The potential for citizen science divers to monitor biodiversity through passive eDNA collection in the Gulf of California

Theodora Mautz

Julie Dinasquet / Nastassia Patin

Brice Semmens

Marine Biology Departmental Exam

Scripps Institution of Oceanography, UC San Diego

May 28, 2024

Abstract

Environmental DNA (eDNA) marker gene sequencing, or "metabarcoding," can be a powerful tool for generating marine community composition and biodiversity estimates. Traditional sampling methods involve manual water collection and active filtration to collect DNA. This process can be laborious, expensive, prone to contamination, and require filtration equipment. To explore the possibility of making eDNA a more accessible tool, we investigated the efficacy of passive eDNA capture during recreational diving for fish and marine mammal detection in the Gulf of California. Filters were attached to citizen science SCUBA divers and passively collected eDNA from seawater as the divers swam. The divers concurrently collected visual survey data on fish species. eDNA from these filters was sequenced using MiFish (12S rRNA gene) and mitochondrial D-loop primers for fish and marine mammals, respectively. The eDNA sequencing results initially yielded 66 fish species, compared to 183 total fish species observed in the divers' visual surveys. Notably, some of those species live in areas inaccessible to divers, demonstrating one advantage of eDNA filters over visual surveys for species detection. However, 90 species observed visually were not listed in the sequence reference database used to assign eDNA taxonomy. To investigate if this is due to lack of representative sequences in the database or to eDNA biases, we are currently collaborating with the Scripps Institution of Oceanography Marine Vertebrate Collection to sequence these missing species. Interestingly, the marine mammal sequencing included Guadalupe fur seal, which is a threatened species, and highlights the potential of eDNA for rare or threatened marine mammal detection. Overall, our results suggest that passive capture of eDNA can be used as a tool for citizen scientists and supplement visual surveys for marine species detection, and is particularly useful for species that are rare or visually inaccessible to divers.

Introduction

Monitoring biodiversity is an important component of conservation because increased biodiversity is linked to ecosystem stability and efficiency.[1–5] In marine ecosystems, Anthropogenic effects ranging from overfishing to rising ocean temperatures threaten biodiversity by decimating fish populations.[6–8] Traditional monitoring schemes include observation-based techniques, including fishing, trapping, or catching species, deploying underwater videos, or conducting underwater visual or acoustic surveys. These methods can be variously restricted by time, cost, visual or acoustic conditions, sampling inefficiencies, and invasiveness.[9,10] Notably, all rely on taxonomic identification based on human sensory ability.[10]

Over the past decade, environmental DNA (eDNA) metabarcoding has emerged as a non-invasive and efficient alternative or supplemental method for aquatic biodiversity monitoring.[9,11,12] eDNA consists of both the intra- and extracellular components of DNA that organisms shed into their surrounding environment and can include feces, sloughed cells, and gametes.[13,14] Researchers use eDNA metabarcoding to extract, amplify, and sequence the eDNA from seawater samples using primer sets for particular taxa groups. These sequences are then compared to known sequences in reference databases to identify the species from which the eDNA originated.[13,15] eDNA has demonstrated high efficacy for detection of the presence and diversity of many marine species, including fishes.[9,10,16] The use of eDNA for species abundance is more complicated, and studies have found unreliable relationships between eDNA sequence read numbers and conventional estimates for abundance, such as trawls.[13,17] Despite this limitation, eDNA metabarcoding offers several advantages over traditional biodiversity approaches in marine ecosystems. First, by applying molecular genetics for taxonomic assignment, eDNA obviates the need for taxonomic expertise.[10,17,18] Second, eDNA metabarcoding can detect species that can elude other monitoring methods, such as rare, cryptic, nocturnal, small, and vulnerable species.[16,17,19] Finally, eDNA provides a non-invasive form of species monitoring that bypasses harmful capture processes.[18]

The most common method for marine eDNA collection is by filtering seawater through a membrane (active filtration). However, because concentrations of eDNA can be low or show

small-scale heterogeneity in aquatic environments, previous studies have recommended that active filtration eDNA studies maximize sampled water volume in order to achieve sampling sufficiency.[9,10,15] Collecting and actively filtering seawater is a labor- and time-intensive process that itself mandates specialized equipment, storage space, and often transportation for the water prior to filtration. These longer processing times can lead to more eDNA degradation and increase the possibility for contamination. Additionally, because samples come from one area of water at one time, active filtration represents only a spatiotemporal snapshot of the ecosystem.[18]

Recently, passive sampling of eDNA has emerged as an attractive alternative to active filtration, and features the direct collection of eDNA from the marine environment through submersion of DNA-adsorbent materials.[13,15,18,20] By removing the active filtration step, passive eDNA collection offers a faster, cheaper, and more simple way to collect DNA from seawater.[15,20] Recent studies have demonstrated that passive eDNA collection can be as effective as active filtration with only 5 minutes of submersion.[15] The downstream effects of these benefits include increased sample replication and expanded access to eDNA collection since it does not require specialized water filtration equipment or expertise in such equipment.[13] In turn, these advantages greatly expand the possible ecological questions that can be addressed through marine eDNA collection.[13]

As passive eDNA collection remains a relatively novel method, few studies have investigated its efficacy in marine environments.[9,13,18,20] Those that have can be separated into studies investigating natural or artificial materials for passive eDNA collection. Among the natural materials group, Mariani et al. (2019), and more recently, Jeunen et al. (2024[21]), successfully extracted tens of thousands of fish DNA reads from natural sponges in Antarctica and the Mediterranean Sea,[22] and Turon et al. (2020) demonstrated their sampling value in tropical biodiversity surveys in Nha Trang Bay.[23] Other studies have also demonstrated the efficacy of shrimp gut contents[24] and aquatic biofilms[25] for indirectly surveying environmental fauna. However, the performances of these natural materials can depend on the environments in which they are deployed and can be difficult to standardize.[9,18]

Alternatively, artificial materials can offer increased application and standardization potential. A few studies have tested various “passive eDNA sampler” (PEDS) adsorbents, including

montmorillonite clay and granular activated carbon,[20] as well as filter membranes made of nylon,[9,13] cellulose ester,[13] and mixed cellulose acetate and nitrate.[9] Additionally, Chen et al. (2022) and Bessey et al. (2022) compared 12 and 9 sorbent materials, respectively, for passive eDNA collection.[15,18] These latter two studies demonstrated that several of these inexpensive materials can be very effective for passive eDNA collection, with Chen et al. (2022) finding that glass fiber filters (GFF) outperformed all other materials in both laboratory and field settings.

