021). These are briefly summarized here, and additional details are in Table S2.

### 2.1 | Water collection, filtration, DNA extraction

Standard water collection volume was 1 L. The data presented were obtained from 19 NJOTS and 23 environmental survey samples, plus tap water controls. Filtration was done using a 47 mm diameter, 0.45  $\mu$ m pore-size nitrocellulose filter, and filters were stored at  $-20^{\circ}\text{C}$  for up to 1 month before DNA extraction. DNA was extracted with DNeasy PowerSoil Pro Kit (QIAGEN), recovered in 100  $\mu$ l Buffer C6, and stored at  $-20^{\circ}\text{C}$  for up to 6 months before library preparation. Negative filtration and extraction controls were prepared from 1 L samples of laboratory tap water and were processed using the same equipment and procedures as for field samples. No animals were housed or experimented upon as part of this study. No endangered or protected species were collected.

### 2.2 | 12S gene DNA standards

We selected ostrich (*Struthio camelus*) and emu (*Dromaius novaehollandiae*) as internal standards, as DNA could be obtained from commercial food products (ostrich, American Ostrich Farms; emu, Newport Jerky Company), both were unlikely to be present in regional environmental samples, and Riaz 12S gene metabarcoding primer sites were identical to those in bony fish. Alternate standards were employed as a step toward understanding the generalizability



**FIGURE 2** Targeted gene regions. Schematic of vertebrate 12S gene with metabarcoding targets and segment used for standard shown.

of this approach. Standards were generated with M13-tailed primers that amplify a 689 bp segment of 12S gene covering the Riaz primer target sites and flanking regions (Figure 2). Primer sequences and thermal cycling parameters are in Table S2. Sanger sequencing confirmed 100% match to *S. camelus* or *D. novaehollandiae* reference sequences. PCR products were purified with AMPure beads at 1:1, concentration measured by Qubit, and a series of 10-fold dilutions were prepared in Buffer EB. Amounts added to replicates in this study ranged from 500 to 0.5 ag, corresponding to 600–0.6 copies, respectively. Linear regression and Fisher's exact test analyses were made with Prism 8.

### 2.3 | PCR

Metabarcoding PCR reactions were carried out in 25  $\mu$ l total volume with TaKaRa High Yield PCR EcoDry™ Premix. Typical conditions were 5  $\mu$ l of extracted DNA or 5  $\mu$ l of molecular biology grade water, and 200 nM of Illumina-tailed Riaz 12S primers (IDT) (Figure 2) (Riaz et al., 2011). Where noted, modified MiFish-U-F/R2 primers were used (Miya et al., 2015; Stoeckle et al., 2022). The modified set has reduced off-target amplification of bacterial 16S gene, facilitating amplicon sequencing without gel purification. Primer sequences and thermal cycling conditions are shown in Table S2. Negative control reactions were included in all amplification sets. Five micro liters of each reaction mix were run on a 2.5% agarose gel to assess amplification, and the remaining 20  $\mu$ l were diluted 1:20 in Buffer EB (QIAGEN) to be used as template for indexing. Indexing was done with Nextera XT kit and Cytiva PuReTaq Ready-To-Go PCR beads, using 5  $\mu$ l of diluted primary PCR product as template (thermal cycling protocol in Table S2).

### 2.4 | Next-generation sequencing, bioinformatic analysis

Sequencing was performed at GENEWIZ on an Illumina MiSeq with Reagent Kit v2, 2 $\times$ 150bp, (2 $\times$ 250bp for MiFish libraries), with 10% PhiX added to each run. Bioinformatic analysis was performed on Illumina FASTQ files using a DADA2 pipeline (Callahan et al., 2016; Callahan et al., 2017) (DADA2 pipeline R script posted in Appendix S1). Taxon assignments were generated by comparison to

an internal 12S gene reference library for regional fishes and other commonly amplified vertebrate ASVs (Appendix S2). In addition, all ASVs were manually submitted to GenBank to recover overlooked matches to 12S gene sequences not included in internal library. All identifications were based on 100% match to a reference sequence. Some near shore samples generated matches to nonlocal fish species, consistent with wastewater origin (Fujii et al., 2019). These were excluded from analysis. In addition, given that cartilaginous fish eDNA amplifies poorly with both primer sets used in this study, elasmobranch detections were set aside. Tap water eDNA and reagent grade water libraries were negative for fish ASVs after filtering DADA2 output tables as previously described. Filtering consisted of excluding detections comprising <1/1000th of the total for that taxon among all libraries in the run. For MiFish amplifications, all ASVs were manually submitted to GenBank. The data presented were obtained from 144 libraries generated from the 42 field samples plus controls, which were analyzed in five MiSeq runs together with other samples not reported here. The DADA2 pipeline was applied to separately to each sequencing run. Except for Riaz vs. MiFish experiment described below, all analyses compared libraries in the same sequencing run. The MiFish 12S gene segment distinguishes some species which have identical Riaz 12S gene sequences. To enable comparison, MiFish reads for species with identical Riaz sequences were collapsed to genus or family level (relevant details in Table S3).

### 3 | RESULTS

To test whether metabarcoding reads were proportional to eDNA copies, we prepared sets of amplification replicates containing the same amount of an eDNA sample and different amounts of ostrich or emu DNA standard, as illustrated in Figure 3a. Replicates were indexed separately, sequenced, and fish eDNA copies were calculated

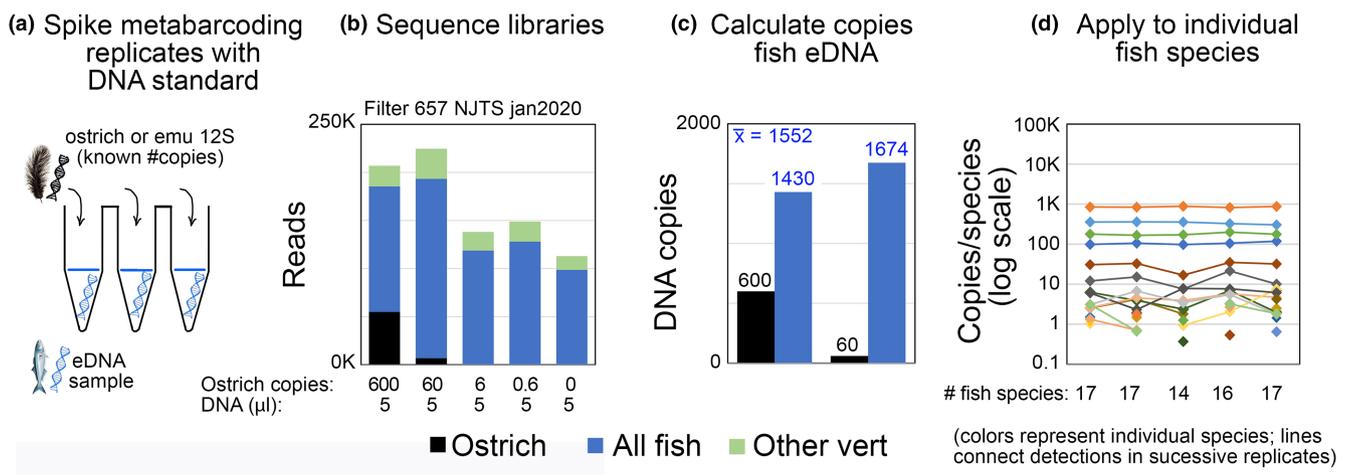
by comparing fish reads to standard reads in each replicate as follows:

$$\text{Copies fish eDNA} = (\text{Fish reads}) \times \left[ \left( \frac{\text{Standard copies}}{\text{Standard reads}} \right) \right]$$

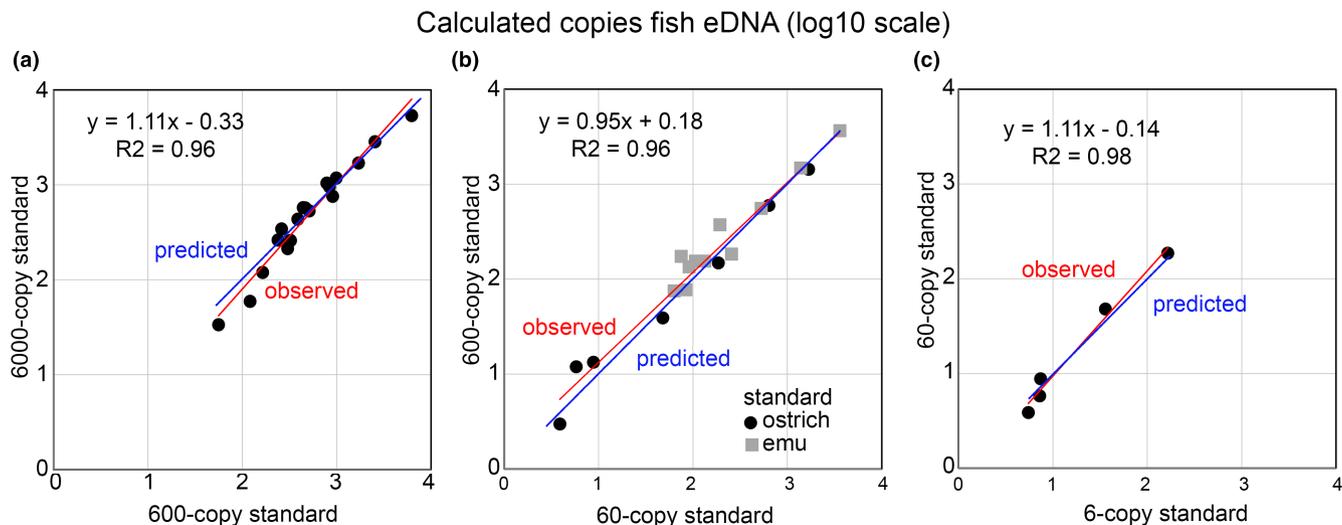
Several findings were evident. First, total vertebrate reads differed among replicates, up to about 2-fold. This might reflect variable efficiency of indexing, PCR inhibition, or other factors. Second, in some replicates prepared with smaller amounts of standard DNA, there were no standard reads (e.g., Figure 3b, see replicate with 6 copies ostrich DNA). In these cases, standard reads were present in FASTQ files but were below the filtering threshold (see Methods). Third, replicates containing different amounts of standard DNA yielded similar estimates of fish eDNA copies, consistent with relative reads being proportional to relative DNA copies (Figure 3c). Fourth, calculated copies per species were uniform across replicates for taxa with more than 10 calculated copies (Figure 3d). Below that level, some species were detected inconsistently. Inconsistent detection, referred to as pickups or dropouts, is commonly observed in metabarcoding studies (Ficetola et al., 2015). These findings were further explored in experiments reported below, with supportive results.

#### 3.1 | Establishing that MiSeq reads are proportional to eDNA copies

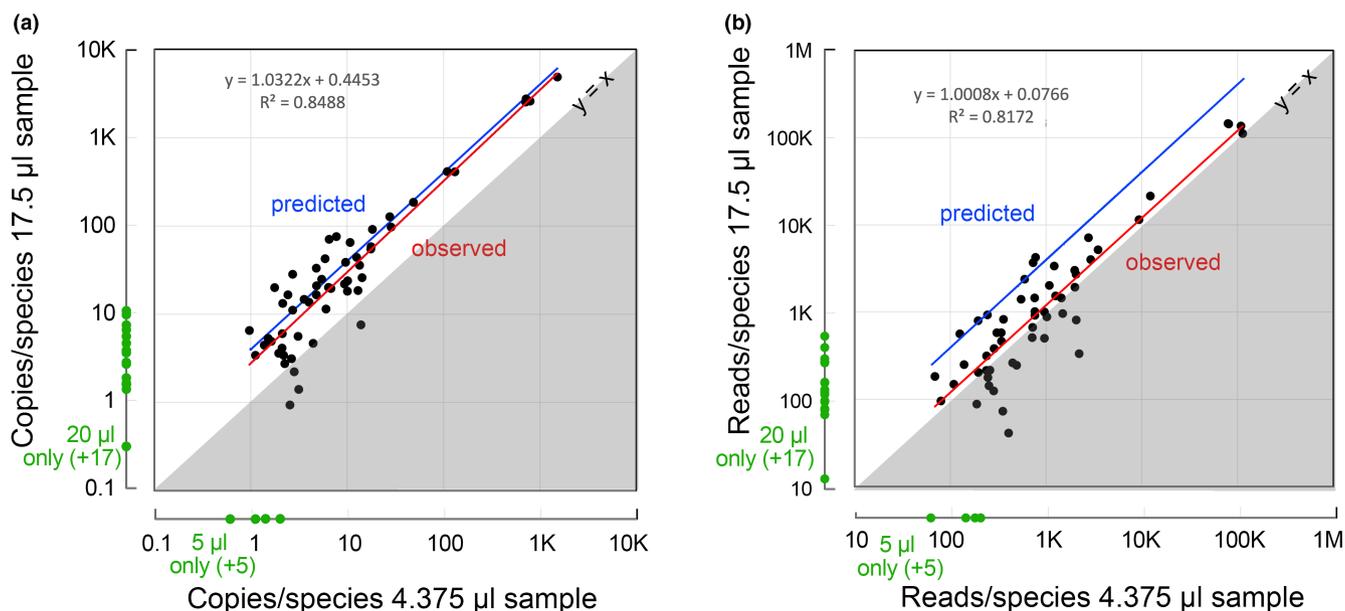
If MiSeq reads are proportional to eDNA copies, then replicates containing different amounts of standard DNA should yield identical estimates of fish eDNA copies. Replicates containing 5 μl of a DNA sample (n = 18) were spiked with 6000 or 600 copies of ostrich DNA, and copies of fish eDNA in each replicate were calculated by comparing fish reads to standard reads as described above. Different amounts of standard DNA generated similar estimates of fish eDNA copies (Figure 4a). The average arithmetic



**FIGURE 3** Experimental protocol, representative results. Beginning at left, (a) technical replicates were spiked with known amounts of ostrich or emu 12S DNA; (b) separate libraries were generated from each replicate; (c) total fish eDNA copies were calculated by comparing fish and DNA standard reads; and (d) eDNA copies for individual fish species were calculated by applying average total fish eDNA copies to each library. Data shown were obtained from one NJOTS sample (source data Table S4a,b).