One potential method for deploying PEDS is to engage citizen scientists. Citizen science refers to science conducted by volunteer members of the public.[26] There has been significant debate over this term, with some institutions favoring "community science" over "citizen science" because the former is more inclusive than the latter. This paper will use citizen science because community science has historically held a different meaning (one in which community members lead the science and project goals are primarily social rather than scientific), and its substitution in this context would be an appropriation of community-lead scientific endeavors.[26,27] Even so, it must be acknowledged that the term citizen science conveys an exclusive and legalistic tone and hopefully one day, a new term can be used in its place.

Citizen science has proven valuable for biodiversity monitoring in many ecosystems.[28–31] It offers a reliable, cost-effective mechanism for biodiversity monitoring while fostering environmental stewardship among the public.[32] In marine ecosystems, recreational SCUBA divers can act as citizen scientists by conducting visual surveys of marine species, and data from these surveys have been invaluable for biodiversity monitoring efforts and subsequent conservation management decisions.[32–35] Citizen scientists have also assisted with aquatic eDNA collection and filtering,[36–38] and citizen scientist observations have been used to supplement eDNA analyses;[39] however, the efficacy of equipping citizen scientist SCUBA divers with PEDS has not yet been explored. Interestingly, in a chapter from the 2018 book, *Advances in Ecological Research*, Pocock et al. (2018) write of two ways to accelerate biodiversity monitoring: "through the application of new technologies, for example, [...] eDNA [...] [or, a] complementary approach is to use citizen science."[28] This study combines the two.

In the present study, we use eDNA metabarcoding to evaluate the effectiveness of passive eDNA collection by citizen science SCUBA divers in the Gulf of California. To our knowledge, this study is the first time citizen scientist SCUBA divers have passively collected eDNA by swimming through seawater. Our objectives were to:

1) Determine whether glass fiber filter PEDS attached to freely swimming divers could successfully adsorb eDNA, and
2) Compare alpha and beta diversity metrics for fish species from the PEDS and from visual surveys concurrently conducted by the divers.

We hypothesized that PEDS attached to recreational divers would facilitate robust eDNA collection via adsorption, and would yield comparable fish biodiversity metrics as the visual surveys. When our data revealed several gaps in online reference databases used for metabarcoding for fish species in the Gulf of California, we added a third objective:

3) Sequence fish species that are commonly observed in diver visual surveys but missing from online reference databases, add these sequences to the databases, then re-evaluate diversity metrics.

We suggest that citizen science divers can successfully collect eDNA through PEDS and that PEDS can complement and offer certain advantages over concurrent visual surveys of fish species.

Methods

Sampling location

Samples were collected from 33 dive sites ranging from 24° 23.536′ N, -110° 18.305′W at the southern end to 29° 32.978′ N, -113° 32.845′ W at the northern end in the Gulf of California. Widely recognized as a hotspot for marine biodiversity, the Gulf is one of the most important fishing regions in Mexico, as well as a marine system of great importance to the international conservation community.[40–42] Figure 1 illustrates the map of all sampling sites within the Gulf of California. Sampling sites were divided into three groups: northern sites (latitudes 28.0°N-30.0°N), middle sites (latitudes 25.7°N-28.0°N), and southern sites (latitudes 24.0° N-25.7°N).

Sampling materials

In line with studies that have shown the outperformance of glass fiber filter (GFF) membranes for passive aquatic eDNA collection over other common materials,[18] we selected GFF membranes with nominal pore sizes of 0.7μm for our passive eDNA filters. These filters were fitted into 3D printed plastic honeycomb pucks. Together, these materials will henceforth be referred to as the PEDS.

eDNA collection

Passive eDNA collection was facilitated by trained volunteer divers on a diving trip organized by Reef Environmental Education Foundation (REEF). Dives were conducted between September 25 and October 5, 2022. The PEDS were attached to divers' first stage regulators by REEF staff wearing gloves, and PEDS were covered with clean gloves or plastic bags before divers entered the water to prevent contamination. Divers then went on dives that ranged from 18 to 75 minutes, freely swimming through the water. When divers surfaced, REEF staff donning clean gloves removed the PEDS from diver regulators, unscrewed the honeycomb pucks, used clean tweezers to remove the GFFs, and placed them into individually marked vials filled with Longmire buffer.[43] Four samples and one blank filter (in which the PEDS was transferred directly from its box to the vial) were collected from each dive. These vials were transported and stored at room temperature until extractions began in March 2023.

Visual survey collection

Visual surveys were conducted using the Roving Diver Technique, in which divers freely swim through dive sites with slates, preprinted data sheets, and pencils and record all observed fish species throughout the water column.[44–46] Abundance values are scored between 1-4, where 1 means "Single" (1 fish), 2 means "Few" (2-10 fish), 3 means "Many" (11-100 fish), and 4 means "Abundant" (>100 fish). At the end of each dive, survey metadata such as habitat, temperature, site, and maximum depth are added to the data sheets, which are then returned to REEF staff and later digitized. All volunteer divers go through extensive training in fish species identification, and the efficacy and value of these surveys have been applied to marine species monitoring, management, and conservation for decades.[45,47,48]

DNA extraction

eDNA was extracted in a dedicated molecular biology laboratory where equipment and benches were regularly cleaned and sterilized. Total nucleic acid was extracted from all samples using the *Quick*-DNA Microprep Kit (Zymo) following the manufacturer's instructions for "Cell Suspensions and Proteinase K Digested Samples," with the following modifications: samples were centrifuged for 1 minute at ≥10,000g after step 1, 400µL of DNA Pre-Wash Buffer was added in step 3, and 200µL of gDNA Wash was added after step 4 with an additional centrifuge for 1 minute at ≥10,000g. DNA was eluted into 50µL of DNA Elution Buffer. Two blank filters were extracted to check for contamination in the filters or during the extraction process, and were discarded after analysis by Qubit fluorometer (Thermo Scientific) revealed no DNA in either control sample.

PCR amplification and sequencing

Two initial PCRs were conducted (Table1). The first was conducted using MiFish-U (Universal) primers, which target a hypervariable region in the 12S rRNA gene (163-185 bp) and was developed for metabarcoding eDNA from fishes.[49] A modified version of the MiFish-U forward primer was used, from Sales et al. (2019).[50] The second was conducted using primers targeting the D-loop from the mitochondrial control region of marine mammals.[51]

Primer name	Forward Sequence	Reverse Sequence	Reference
MiFish-U (with reverse modification)	5'-GCCGGTAAAAC TCGTGCCAGC-3'	5'-CATAGTGGGGT ATCTAATCCCAGT TTG-3'	Sales et al. (2019)[50]
D-loop	5'-TCACCCAAAGC TGRARTTCTA-3'	5'-GCGGGTTGCTG GTTTCACG-3'	Baker et al. (2018)[51]

Table 1. Forward and reverse primers used to amplify fish (MiFish-U) and marine mammal (D-loop) eDNA.

PCR reagents included 2X Phusion Green Hot Start II High-Fidelity PCR Master Mix (Thermo Scientific), 10µM forward and reverse primers, Dimethyl sulfoxide (DMSO), nuclease-free water, Recombinant Albumin (rAlbumin), and 2µL of template DNA from each sample. PCR was performed using the following conditions: initial denaturation at 98°C for 30 seconds, followed by 35

cycles of 10 seconds at 98°C for denaturation, 30 seconds at 60°C for primer annealing, 45 seconds at 72°C, and a final extension for 10 minutes at 72°C. PCR products were purified using Exonuclease I Reaction Buffer (New England Biolabs) and Shrimp Alkaline Phosphatase (rSAP) (New England Biolabs). An additional PCR was then run on all samples with the addition of Illumina tag primers unique to each sample, as well as an additional purification using AxyPrep Mag PCR Clean-Up Protocol (Axygen Biosciences). The sequencing library was created by quantifying all PCR products using Quibit fluorometer (Thermo Scientific) and then pooling into equimolar ratios (range: 0.814 - 44.4μg/mL). Pooled samples were submitted to the UC Davis sequencing core (https://dnatech.genomecenter.ucdavis.edu/illumina-high-throughput-sequencing/) for 2 x 250 bp paired-end Illumina MiSeq sequencing.

Data processing

Following Illumina sequencing, data were filtered through quality control workflows. Primers were trimmed from sequences using Cutadapt. Reads longer than 110 bps were trimmed for forward and reverse sequences, empty samples were pruned, paired-end reads were merged, chimeras were removed, only merged reads between 163-185 bps were kept, and samples with fewer than 10 sequences were removed. Taxonomic assignment was conducted with DADA2 using the rCRUX Generated Universal MiFish Expanded reference database.[52,53]

Statistical analysis

All statistics and graphics were conducted or produced using R (version 4.3.2; R Development Core Team, 2023). For beta diversity analyses, PEDS data were standardized using the Hellinger transformation,[54,55] and NMDS values were calculated using the Gower dissimilarity index for categorical abundance data.[56] Phyloseq and vegan packages in R were used to calculate alpha and beta diversity values, SF package in R was used to generate maps and heatmaps, and and ANOVA, PERMANOVA, and linear regressions were used to assess significances of various metadata variables on diversity values.

Results

129 of 136 PEDS (95%) were fully sequenced and processed. A total of 359 REEF Visual Surveys (RVS) were included in analyses, with 14,137 total species entries. The PEDS were almost always attached to the same four divers (D1, D2, D3, D4), while fifteen total divers conducted RVS (D1 - D15). The habitats values for each RVS species entry were assigned categorically from 1 to 7, and the habitats reported in these surveys were 7 for rocky slope (8,364 entries), 1 for rock boulder or shale reef (5,105 entries), 3 for pinnacle (557 entries), 4 for sandy bottom/mud (68 entries), and 8 for seamount (43 entries). Maximum depth was represented categorically in RVS with 1 for snorkel, 2 for <10 feet, 3 for 10-19 feet, up to 13 for 110-119 feet, and ranged from 10-19 to 110-119 feet, and for PEDS, maximum depth ranged from 20 to 113 feet.

Fish species detection counts

In total, 71 species were recovered from eDNA analysis from the PEDS. After removing 4 terrestrial contaminant species and one "NA" for samples in which species-level taxonomy assignment could not be completed, the PEDS had detected 66 marine fish species. A list of these species can be found in Supplementary File 1.

The visual surveys conducted by REEF citizen science divers contained a total of 189 unique entries for marine fish species. After removing all species labeled with "sp." (indicating that species-level taxonomy could not be visually resolved), the REEF visual surveys (henceforth, RVS) yielded a total of 183 marine fish species. A list of these species can be found in Supplementary File 2. The PEDS and RVS had 30 species of overlap (Appendix Table 1), representing 45% of PEDS species and 16% of RVS species. The RVS included a total of 183 species and PEDS included a total of 66 species. A Venn diagram depicting species and family overlap can be seen in Figure 2.