**FIGURE 4** Calculated copies of fish eDNA according to DNA standards. (a) separate libraries were generated from replicates spiked with 6000 or 600 copies of ostrich DNA, and copies of fish eDNA were calculated as described. Each point represents one eDNA sample analyzed with one pair of replicates. Blue line denotes relative reads directly proportional to relative copies. Red line is linear regression of log-transformed experimental data. (b) Protocol applied to a different set of eDNA samples, using 600 or 60 copies of ostrich or emu DNA. (c) Results with subset of samples in center panel, using 60 or 6 copies ostrich standard (source data Tables S4A–S6).



**FIGURE 5** Copies and reads compared to amount eDNA analyzed. Replicates were prepared with indicated amounts of an eDNA sample plus ostrich standards. Each point represents one species in one pair of replicates. (a) black points represent paired values for copies/species or (b) reads/species (right). In both graphs, blue line corresponds to copies or reads per species being 4-fold higher in 17.5 µl replicate, red line is linear regression of log-transformed experimental data, and the  $y = x$  boundary indicates plot if values were equal. Green points represent pickups, i.e., species detected in one of the paired replicates (source data Table S5).

deviation from identity was 1.3-fold, and the maximum was 2-fold. The direct relationship between relative reads and relative DNA copies held over a wide range of sample eDNA content, from about 50 to about 5000 calculated copies fish eDNA. This protocol was applied to additional samples ( $n = 18$ ), except that 600 or 60 copies of ostrich or emu DNA were added. As in prior experiment, different amounts of standard generated nearly identical

estimates of fish eDNA copies (Figure 4b). The average and maximum difference in replicate pairs was 1.2- and 1.7-fold, respectively. A proportional relationship between reads and copies was observed in samples with widely varying eDNA content ranging from about 5 to about 5000 copies of fish eDNA. A direct relationship of relative reads to relative DNA copies extended to 60- vs. 6-copy replicates (Figure 4c).

To further test whether metabarcoding reads are proportional to eDNA copies, we varied the amount of DNA analyzed. Paired replicates with standards were prepared with 4.375  $\mu\text{l}$  or a 4-fold greater amount, 17.5  $\mu\text{l}$ , of eDNA samples ( $n = 4$ ). Increased DNA input was evident in copies/species, but not in reads/species (Figure 5). Consistent with expectations, there were four-fold more pickups in 17.5  $\mu\text{l}$  replicates, and among all replicates, most pickups (21/22, 95%) were eDNAs present at fewer than 10 copies.

### 3.2 | Defining lower limit to reproducible detection

Drop-outs, as in illustrated Figures 3 and 5, indicate inconsistent amplification of a species eDNA. Conversely, the absence of drop-outs defines reproducible detection. To evaluate a potential threshold, we compiled detections from samples ( $n = 8$ ) analyzed with five replicates (Figure 6). Copies/species and reads/species were calculated by averaging positive values among replicates, that is, excluding non-detections. Higher eDNA copy number was associated with fewer dropouts (e.g., detected in all replicates: 4th vs. 1st quartile, 97% vs. 0%;  $p < 0.0001$ , Fisher's exact test). No dropouts were observed in detections with more than 10 calculated copies (Figure 6). A weaker trend was observed with raw reads (detected in all replicates, 4th quartile, copies vs. reads, 97% vs 74%,  $p = 0.007$ , Fisher's exact test).

### 3.3 | Overcoming distortion due to amplification of nontarget DNA

We analyzed copies in replicates with different proportions of fish vs. non-fish vertebrate reads, the latter including those due to DNA standard itself (Figure 7). Despite large differences within replicate sets in the proportion of total reads occupied by fish, fish copies and fish species detection were consistent across replicates. In addition, calculated fish copies were approximately two-fold higher in replicates prepared from a two-fold larger volume water sample, as expected, whereas fish reads were lower (Figure 7, columns 2 and 3). Most non-fish vertebrate DNA reads not from standards were

human or domestic animal, and the remainder were derived from non-fish wildlife (Tables S4–S6).

### 3.4 | Evaluating PCR bias

PCR bias may arise from primer mismatch, effects of flanking sequences on primer binding, or intrinsic differences in the intervening segment that affect amplification efficiency (Silverman et al., 2021). We reasoned that if PCR bias differed among species, then primer sets targeting different gene regions should give different results. Conversely, similar results would be evidence of absence of significant PCR bias, at least for the species present in the sample. To test this hypothesis, we compared results obtained with Riaz primers to those generated with modified MiFish primers (see Methods), which amplify an adjacent region of vertebrate 12S gene (Figure 2). Libraries were generated from NJOTS samples ( $n = 19$ ) from January 2020, and the results were pooled. Reads/species were well-correlated between the two primer sets (Figure 8, Tables S7–S9), although there were greater differences below about 1000 reads/species, which might be due to variability inherent in low-copy number eDNA as noted above (more than 1000 reads, linear regression log-transformed values: slope, 1.00;  $R^2$ , 0.94; fewer than 1000 reads, no significant relationship). The number of dropouts was similar between primer sets, and most (10/11; 91%) were represented by fewer than 1000 reads, consistent with these corresponding to low-copy number eDNA.