Alpha diversity

Alpha diversity metrics were calculated for both PEDS and RVS data. Simpson's diversity index values organized by four metadata variables are shown for PEDS and RVS in Figures 3 and 4, respectively. ANOVAs were then run for each metadata variable and corresponding Simpson's values. None were found to have statistically significant relationships in PEDS but for RVS, the ANOVA

revealed a statistically significant ($p < 0.05$) relationship between maximum depth and Simpson's value and diver and Simpson's value (Table 2). We also ran a linear regression on Simpson's values and latitude for both PEDS and RVS but found no significant relationships.

	PEDS		**RVS**	
	F	*p*	*F*	*p*
Latitude group	0.248	0.781	1.919	0.148
Habitat	1.681	0.197	2.493	0.115
Depth	1.154	0.285	13.47	**0.000279***
Diver	0.704	0.647	41.78	**<2E-16***

Table 2. F-statistics and p-values for ANOVAs of four metadata variables and their respective Simpson's diversity index values for PEDS samples and RVS surveys. * denotes statistical significance with threshold $p < 0.05$.

Beta diversity

Beta diversity metrics were calculated for PEDS and RVS. NMDS values organized by four metadata variables are shown for PEDS and RVS in Figure 5. PERMANOVAs were then run for each metadata variable and corresponding NMDS values. Statistically significant relationships ($p < 0.05$) were found between latitude group and NMDS values and depth and NMDS values for both PEDS and RVS. Additionally, RVS PERMANOVA produced a statistically significant relationship between diver and NMDS values as well as between the interaction of latitude group and depth and NMDS values. And finally, PEDS PERMANOVA produced a statistically significant relationship between the interaction between latitude group and habitat and NMDS values.

	PEDS		**RVS**	
	F	*p*	*F*	*p*
Latitude group	2.6952	**0.001***	74.007	**0.001***
Habitat	1.3041	0.195	1.669	0.105
Depth	2.5761	**0.003***	10.991	**0.001***

Diver	1.0159	0.443	2.9311	**0.001***
Lat group, Habitat	2.1791	**0.001***	1.0165	0.380
Lat group, Depth	1.2109	0.206	6.2502	**0.001***
Habitat, Depth	1.2546	0.219	1.0141	0.366

Table 2. F-statistics and p-values for PERMANOVAs of four metadata variables and their respective NMDS values for PEDS samples and RVS surveys. Where multiple variables are listed (last three rows), multifactorial PERMANOVAs were run and variable interaction statistics are listed. * denotes statistical significance with threshold $p < 0.05$.

Comparison of methods

Due to the large differences in the numbers of species detected by each method, bar plots were created to visualize site composition by taxa using family data. Of the top 20 most common families in RVS, 9 families overlapped with PEDS families and these were the 9 families depicted in Figure 6a and 6b. Figure 6 highlights that Site 10 yielded notably more sightings of species within those families (RVS) and read abundances (PEDS) across both RVS and PEDS compared to other sites.

To examine the similarity in presence / absence data between PEDS and RVS, a heatmap was created that plots the similarity by site between PEDS and RVS using the number of species that were either detected by both methods (present in both) or not detected by both methods (absent in both) (Figure 7). Each site's similarity score ranged from 6 to 20. Appendix Table 2 lists these scores for each site and their proportion to the total number of overlapping species (30). To quantify this relationship, we plotted similarity scores against latitude and ran a linear regression, which yielded a p-value of 0.0508 (Figure 8).

Species in reference database

As 83% of species observed in RVS were not detected by the PEDS, we wanted to determine whether this discrepancy was attributable to methodological limitations of the PEDS or to other factors. We first searched through the sequence reference database used to assign taxonomy to PEDS sequences. This revealed that over half of all detected RVS species, or 95 species, were not in the

reference database. Efforts to mine publicly available sequence data provided only 5 additional species, yielding 4 new species detections in the eDNA, which are included in the 66 species count for PEDS.

Marine mammals

Marine mammal sequences matched only three species: *Tursiops truncatus* (common bottlenose dolphin), *Zalophus californianus* (California sea lion), and *Arctocephalus townsendi* (Guadalupe fur seal).

Discussion

Our results demonstrate that GFF PEDS can be effectively used by citizen science SCUBA divers to collect passive eDNA. 95% of all samples contained eDNA that was successfully amplified using MiFish-U primers, demonstrating the utility of this new technology for marine fish detection in the Gulf of California. Additionally, the eDNA in these PEDS was stored for several months in Longmire buffer at room temperature before being extracted. The PEDS in this study required minimal time or specialized equipment, and necessitated no technical or scientific expertise. The only materials that were needed were the filters, pucks and zipties, pre-filled vials with Longmire buffer, gloves, and tweezers. The only time and actions required were the few seconds of attaching and removing pucks and filters. These findings therefore greatly expand the accessibility of eDNA research as a tool for marine biodiversity monitoring, particularly through engaging citizen scientist SCUBA divers for passive eDNA collection.

RVS and PEDS had 30 species of overlap, suggesting that PEDS can complement visual surveys for species detection. Interestingly, PEDS also detected 36 species that were not observed in RVS. Some of these species include *Engraulis mordax* (Californian anchovy), *Hemanthias signifer* (Damsel bass), and *Umbrina roncador* (Yellowfin croaker). It makes sense that *Engraulis mordax* would not have been detected by RVS: 25 entries in RVS are identified only as “Atheriniformes sp./ Clupeiformes sp.,” or as “Unidentified Silvery Fish (Silversides / Anchovies / Herrings).” Silvery fish such as anchovies can be hard to visually distinguish at the species level, indicating one possible advantage of PEDS over RVS as PEDS do not depend on visual cues that may be subtle or absent and can resolve species-level

differences in an quantitative manner. As for *Hemanthias signifer* and *Umbrina roncador*, the former's habitat extends to over 1000 feet and can often be found towards the bottom, too deep for divers,[57] while the latter spends much of its time in the surf zone, too shallow to be seen by recreational divers.[58] Thus, PEDS offer additional advantages over visual surveys by enabling detection of species that do not live in waters accessible to divers.