## 4 | DISCUSSION

We tested whether Riaz 12S gene metabarcoding with an internal DNA standard quantifies marine bony fish eDNA. We present several lines of evidence supporting this hypothesis. First, reads were directly proportional to DNA copies over a 1000-fold range of standard DNA copies (6–6000) and a 1000-fold range of calculated eDNA copies (about 5–5000). Paired replicate measurements differed by an average and maximum of 1.3- and 2-fold,

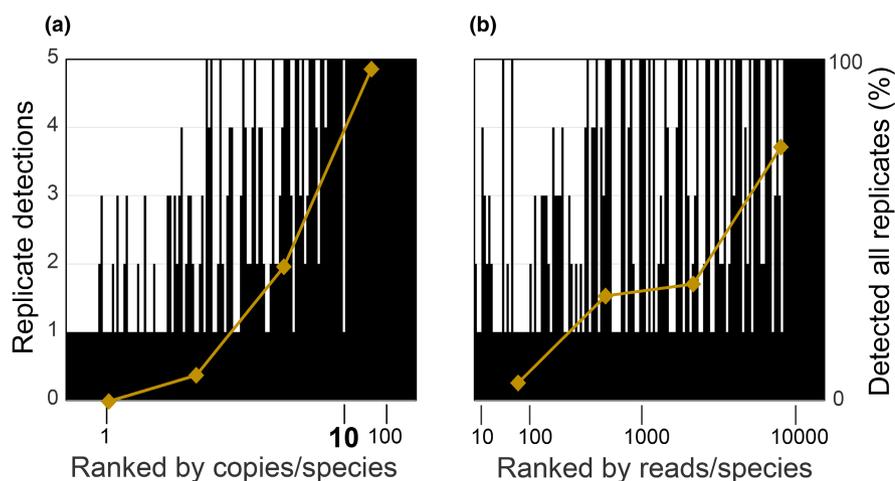
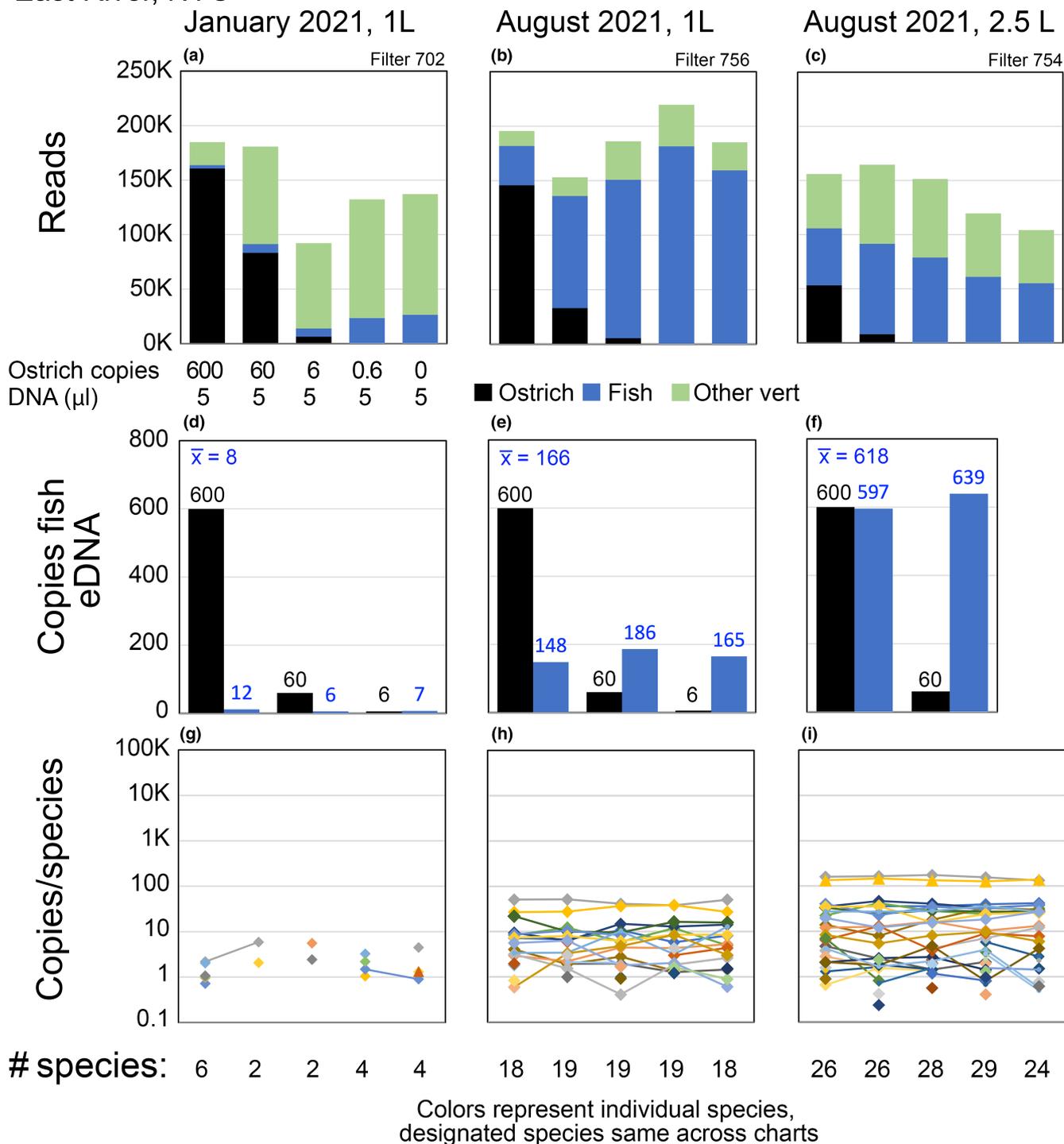


FIGURE 6 Replicate detection. Each black line represents one species in one eDNA sample analyzed with five technical replicates ( $n = 154$ ). (a) samples ranked by copies per species; (b) samples ranked by reads per species. Superimposed gold line charts percent detections present in all replicates for each quartile (source data Table S4c).

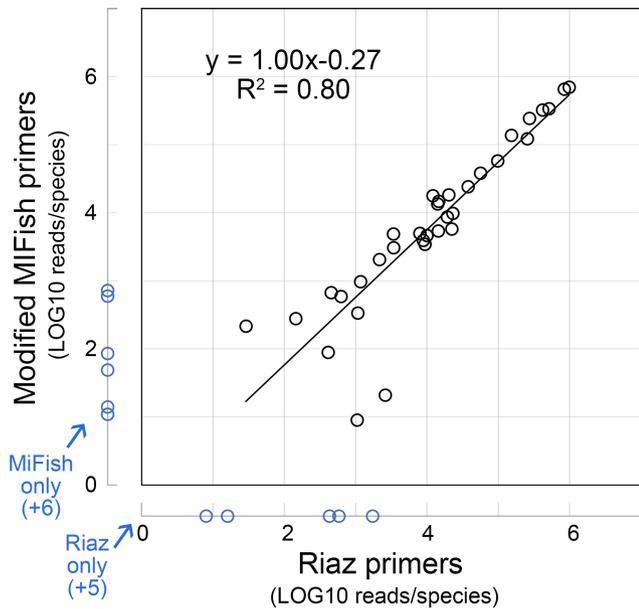
## East River, NYC



**FIGURE 7** Evaluating suppression due to non-fish vertebrate reads. Three eDNA samples were analyzed with protocol illustrated in Figure 3. (a–c) reads by category; (d–f) calculated total copies fish eDNA; (g–i) calculated eDNA copies per fish species. See text for discussion (source data Table S4a,b)

respectively. Second, the DNA standard accurately measured the proportion of a DNA sample used for PCR, information that could not be extracted from read data alone. Third, we found a threshold for reproducible detection of about 10 copies/species per PCR reaction and demonstrated improved detection of low-copy number eDNA by analyzing a larger proportion of a DNA sample.