While the marine mammal D-loop sequencing only yielded three species, the *Arctocephalus townsendi* (Gudalupe fur seal) is an interesting detection. This species was believed to be driven to extinction by 19th century hunting, but was seen again in the 1950s. Since then, the species recovery has been slow, and it is included in the Convention on International Trade in Endangered Species of Wild Fauna and Flora (CITES) 2017 Appendix.[59] Its detection by PEDS in this study suggests that PEDS may be effective for the detection of rare or threatened species, and may contribute to the monitoring of these species.

For alpha diversity metrics, RVS ANOVAs produced significant results for the relationships between maximum depth and Simpson's diversity index values, and diver and Simpson's values. The former finding reveals a relationship between community composition and depth, which is corroborated by other studies that have found that marine fish community compositions differ at sampling depths that differ by as little as 33 feet.[9,60] Interestingly, a PEDS study by Zhang et al. (2024) found that PEDS can be more sensitive to the detection of differently abundant species at different sampling depths, which could explain why the ANOVA for PEDS did not identify maximum depth as having a significant relationship with Simpson's values.[9] As for the effects of diver (which RVS also found had a significant relationship with NMDS values in beta diversity metrics), this finding suggests observer bias among the citizen scientist divers. Other research has previously documented observer bias in visual survey detection probabilities of marine fishes, which can be mitigated in future studies by independent double-observer approaches.[61,62] It is worth noting that PEDS diversity statistics did not highlight diver as a statistically significant variable at all in alpha or beta diversity statistical analyses, raising the possibility that citizen science divers can use PEDS equally effectively, regardless of individual diving preferences or swimming efficiency.

For beta diversity metrics, both PEDS and RVS beta diversity PERMANOVAs suggested significant relationships between latitude and NMDS values and maximum depth and NMDS values. However, the F-statistic for RVS was much higher than for PEDs (74.0 versus 2.7 for latitude group and 11.0 versus 2.6 for maximum depth), which is consistent with the robustness of RVS species data compared to PEDS species data. RVS also found a significant relationship between the interaction of latitude group and depth with NMDS values, while PEDS found a significant relationship between the interaction of latitude group and habitat.

The Gulf of California has several features that could account for the significance of latitude group on beta diversity. First, much of the surface circulation of the southern gulf (specifically the peninsula side, where all of the sites in this study were) is driven by mesoscale eddies in the summer,[63] and these eddies last from August until October,[64] which encompasses the dates of PEDS and RVS data collection. These eddies cause additional water mixing, which could impact the ecology of eDNA in the southern group compared to the northern or middle groups. Second, the northern Gulf of California is less tropical, subject to more annual fluctuations in productivity, and has a different Temperature-Salinity profile than the central and southern regions.[65,66] Finally, a previous study compared fish biodiversity between three latitude groups in the Gulf of California and similarly found that Shannon diversity values showed statistically significant differences among latitude groups. Their northern latitude group matches the one in this study (≥28°N), although their "central zone" encompasses both of our southern and central zones and ranges from 27°N to 24°N.[66] Other studies have found latitudinal gradients or biogeographic regionalizations of species composition in the Gulf of California using both eDNA metabarcoding and underwater visual surveys,[67,68] further supporting the significance of latitude group on diversity values in the present study.

RVS surveys detected 117 more species than PEDS, or nearly 2.8 times the number of species that PEDS detected. However, as our genetic database mining revealed, around half of all RVS species did not have verified sequence vouchers in any online database. As a result, it's impossible to precisely evaluate the efficacy of the PEDS when there are so many gaps in the database that assigns taxonomy to

PEDS sequences. As will be discussed in the next section, sequencing the missing species and adding them to online databases will be a priority for this project going forward.

Despite this limitation, beta diversity metrics were similar between PEDS and RVS. While the effectiveness of eDNA to measure species abundances has had mixed results,[10,17,69] Figure 6 demonstrates some overlap between the relative abundances of the PEDS (as computed from the number of reads for each sequence in each sample), and the relative numbers of survey entries (or "sightings") of each species in RVS surveys, particularly around Site 10. Additionally, site-wise comparisons indicate significant overlap in common species that had identical presence / absence data in both RVS and PEDS (Figure 7a), with 70% of sites demonstrating equal to or over 50% similarity in presence / absence data for the 30 overlapping species. Finally, PEDS and RVS PERMANOVAs both highlighted latitude group and depth as having statistically significant relationships with NMDS values for each group. Thus, even with several species observed in RVS missing from the reference database, PEDS beta diversity statistics still reveal similar relationships as RVS beta diversity statistics.

Overall, citizen science SCUBA divers can effectively use PEDS for passive eDNA collection. Moreover, PEDS can serve as a quantitative, genetics-based complement to citizen science visual surveys for marine biodiversity. Improving sequence reference databases will be an important next step in determining just how effective PEDS are when compared to visual surveys of marine fish species in the Gulf of California.

Future plans

I have initiated a collaboration with Professor Dahiana Arcila and Ben Frable and intend to extract and sequence the DNA of all species in the Scripps Institution of Oceanography Marine Vertebrate Collection that were detected in RVS but absent from the rCRUX Generated Universal MiFish Expanded reference database.[52,53] We will then add these sequences to open access databases. One thing that has become clear is that even for species that have available sequences in reference databases, sometimes taxonomy assignment will assign a closely related but separate species to our sequences. For example, one of the species detected by eDNA is *Caranx hippos*, or the Crevalle jack.

However, this species can exclusively be found in the Atlantic Ocean.[70] When we ran the sequence from our eDNA against GenBank using NCBI's Basic Local Alignment Search Tool (BLAST), we found that in addition to its 100% identity match to *Caranx hippos*, the sequence also had a 98.82% identity match to *Caranx sexfasciatus*. Notably, *C. sexfasciatus* was observed in RVS on the same dive as PEDS detected *C. hippos*. As a result, it seems likely that because the MiFish-U 12S sequence for these two species is very similar, the *C. hippos* detection represents an error with taxonomic assignment. As such, part of this process will also involve double checking by alternative taxonomic assignment processes that all PEDS species are correctly identified. After this is done, we will rerun all bioinformatic analyses.

Additionally, to improve eDNA capture on the PEDS, we are testing different filter materials, pore and chemical pre-treatments for better DNA adsorption; We hope these results will help guide future PEDS research and citizen science eDNA initiatives.

References

1. Cardinale, B. J. *et al.* Biodiversity loss and its impact on humanity. *Nature* **486**, 59–67 (2012).
2. Taberlet, P., Bonin, A., Zinger, L. & Coissac, E. Environmental DNA for functional diversity. in *Environmental DNA: For Biodiversity Research and Monitoring* 90–98 (Oxford University PressOxford, 2018). doi:10.1093/oso/9780198767220.003.0010.
3. Duffy, J. E. *et al.* The functional role of biodiversity in ecosystems: incorporating trophic complexity. *Ecol. Lett.* **10**, 522–538 (2007).
4. Loreau, M. & de Mazancourt, C. Biodiversity and ecosystem stability: a synthesis of underlying mechanisms. *Ecol. Lett.* **16**, 106–115 (2013).
5. Pennekamp, F. *et al.* Biodiversity increases and decreases ecosystem stability. *Nature* **563**, 109–112 (2018).
6. Worm, B. & Lotze, H. K. Chapter 21 - Marine biodiversity and climate change. in *Climate Change (Third Edition)* (ed. Letcher, T. M.) 445–464 (Elsevier, 2021). doi:10.1016/B978-0-12-821575-3.00021-9.
7. Ramírez, F., Afán, I., Davis, L. S. & Chiaradia, A. Climate impacts on global hot spots of marine biodiversity. *Sci. Adv.* **3**, e1601198 (2017).
8. Jackson, J. B. C. *et al.* Historical Overfishing and the Recent Collapse of Coastal Ecosystems. *Science* **293**, 629–637 (2001).
9. Zhang, L. *et al.* Use of passive sampling in environmental DNA metabarcoding technology: Monitoring of fish diversity in the Jiangmen coastal waters. *Sci. Total Environ.* **908**, 168298 (2024).
10. Bessey, C. *et al.* Maximizing fish detection with eDNA metabarcoding. *Environ. DNA* **2**, 493–504 (2020).
11. Foote, A. D. *et al.* Investigating the Potential Use of Environmental DNA (eDNA) for Genetic Monitoring of Marine Mammals. *PLOS ONE* **7**, e41781 (2012).
12. Thomsen, P. F. *et al.* Detection of a Diverse Marine Fish Fauna Using Environmental DNA from Seawater Samples. *PLOS ONE* **7**, e41732 (2012).
13. Bessey, C. *et al.* Passive eDNA collection enhances aquatic biodiversity analysis. *Commun. Biol.* **4**, 1–12 (2021).
14. Taberlet, P., Bonin, A., Zinger, L. & Coissac, E. Introduction to environmental DNA (eDNA). in *Environmental DNA: For Biodiversity Research and Monitoring* 1–6 (Oxford University PressOxford, 2018). doi:10.1093/oso/9780198767220.003.0001.
15. Bessey, C. *et al.* Comparison of materials for rapid passive collection of environmental DNA. *Mol. Ecol. Resour.* **22**, 2559–2572 (2022).
16. Dalongeville, A. *et al.* Benchmarking eleven biodiversity indicators based on environmental DNA surveys: More diverse functional traits and evolutionary lineages inside marine reserves. *J. Appl.*

Ecol. **59**, 2803–2813 (2022).
17. Veron, P. *et al.* Environmental DNA complements scientific trawling in surveys of marine fish biodiversity. *ICES J. Mar. Sci.* **80**, 2150–2165 (2023).
18. Chen, X. *et al.* Comparative Evaluation of Common Materials as Passive Samplers of Environmental DNA. *Environ. Sci. Technol.* **56**, 10798–10807 (2022).
19. Nester, G. M. *et al.* Development and evaluation of fish eDNA metabarcoding assays facilitate the detection of cryptic seahorse taxa (family: Syngnathidae). *Environ. DNA* **2**, 614–626 (2020).
20. Kirtane, A., Atkinson, J. D. & Sassoubre, L. Design and Validation of Passive Environmental DNA Samplers Using Granular Activated Carbon and Montmorillonite Clay. *Environ. Sci. Technol.* **54**, 11961–11970 (2020).
21. Jeunen, G.-J. *et al.* Characterizing Antarctic fish assemblages using eDNA obtained from marine sponge bycatch specimens. *Rev. Fish Biol. Fish.* **34**, 221–238 (2024).
22. Mariani, S., Baillie, C., Colosimo, G. & Riesgo, A. Sponges as natural environmental DNA samplers. *Curr. Biol.* **29**, R401–R402 (2019).
23. Turon, M., Angulo-Preckler, C., Antich, A., Præbel, K. & Wangensteen, O. S. More Than Expected From Old Sponge Samples: A Natural Sampler DNA Metabarcoding Assessment of Marine Fish Diversity in Nha Trang Bay (Vietnam). *Front. Mar. Sci.* **7**, (2020).
24. Siegenthaler, A. *et al.* Metabarcoding of shrimp stomach content: Harnessing a natural sampler for fish biodiversity monitoring. *Mol. Ecol. Resour.* **19**, 206–220 (2019).
25. Rivera, S. F. *et al.* Fish eDNA metabarcoding from aquatic biofilm samples: Methodological aspects. *Mol. Ecol. Resour.* **22**, 1440–1453 (2022).
26. Lin Hunter, D. E., Newman, G. J. & Balgopal, M. M. What's in a name? The paradox of citizen science and community science. *Front. Ecol. Environ.* **21**, 244–250 (2023).
27. Cooper, C. B. *et al.* Inclusion in citizen science: The conundrum of rebranding. *Science* **372**, 1386–1388 (2021).
28. Pocock, M. J. O. *et al.* Chapter Six - A Vision for Global Biodiversity Monitoring With Citizen Science. in *Advances in Ecological Research* (eds. Bohan, D. A., Dumbrell, A. J., Woodward, G. & Jackson, M.) vol. 59 169–223 (Academic Press, 2018).
29. Bonney, R. *et al.* Citizen Science: A Developing Tool for Expanding Science Knowledge and Scientific Literacy. *BioScience* **59**, 977–984 (2009).
30. Chandler, M. *et al.* Contribution of citizen science towards international biodiversity monitoring. *Biol. Conserv.* **213**, 280–294 (2017).
31. Amano, T., Lamming, J. D. L. & Sutherland, W. J. Spatial Gaps in Global Biodiversity Information and the Role of Citizen Science. *BioScience* **66**, 393–400 (2016).
32. Branchini, S. *et al.* Using a citizen science program to monitor coral reef biodiversity through space and time. *Biodivers. Conserv.* **24**, 319–336 (2015).