These results provide an intuitive mechanistic explanation for the empirical phenomenon of metabarcoding dropouts. Fourth, incorporating a DNA standard corrected for suppression of fish reads due to amplification of non-fish vertebrate DNA and variation in total reads among replicates. Finally, a comparison of eDNA reads obtained with primer sets targeting different 12S gene regions



**FIGURE 8** Comparison of reads per bony fish species with different 12S primer sets. Black circles, species with shared detections ( $n = 33$ ). Blue circles, species detected with one primer set only. Line is linear regression of log-transformed data (source data Tables S7–S9).

was consistent with absence of significant PCR bias among marine bony fish species.

#### 4.1 | Limitations

The bioinformatic pipeline filters out low-level detections, those present at  $<1/1000$  of total reads per taxon in sequencing run. This threshold aims to reduce effects of tag jumping and sequencing error. However, it may eliminate true positives if total reads for a taxon are high, which was the case for DNA standards. As expected, in some replicates prepared with smaller amounts of standard DNA, there were no standard reads after filtering (e.g., Figures 3 and 7), thus we could not use these to test our hypothesis. This limit could be addressed with unique dual indexing (Bohmann et al., 2021), which should reduce tag jumping and need for threshold. Although not a factor in this study, on a MiSeq, a high level of DNA standard ( $>100,000$  copies per reaction) suppresses detection of fish eDNA in near shore samples (Stoeckle et al., 2022). If desired, this could potentially be overcome with deeper sequencing using NovaSeq or a similar platform (Singer et al., 2019).

Comparison of reads/species with MiFish and Riaz primer sets, which target different gene regions, was consistent with absence of significant PCR bias among bony fish species. However, we cannot exclude PCR bias, particularly for species not represented in the NJOTS samples analyzed. Both primer sets amplify cartilaginous fish eDNA poorly, and this limitation might apply to other species. It is well established that broad range metazoan primers can give very different results in terms of taxon detection and relative

abundance (Kumar et al., 2022; Liu & Zhang, 2021). Metabarcoding mock community samples may enable adjustments that correct for PCR bias (Shelton, Ramón-Laca, et al., 2022). Regarding the latter, it may be worthwhile to test mock communities that closely mimic eDNA samples, that is, in which target DNAs are present at eDNA-like concentrations and comprise a small fraction of total DNA. Primer mismatch may be the main factor determining 12S PCR bias in marine fish. For example, in Shelton, Ramón-Laca, et al. (2022), in analyzing fish eDNA with MiFish-UF/R primers, outliers were cartilaginous species, which are known to amplify poorly with this primer set, while teleost taxa amplification was relatively uniform. Similar amplification with MiFish and Riaz 12S primers in this report may reflect that both sets are exact matches for most bony fish species. Recent reports demonstrate correlation between relative reads and relative organism abundance, which implies that even without internal standards, metabarcoding correctly reports relative eDNA abundance in some settings (Afzali et al., 2021; Di Muri et al., 2022; Ershova et al., 2021; Klymus et al., 2017). The reverse MiFish primer overlaps with the forward Riaz primer, which might mean there is less difference between these primer sets than one would otherwise expect. PCR bias could be further evaluated by testing both primer sets on samples containing eDNA of other species and on mock communities, calibrating modified MiFish primers with protocols similar to those here, and comparing NGS metabarcoding copies to results with single-species qPCR or ddPCR assays.

Copy number calculations in this study assumed that the ostrich and emu amplicons employed as DNA standards were accurately measured and that they amplified without PCR bias as compared to fish eDNA. Regarding the former, dilution experiments analyzed with Poisson distribution indicate these standards are within about 3-fold of nominal value (Stoeckle et al., 2022). Similar support is provided by observation that the lowest values for calculated copies/species per replicate were regularly at or close to 1 (e.g., Figures 3, 6 and 7). The latter aspect, PCR bias, could be evaluated by adding standards to mock eDNA community samples. Alternatively, synthetic DNA standards closely related to target sequences may help minimize PCR bias (Tsuji et al., 2020).

#### 4.2 | Other reports

As noted in Introduction, Ushio et al. (2018) were the first to apply internal standards to quantify MiSeq metabarcoding of marine fish eDNA. In that report, eDNA samples were spiked with four internal standards present at 50–1000 copies, amplified in triplicate with unmodified MiFish-U primers, and resulting amplicons were pooled, indexed, and sequenced on a MiSeq. eDNA copies were calculated by comparing fish reads to standard reads, and these values were compared to qPCR copies for total fish eDNA and two fish species. Some samples amplified poorly or not at all, possibly due to PCR inhibition, and correlations with qPCR were modest even after excluding outliers. The authors note that off-target amplification with unmodified MiFish-U primers may have interfered with qPCR assay for total fish

eDNA. In addition, we observe that calculated fish eDNA copies/species were mostly below 10 copies per reaction, which may have hampered reproducible measurement. This methodology has been applied to quantify marine fish eDNA around artificial reefs (Sato et al., 2021), but to our knowledge accuracy for marine fish has not been further evaluated.

### 4.3 | Looking ahead

Our findings support further testing of the protocol outlined here for marine bony fish metabarcoding, namely, incorporation of an internal DNA standard and separate analysis of replicates. A reasonable starting protocol would be a pair of replicates for each eDNA sample, one with 600 and one with 60 copies of DNA standard. Potential avenues for further investigation include comparing metabarcoding eDNA measurement to that obtained with qPCR, applying MiFish and Riaz primers to additional samples, testing different DNA standards, and adapting this approach to analyze cartilaginous fish eDNA. In the long run, just as a net catches some fish while others swim away, a metabarcoding assay does not need to perfectly report eDNA abundance to be useful, but it does need to be reproducible over a range of eDNA concentrations. Incorporating a DNA standard will likely help. In evaluating how eDNA abundance relates to fish abundance, it may be helpful to consider allometric scaling of fish biomass, which improves correlation of eDNA reads with fish abundance (Stoeckle et al., 2021; Yates et al., 2020).