33. Lucrezi, S., Milanese, M., Palma, M. & Cerrano, C. Stirring the strategic direction of scuba diving marine Citizen Science: A survey of active and potential participants. *PLOS ONE* **13**, e0202484 (2018).
34. Cerrano, C., Milanese, M. & Ponti, M. Diving for science - science for diving: volunteer scuba divers support science and conservation in the Mediterranean Sea. *Aquat. Conserv. Mar. Freshw. Ecosyst.* **27**, 303–323 (2017).
35. Ben Lamine, E., Di Franco, A., Romdhane, M. S. & Francour, P. Can citizen science contribute to fish assemblages monitoring in understudied areas? The case study of Tunisian marine protected areas. *Estuar. Coast. Shelf Sci.* **200**, 420–427 (2018).
36. Valsecchi, E. *et al.* Playing "hide and seek" with the Mediterranean monk seal: a citizen science dataset reveals its distribution from molecular traces (eDNA). *Sci. Rep.* **13**, 2610 (2023).
37. Agersnap, S. *et al.* A National Scale "BioBlitz" Using Citizen Science and eDNA Metabarcoding for Monitoring Coastal Marine Fish. *Front. Mar. Sci.* **9**, (2022).
38. Miya, M., Sado, T., Oka, S. I. & Fukuchi, T. The use of citizen science in fish eDNA metabarcoding for evaluating regional biodiversity in a coastal marine region: A pilot study. *Metabarcoding Metagenomics* **6**, 133–144 (2022).
39. Feng, W. & Lougheed, S. C. Integrating eDNA and citizen science observations to model distribution of a temperate freshwater turtle near its northern range limit. *PeerJ* **11**, e15120 (2023).
40. Lluch-Cota, S. E. *et al.* The Gulf of California: Review of ecosystem status and sustainability challenges. *Prog. Oceanogr.* **73**, 1–26 (2007).
41. Lluch-Cota, S. E. *et al.* Changing climate in the Gulf of California. *Prog. Oceanogr.* **87**, 114–126 (2010).
42. Munguia-Vega, A. *et al.* Ecological guidelines for designing networks of marine reserves in the unique biophysical environment of the Gulf of California. *Rev. Fish Biol. Fish.* **28**, 749–776 (2018).
43. Longmire, J. L., Maltbie, M. & Baker, R. J. Use of 'Lysis Buffer' in DNA isolation and its implication for museum collections. *Mus. Tex. Tech Univ.* **163**, (1997).
44. Schmitt, E. F. & Sullivan, K. M. Analysis of a Volunteer Method for Collecting Fish Presence and Abundance Data in the Florida Keys. *Bull. Mar. Sci.* **59**, 404–416 (1996).
45. Pattengill-Semmens, C. & Semmens, B. Assessment and Monitoring Applications of a Community-Based Monitoring Program: The Reef Environmental Education. (1999).
46. Ward-Paige, C. A. *et al.* Large-Scale Absence of Sharks on Reefs in the Greater-Caribbean: A Footprint of Human Pressures. *PLoS ONE* **5**, e11968 (2010).
47. Greenberg, D. A., Pattengill-Semmens, C. V. & Semmens, B. X. Assessing the value of citizen scientist observations in tracking the abundance of marine fishes. *Conserv. Lett.* **n/a**, e13009.