Our results add evidence that current metabarcoding protocols suffice for relatively abundant marine fish eDNA, which presumably includes that of most commercial species. Conversely, we find that eDNA rarity is the biggest challenge to fish eDNA metabarcoding. With a typical protocol applied to near shore mid-Atlantic samples, eDNA of many bony fish species was below the limit of reproducible detection. This could be addressed by amplifying a larger proportion of a DNA sample, starting with multiple or larger water samples, or utilizing a different collection strategy such as targeted collection near to organisms, passive collection, or examining DNA trapped by filter feeders (Baker et al., 2018; Bessey et al., 2020; Bessy et al., 2021; Hunter et al., 2019; Mächler et al., 2016). Assuming complete DNA recovery during extraction, the apparent detection threshold of about 10 eDNA copies per PCR reaction translates to about 200 eDNA copies per liter of seawater. Regarding rare species, qPCR studies indicate that in some settings endangered marine vertebrates have eDNA concentrations above this level (Valsecchi et al., 2022; Weltz et al., 2017), so likely could be reliably detected with this 12S metabarcoding protocol. Our analysis benefitted from a well-populated genetic reference library for northwest Atlantic coastal marine fish. However, there is no obvious roadblock to application in areas with higher diversity or less complete reference libraries, such as tropical reefs or abyssal zone ocean. As reference libraries become more comprehensive and computational power increases, eDNA will be more capable of identifying and quantifying rarer species and those in areas with less historical data. The field

will benefit from sharing protocols and standardized metadata tags and categories (Samuel et al., 2021).

### 4.4 | Summary

Effective ocean management relies on accurate, up-to-date information on marine life. eDNA provides a relatively low-cost, harmless, widely applicable, and potentially autonomous method that will supplement traditional ocean survey techniques. Reliable measurement of eDNA concentration will enhance the value of this new technology in accurate monitoring of marine animal populations.

#### AUTHOR CONTRIBUTIONS

Designed study: MYS. Updated bioinformatics: MC. Collected samples: MYS, JHA. Analyzed samples: MYS. Analyzed data: MYS, JHA. Prepared figures: MYS. Wrote first draft of manuscript: MYS. Revised manuscript: MYS, JHA, MC.

#### ACKNOWLEDGMENTS

This work was supported by NOAA Grant NA20OAR01104XX. Sampling during the New Jersey Ocean Trawl Survey was supported by US Fish and Wildlife Service Wildlife and Sport Fish Restoration Program. We thank Greg Hinks and Stacy VanMorter at Bureau of Marine Fisheries, New Jersey Department of Environmental Protection for sharing NJOTS catch data. At The Rockefeller University, we thank Jeanne Garbarino, Disan Davis, and Odaelys Walwyn for laboratory resources and assistance. The funders had no role in the study design, data collection and analysis, decision to publish, or preparation of the manuscript.

#### CONFLICT OF INTEREST

This is to state that the authors have no conflict of interest to declare. This includes absence of any interest or relationship, financial or otherwise, that might be perceived as influencing an author's objectivity.

#### DATA AVAILABILITY STATEMENT

Illumina FASTQ files underlying this article are deposited in NCBI Bioproject ID PRJNA864636.

#### ORCID

Mark Y. Stoeckle  <https://orcid.org/0000-0002-8280-4795>

#### REFERENCES

- Able, K. W. (1992). Checklist of New Jersey saltwater fishes. *Bulletin New Jersey Academy Sciences*, 37, 1–11.
- Afzali, S. F., Bourdages, H., Laporte, M., Mérot, C., Normandeau, E., Audet, C., & Bernatchez, L. (2021). Comparing environmental metabarcoding and trawling survey of demersal fish communities in the Gulf of St. Lawrence, Canada. *Environmental DNA*, 3, 22–42.
- Ahmadia, G. N., Glew, G., Provost, M., Gill, D., Hidayat, N. I., Mangubhai, S., Purwanto, X., & Fox, H. E. (2015). Integrating impact evaluation in the design and implementation of marine protected areas.