48. Pattengill-Semmens, C. V. & Semmens, B. X. Conservation and Management Applications of the Reef Volunteer Fish Monitoring Program. in *Coastal Monitoring through Partnerships: Proceedings of the Fifth Symposium on the Environmental Monitoring and Assessment Program (EMAP) Pensacola Beach, FL, U.S.A., April 24–27, 2001* (eds. Melzian, B. D., Engle, V., McAlister, M., Sandhu, S. & Eads, L. K.) 43–50 (Springer Netherlands, Dordrecht, 2003). doi:10.1007/978-94-017-0299-7_5.
49. 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**, 150088 (2015).
50. Sales, N. G., Wangensteen, O. S., Carvalho, D. C. & Mariani, S. Influence of preservation methods, sample medium and sampling time on eDNA recovery in a neotropical river. *Environ. DNA* **1**, (2019).
51. Baker, C. S., Steel, D., Nieukirk, S. & Klinck, H. Environmental DNA (eDNA) From the Wake of the Whales: Droplet Digital PCR for Detection and Species Identification. *Front. Mar. Sci.* **5**, (2018).
52. Gold, Z., Curd, E., Gallego, R., Gal, L. & Nielsen, S. rCRUX Generated MiFish Universal 12S Expanded Reference Database. Zenodo https://doi.org/10.5281/zenodo.7908865 (2023).
53. Curd, E. E. *et al.* rCRUX: A rapid and versatile tool for generating metabarcoding reference libraries in R. *Environ. DNA* **6**, e489 (2024).
54. Legendre, P. & Gallagher, E. D. Ecologically meaningful transformations for ordination of species data. *Oecologia* **129**, 271–280 (2001).
55. Legendre, P., Borcard, D. & Peres-Neto, P. R. Analyzing Beta Diversity: Partitioning the Spatial Variation of Community Composition Data. *Ecol. Monogr.* **75**, 435–450 (2005).
56. Fong, C. R., Ryznar, E. R., Smith, L. L. & Fong, P. Towards a trait-based framework for marine macroalgae: Using categorical data to explore the nature of emergent functional groups. *J. Ecol.* **111**, 1848–1865 (2023).
57. Hemanthias signifer, Damsel bass. *FishBase* https://www.fishbase.se/summary/Hemanthias-signifer.html.
58. Pondella, D. J. *et al.* Demographic Parameters of Yellowfin Croaker, Umbrina roncador (Perciformes: Sciaenidae), from the Southern California Bight1. *Pac. Sci.* **62**, 555–568 (2008).
59. García-Aguilar, M. C., Elorriaga-Verplancken, F. R., Rosales-Nanduca, H. & Schramm, Y. Population status of the Guadalupe fur seal (Arctocephalus townsendi). *J. Mammal.* **99**, 1522–1528 (2018).
60. Andruszkiewicz, E. A. *et al.* Biomonitoring of marine vertebrates in Monterey Bay using eDNA metabarcoding. *PLOS ONE* **12**, e0176343 (2017).
61. Williams, I. D., Walsh, W. J., Tissot, B. N. & Hallacher, L. E. Impact of observers' experience level

on counts of fishes in underwater visual surveys. *Mar. Ecol. Prog. Ser.* **310**, 185–191 (2006).

62. Bernard, A. T. F., Götz, A., Kerwath, S. E. & Wilke, C. G. Observer bias and detection probability in underwater visual census of fish assemblages measured with independent double-observers. *J. Exp. Mar. Biol. Ecol.* **443**, 75–84 (2013).
63. Lavín, M. F. *et al.* Surface circulation in the Gulf of California in summer from surface drifters and satellite images (2004–2006). *J. Geophys. Res. Oceans* **119**, 4278–4290 (2014).
64. Lavín, M. F., Castro, R., Beier, E. & Godínez, V. M. Mesoscale eddies in the southern Gulf of California during summer: Characteristics and interaction with the wind stress. *J. Geophys. Res. Oceans* **118**, 1367–1381 (2013).
65. Bray, N. A. Water mass formation in the Gulf of California. *J. Geophys. Res. Oceans* **93**, 9223–9240 (1988).
66. Fernández-Rivera Melo, F. J. *et al.* Latitudinal variation in structure and function of conspicuous reef fish assemblages along the western Gulf of California. *Rev. Mex. Biodivers.* **89**, 1154–1166 (2018).
67. Valdivia-Carrillo, T., Rocha-Olivares, A., Reyes-Bonilla, H., Domínguez-Contreras, J. F. & Munguia-Vega, A. Integrating eDNA metabarcoding and simultaneous underwater visual surveys to describe complex fish communities in a marine biodiversity hotspot. *Mol. Ecol. Resour.* **21**, 1558–1574 (2021).
68. Mac Loughlin, C. *et al.* eDNA metabarcoding warms up a hotspot of marine biodiversity: revealing underrepresented taxa in visual surveys and historical records from the Gulf of California. *Mar. Biodivers.* **54**, 22 (2024).
69. Stoeckle, M. Y. *et al.* Trawl and eDNA assessment of marine fish diversity, seasonality, and relative abundance in coastal New Jersey, USA. *ICES J. Mar. Sci.* **78**, 293–304 (2021).
70. Caranx hippos, Crevalle jack. *FishBase* https://fishbase.mnhn.fr/summary/Caranx-hippos.

Appendix

Appendix Table 1. Species detected by both PEDS and RVS

Scientific name
Anisotremus davidsonii
Balistes polylepis
Bodianus diplotaenia
Calamus brachysomus
Cephalopholis colonus
Chromis atrilobata
Cirrhitichthys oxycephalus
Coryphopterus urospilus
Diodon holocanthus
Diodon hystrix
Girella simplicidens
Haemulon flaviguttatum
Halichoeres dispilus
Harengula thrissina
Kyphosus analogus
Kyphosus azureus
Labrisomus xanti
Lutjanus argentiventris
Lythrypnus dalli
Microlepidotus inornatus
Microspathodon dorsalis
Paralabrax auroguttatus
Pomacanthus zonipectus

Scarus ghobban
Selar crumenophthalmus
Sphoeroides annulatus
Sphoeroides lobatus
Stegastes rectifraenum
Synodus lacertinus
Thalassoma lucasanum

Appendix Table 2. Species Overlap by Site with Proportional Score to Total Common Species

Site #	# Species Overlap	% of Common Species
1	16	0.53
2	16	0.53
3	20	0.67
4	17	0.57
5	16	0.53
6	14	0.47
7	6	0.20
8	15	0.50
9	14	0.47
10	15	0.50
11	18	0.60
12	16	0.53
13	18	0.60

14	14	0.47
15	18	0.60
16	15	0.50
17	17	0.57
18	20	0.67
19	19	0.63
20	20	0.67
21	18	0.60
22	15	0.50
23	17	0.57
24	12	0.40
25	13	0.43
26	13	0.43
27	15	0.50
28	17	0.57
29	13	0.43
30	15	0.50
31	18	0.60
32	14	0.47
33	14	0.47