- Philosophical Transactions of the Royal Society B: Biological Sciences*, 370, 20140275.
- Allan, E. A., Zhang, W. G., Lavery, A. C., & Govindarajan, A. (2020). Environmental DNA shedding decay rates from diverse animal forms and thermal regimes. *Environmental DNA*, 3, 1–23.
- Andruszkiewicz, E. A., Koseff, J. R., Fringer, O. B., Ouellette, N. T., Lowe, A. B., Edwards, C. A., & Boehm, A. B. (2019). Modeling environmental DNA transport in the coastal ocean using Lagrangian particle tracking. *Frontiers in Marine Science*, 6, 477.
- Andruszkiewicz, E. A., Yamahara, K. M., Closek, C. J., & Boehm, A. B. (2020). Quantitative PCR assays to detect whales, rockfish, and common murre environmental DNA in marine water samples of the northeastern Pacific. *PLoS One*, 15, e0242689.
- Arreguín-Sánchez, F. (1996). Catchability: A key parameter for fish stock assessment. *Reviews in Fish Biology and Fisheries*, 6, 221–242.
- Atlantic States Marine Fisheries Commission. (2022). Stock status, June 2022. [http://www.asmfc.org/files/pub/ASMFC\\_StockStatus\\_June2022.pdf](http://www.asmfc.org/files/pub/ASMFC_StockStatus_June2022.pdf)
- Baker, C. S., Steel, D., Nieukirk, S., & Klinck, H. (2018). Environmental DNA (eDNA) from the wake of the whales: Droplet digital PCR for detection and species identification. *Frontiers in Marine Science*, 5, 133.
- Bessey, C., Jarman, S. N., Berry, O., Olsen, Y. S., Bunce, M., Simpson, T., Power, M., McLaughlin, J., Edgar, G. J., & Keesing, J. (2020). Maximizing fish detection with eDNA metabarcoding. *Environmental DNA*, 2, 493–504.
- Bessy, C., Jarman, S. N., Simpson, T., Miller, H., Stewart, T., Keesing, J. K., & Berry, O. (2021). Passive eDNA collection enhances aquatic biodiversity analysis. *Communications Biology*, 4, 236.
- Bohmann, K., Elbrecht, V., Carøe, C., Bista, I., Leese, F., Bunce, M., Yu, D. W., Seymour, M., Dumbrell, A. J., & Creer, S. (2021). Strategies for sample labelling and library preparation in DNA metabarcoding studies. *Molecular Ecology Resources*, 22, 1231–1246.
- Bourlat, S. J., Borfa, A., Gilbert, J., Taylor, M. I., Davies, N., Weisberg, S. B., Griffith, J. F., Lettieri, T., Field, D., Benzie, J., Glöckner, F. O., Rodríguez-Ezpeleta, N., Faith, D. P., Bean, T. P., & Obst, M. (2013). Genomics in marine monitoring: New opportunities for assessing marine health status. *Marine Pollution Bulletin*, 74, 19–31.
- Bush, A., Monk, W. A., Compson, Z. G., Peters, D. L., Porter, T. M., Shokralla, S., Wright, M. T. G., Hajibabaei, M., & Baird, D. J. (2020). DNA metabarcoding reveals metacommunity dynamics in a threatened boreal wetland wilderness. *Proceedings of the National Academy of Sciences of the United States of America*, 117, 8539–8545.
- Callahan, B. J., McMurdie, P. J., & Holmes, S. P. (2017). Exact sequence variants should replace operational taxonomic units in marker-gene data analysis. *ISME Journal*, 11, 2639–2643.
- Callahan, B. J., McMurdie, P. J., Rosen, M. J., Han, A. W., Johnson, J. A., & Holmes, S. P. (2016). DADA2: High-resolution sample inference from Illumina amplicon data. *Nature Methods*, 13, 581–583.
- Collins, R. A., Wangensteen, O. S., O’Gorman, E. J., Mariani, S., Sims, D. W., & Genner, M. J. (2018). Persistence of environmental DNA in marine systems. *Communications Biology*, 1, 185.
- De Groot, S. J. (1984). The impact of bottom trawling on benthic fauna of the North Sea. *Ocean Management*, 9, 177–190.
- Di Muri, C., Handley, L. L., Bean, C. W., Benucci, M., Harper, L. R., James, B., Li, J., Winfield, I. J., & Hänfling, B. (2022). Spatio-temporal monitoring of lake fish spawning activity using environmental DNA metabarcoding. *Environmental DNA*, 1–12.
- Doi, H., Uchii, K., Takahara, T., Matsushashi, S., Yamanaka, H., & Minamoto, T. (2015). Use of droplet digital PCR for estimation of fish abundance and biomass in environmental DNA surveys. *PLoS One*, 10, e0122763.
- Ershova, E. A., Wangensteen, O. S., Descoteaux, R., Barth-Jensen, C., & Præbel, K. (2021). Metabarcoding as a quantitative tool for estimating biodiversity and relative biomass of marine zooplankton. *ICES Journal of Marine Science*, 78, 3342–3355.
- Fernandes, T. F., Eleftheriou, A., Ackefors, H., Eleftheriou, M., Ervik, A., Sanchez-Mata, A., Scanlon, T., White, P., Cochrane, S., Pearson, T. H., & Read, P. A. (2001). The scientific principles underlying the monitoring of the environmental impacts of aquaculture. *Journal of Applied Ichthyology*, 17, 181–193.
- Ficetola, G. F., Pansu, J., Bonin, A., Coissac, E., Giguet-Coxev, C., de Barba, M., Gielly, L., Lopes, C. M., Boyer, F., Rayé, F. P. G., & Tablerlet, P. (2015). Replication levels, false presences and the estimation of the presence/absence from eDNA metabarcoding data. *Molecular Ecology Resources*, 15, 543–556.
- Fraser, H. M., Greenstreet, S. P. R., & Piet, G. J. (2007). Taking account of catchability in groundfish survey trawls: Implications for estimating demersal fish biomass. *ICES Journal of Marine Science*, 9, 1800–1819.
- Fujii, K., Doi, H., Matsuoka, S., Nagano, M., Sato, H., & Yamanaka, H. (2019). Environmental DNA metabarcoding for fish community analysis in backwater lakes: A comparison of capture methods. *PLoS One*, 14, e210357.
- Gilbey, J., Carvalho, G., Castilho, R., Coscia, I., Coulson, M. W., Dahle, G., Derycke, S., Francisco, S. M., Helyar, S. J., Johansen, T., Junge, C., Layton, K. K. S., Martinsohn, J., Matejusova, I., Robalo, J. I., Rodríguez-Ezpeleta, N., Silva, G., Strammer, I., Vasemägi, A., & Volckaert, F. A. M. (2021). Life in a drop: Sampling environmental DNA for marine fishery management and ecosystem monitoring. *Marine Policy*, 124, 104331.
- Halpern, B. S., Walbridge, S., Selkoe, K. A., Kappel, C. V., Micheli, F., D’Agrosa, C., Bruno, J. F., Casey, K. S., Ebert, C., Fox, H. E., Fujita, R., Heinemann, D., Lenihan, H. S., Madin, E. M. P., Perry, M. T., Selig, E. R., Spalding, M., Steneck, R., & Watson, R. (2008). A global map of human impact on marine ecosystems. *Science*, 319, 948–952.
- Hansen, B. K., Bekkevold, D., Clausen, L. W., & Nielsen, E. E. (2018). The skeptical optimist: Challenges and perspectives for the application of environmental DNA in marine fisheries. *Fish Fisheries*, 19, 751–768.
- Hinks, G., & Barry, L. (2020). *Ocean trawling survey, 2020*. New Jersey Department of Environmental Protection, Division of Fish and Wildlife, Marine Fisheries Administration, Bureau of Marine Fisheries.
- Hinz, S., Coston-Guarini, J., Marnane, M., & Guarini, J.-M. (2022). Evaluating eDNA for use within marine environmental impact assessments. *Journal of Marine Science and Engineering*, 10, 375.
- Hollingworth, C. (2000). Ecosystem effects of fishing. *ICES Journal of Marine Science*, 57, 465–792.
- Hunter, M. E., Ferrante, J. A., Meigs-Friend, G., & Ulmer, A. (2019). Improving eDNA yield and inhibitor reduction through increased water volumes and multi-filter isolation techniques. *Scientific Reports*, 9, 5259.
- Jeunen, G.-J., Knapp, M., Spencer, H. G., Taylor, H. R., Lamare, M. D., Stat, M., Bunce, M., & Gemmill, N. J. (2019). Species-level biodiversity assessment using marine environmental DNA metabarcoding requires protocol optimization and standardization. *Ecology and Evolution*, 9, 1323–1335.
- Kelly, R. P., Shelton, A. O., & Gallego, R. (2019). Understanding PCR processes to draw meaningful conclusions from environmental DNA studies. *Scientific Reports*, 9, 12133.
- Klymus, K. E., Marshall, N. T., & Stepien, C. A. (2017). Environmental DNA (eDNA) metabarcoding assays to detect invasive invertebrate species in the Great Lakes. *PLoS One*, 12, e0177643.
- Krehenwinkel, H., Wolf, M., Lim, J. Y., Rominger, A. J., Simison, W. B., & Gillespie, R. G. (2017). Estimating and mitigating amplification bias in qualitative and quantitative arthropod metabarcoding. *Scientific Reports*, 7, 17668.
- Kumar, G., Reaume, A. M., Farrell, E., & Gaither, M. R. (2022). Comparing eDNA metabarcoding primers for assessing fish communities in a biodiverse estuary. *PLoS One*, 17, e0266720.
- Leray, M., & Knowlton, N. (2015). DNA barcoding and metabarcoding of standardized samples reveal patterns of marine benthic diversity.

- Proceedings of the National Academy of Sciences of the United States of America*, 112, 2076–2081.
- Levesque, J. C. (2019). Spatio-temporal patterns of the oceanic conditions and nearshore marine community in the mid-Atlantic bight (New Jersey, USA). *PeerJ*, 7, e7927.
- Liu, J., & Zhang, H. (2021). Combining multiple markers in environmental DNA metabarcoding to assess deep-sea benthic biodiversity. *Frontiers in Marine Science*, 8, 684955.
- Lubchenco, J., Haugan, P. M., & Pangestu, M. E. (2020). Five priorities for a sustainable ocean economy. *Nature*, 588, 30–32.
- Mächler, E., Deiner, K., Spahn, F., & Altermatt, F. (2016). Fishing in the water: Effect of sampled water volume on environmental DNA-based detection of macroinvertebrates. *Environmental Science and Technology*, 50, 305–312.
- Miya, M., Sato, Y., Fukunaga, T., Sado, T., Poulsen, J. Y., Sato, K., Minamoto, T., Yamamoto, S., Yamanaka, H., Araki, H., Kondoh, M., & Iwasaki, W. (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, 150088.
- Riaz, T., Shehzad, W., Viari, A., Pompanon, F., Taberlet, P., & Coissac, E. (2011). ecoPrimers: Inference of new DNA barcode markers from whole genome sequence analysis. *Nucleic Acids Research*, 39, e145.
- Samuel, R. M., Meyer, R., Buttigieg, P. L. L., Davies, N., Jeffery, N. W., Meyer, C., Pavloudi, C., Pitz, K. J., Sweetlove, M., Theroux, S., van de Kamp, J., & Watts, A. (2021). Toward a global public repository of community protocols to encourage best practices in biomolecular ocean observing and research. *Frontiers in Marine Sciences*, 8, 758694.
- Sato, M., Inoue, N., Nambu, R., Furuichi, N., Imaizumi, T., & Ushio, M. (2021). Quantitative assessment of multiple fish species around artificial reefs combining environmental DNA metabarcoding and acoustic survey. *Scientific Reports*, 11, 19477.
- Shelton, A. O., Gold, Z. J., Jensen, A. J., D'Agnese, E., Allan, E. A., Van Cise, A., Gallego, R., Ramón-Laca, A., Garber-Yonts, M., Parsons, K., & Kelly, R. P. (2022). Toward quantitative metabarcoding. *Ecology*, 103, e396. <https://doi.org/10.1101/2022.04.26.489602>
- Shelton, A. O., Ramón-Laca, A., Wells, A., Clemons, J., Chu, D., Feist, B. E., Kelly, R. P., Parker-Stetter, S. L., Thomas, R., Nichols, K. M., & Park, L. (2022). Environmental DNA provides quantitative estimates of Pacific hake abundance and distribution in the open ocean. *Proceedings of the Royal Society B: Biological Sciences*, 289, 20212613.
- Silverman, J. D., Bloom, R. J., Jiang, S., Durand, H. K., Dallow, E., Mukherjee, S., & David, L. A. (2021). Measuring and mitigating PCR bias in microbiota datasets. *PLoS Computational Biology*, 17, e1009113.
- Singer, G. A. C., Fahner, N. A., Barnes, J. G., McCarthy, A., & Hajibabaei, M. (2019). Comprehensive biodiversity analysis via ultra-deep patterned flow cell technology: A case study of eDNA metabarcoding seawater. *Scientific Reports*, 9, 5991.
- Stoeckle, M. Y., Adolf, J., Ausubel, J. H., Charlop-Powers, Z., Dunton, K. J., & Hinks, G. (2022). Current laboratory protocols for detecting fish species with environmental DNA optimize sensitivity and reproducibility, especially for more abundant populations. *ICES Journal of Marine Science*, 79, 403–412.
- Stoeckle, M. Y., Adolf, J., Charlop-Powers, Z., Dunton, K. J., Hinks, G., & VanMorter, S. M. (2021). Trawl and eDNA assessment of marine fish diversity, seasonality, and relative abundance in coastal New Jersey, USA. *ICES Journal of Marine Science*, 78, 293–304.
- Taberlet, P., Bonin, A., Zinger, L., & Coissac, E. (2018). *Environmental DNA for biodiversity research and monitoring*. Oxford University Press.
- Tsuji, S., Shibata, N., Sawada, H., & Ushio, M. (2020). Quantitative evaluation of intraspecific genetic diversity in a natural fish population using environmental DNA analysis. *Molecular Ecology Resources*, 20, 11323–11332.
- Turunen, J., Mykrä, H., Elbrecht, V., Steinke, D., Braukmann, T., & Aroviita, J. (2021). The power of metabarcoding: Can we improve bioassessment and biodiversity surveys of stream macroinvertebrate communities? *Metabarcoding Metagenomics*, 5, e68938.
- Ushio, M., Murakami, H., Masuda, R., Sado, T., Miya, M., Sakurai, S., Yamanaka, H., Minamoto, T., & Kohdoh, M. (2018). Quantitative monitoring of multispecies fish environmental DNA using high-throughput sequencing. *Metabarcoding Metagenomics*, 2, 1–15.
- Valsecchi, E., Coppola, E., Pires, R., Parmegiani, A., Casiraghi, M., Galli, P., & Bruno, A. (2022). A species-specific qPCR assay provides novel insight into range expansion of the Mediterranean monk seal (*Monachus monachus*) by means of eDNA analysis. *Biodiversity and Conservation*, 31, 1175–1196.
- Weltz, K., Lyle, J. M., Ovenden, J., Morgan, J. A. T., Morendo, D. A., & Semmens, J. M. (2017). Application of environmental DNA to detect an endangered marine skate species in the wild. *PLoS One*, 12, e0178124.
- Yates, M. C., Glaser, D., Post, J., Cristescu, M. E., Fraser, D. J., & Derry, A. M. (2020). The relationship between eDNA particle concentration and organism abundance in nature is strengthened by allometric scaling. *Molecular Ecology*, 30, 1–15.
- Zemb, O., Achard, C. S., Hamelin, J., De Almeida, M.-L., Gabinaud, B., Cauquil, L., Verschuren, L. M. G., & Godon, J. J. (2019). Absolute quantification of microbes using 16S rRNA gene metabarcoding: A rapid normalization of relative abundances by quantitative PCR targeting a 16S rRNA gene spike-in standard. *Microbiology Open*, 9, e977.

## SUPPORTING INFORMATION

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

**How to cite this article:** Stoeckle, M. Y., Ausubel, J. H., & Coogan, M. (2022). 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. *Environmental DNA*, 00, 1–11. <https://doi.org/10.1002/edn3.376>