eDNA metabarcoding as a biomonitoring tool for marine protected areas

Zachary Gold 1,\*, Joshua Sprague 2 , David J. Kushner 2 , Erick Zerecero1, Paul H. Barber 1

1Department of Ecology and Evolutionary Biology, University of California – Los Angeles

2Channel Islands National Park Service

**\*** Corresponding Author

Email: zack.gold@ucla.edu

**Abstract**

Monitoring of marine protected areas (MPAs) is critical for marine ecosystem management, yet current protocols rely on SCUBA-based visual surveys that are costly and time consuming, limiting their scope and effectiveness. Environmental DNA (eDNA) metabarcoding is a promising alternative for marine ecosystem monitoring, but more direct comparisons to visual surveys are needed to understand the strengths and limitations of each approach. This study compares fish communities inside and outside the Scorpion State Marine Reserve off Santa Cruz Island, CA using eDNA metabarcoding and underwater visual census surveys. Results from eDNA captured 76% (19/25) of fish species and 95% (19/20) of fish genera observed during pairwise underwater visual census. Species missed by eDNA were due to the inability of MiFish *12S* barcodes to differentiate species of rockfishes (*Sebastes,* n=4) or low site occupancy rates of crevice-dwelling *Lythrypnus* gobies. However, eDNA detected an additional 23 fish species not recorded in paired visual surveys, but previously reported from prior visual surveys, highlighting the sensitivity of eDNA. Significant variation in eDNA signatures by location (50m) and site (~1000m) demonstrates the sensitivity of eDNA to address key questions such as community composition inside and outside MPAs. Results demonstrate the utility of eDNA metabarcoding for monitoring marine ecosystems, providing an important complementary tool to visual methods.

**Keywords:** biomonitoring; Channel Islands National Park; eDNA; environmental DNA; marine protected areas; MPAs.

Introduction

Marine Protected Areas (MPAs) promote sustainability of marine ecosystems and the ecological goods and services they provide [1]. However, ensuring MPA effectiveness requires regular monitoring to document that ecosystem health is stable or improving [1]. MPA monitoring also provides an essential opportunity to assess the impact of management practices, allowing resource managers to adjust management plans as required [2].

Current MPA monitoring protocols typically assess the diversity and abundance of fish and benthic invertebrates, as well as community trophic structure [3]. Much of this assessment is based on underwater visual census surveys conducted on SCUBA [3], which are costly, and time and labor intensive [3]. For example, to survey 33 sites within the Channel Islands National Park once per year, the National Park Service Kelp Forest Monitoring Program spends over $395,000 on ~1,000 hours of dive time, ~1,400 hours of data entry to survey 33 sites within the Channel Islands National Park once per year (J. Sprague per obs., 2020). Furthermore, SCUBA-based surveys are constrained by weather, diving conditions, and personnel (J. Sprague per. obs., 2020), and can require extended and repeated dives to accurately document marine communities that place divers at risk for dive-related injuries. SCUBA surveys can also introduce significant observer bias, as fish react differently to divers, particularly inside and outside of MPAs, potentially impacting survey results [5].

Given the above logistical and methodological constraints, MPA monitoring efforts are largely limited to shallow depths (e.g. <30m) and the most economically or ecologically important taxa as proxies for ecosystem health [6]. Moreover, examining a predetermined subset of community diversity potentially excludes crucial functional groups, biasing ecosystem assessment [7]. Combined, these issues restrict the scope, scale, and frequency of visual surveys, limiting the utility of SCUBA-based MPA surveys to quantify biodiversity and trophic structure, [3]*,* data essential for assessing MPA effectiveness*.*

One promising new approach for assessing and monitoring marine ecosystems is environmental DNA, or “eDNA”, a technique based on isolation and sequencing of freely associated DNA from soil or water samples [8]. Through metabarcoding and high-throughput next generation sequencing, eDNA can broadly survey community biodiversity in a rapid, repeatable, and affordable manner [8]. As such, eDNA is ideally suited to measure the biodiversity for intensive monitoring programs, such as those required for MPAs [9].

eDNA has some key advantages over traditional SCUBA-based survey methods for biodiversity measurments. First, eDNA can capture a wide diversity of marine vertebrate taxa, frequently detecting more species than traditional fish survey methods [10]. Second, eDNA detects rare and cryptic species that are frequently overlooked or ignored in traditional survey methods [11,12], including both endangered and invasive species [8]. Third, eDNA collection is relatively simple, requiring only small volumes of seawater (< 3L) and simple filtering techniques, allowing sampling by individuals with limited training, even in remote locations (Miya et al., 2016). Forth, because eDNA doesn’t require diving, there are significant worker safety advantages. Lastly, eDNA is affordable (e.g. ~$50/sample) and has the potential for automation, allowing for remote sample collection and high throughput autonomous lab processing [14].

Despite these advantages, eDNA also has limitations. Of particular concern is PCR bias that can result in preferential amplification of particular taxa [15]. Additionally, detection probabilities can be influenced by species specific eDNA generation and degradation rates [8], an issue potentially further complicated by the transport of eDNA on ocean currents [16]. Furthermore, primer design, bioinformatic, and reference database limitations can also affect the accuracy of taxonomic assignment from eDNA [17].

Unlike well-established visual surveys, the impact of biases in eDNA metabarcoding are not well characterized, and may be less problematic than believed. For example recent studies show that impacts of PCR bias can be mitigated by technical replicates and site occupancy modelling [17–19]. Similarly, because eDNA signals decay relatively rapidly (e.g. hours to days; [20,21]), eDNA signatures are surprisingly stable [22]. As such, eDNA holds tremendous promise for monitoring marine ecosystems. Realizing that promise, however, requires a better understanding of how visual surveys and eDNA metabarcoding approaches compare in direct field applications.

The Channel Islands MPA Network spans >1000 reefs across 8 islands of the coast of Southern California. It is monitored by several programs including the Kelp Forest Monitoring Program, which conducts visual monitoring surveys of 41 invertebrates and over 100 fishes [23]. In total only 94 of the >1000 Channel Island reefs are surveyed, and just once per year [6], missing the seasonal dynamics in the variable Southern California Bight, limiting the scope and scale of assessment [2]. While born of logistical necessity, the spatial and temporal limits of this survey protocol makes accurately assessing the health of this MPA network difficult [2,24] and suggests the need for new approaches that produce data on broader taxonomic, spatial and temporal scales.

This study tests the efficacy of eDNA for MPA monitoring and to better understand the advantages and shortcomings of eDNA methods. We do this through a side-by-side comparisons of eDNA metabarcoding and visual surveys of fish communities conducted by the National Parks Service.

Materials and Methods

Sample collection

We conducted our study at Scorpion State Marine Reserve within the Channel Islands National Park and National Marine Sanctuary. To determine the degree to which eDNA could capture documented differences inside and outside this MPA, we sampled three sites: 1) inside the MPA (34.05223 N , 119.58253 W) 2) outside but adjacent (<0.5km) to the MPA (“edge site”; 34.04415 N, 119.54245 W), and 3) 2.3km outside the MPA boundary (“outside site”; 34.03837 N, 119.5253 W; Fig 1). At each of these three sites, we sampled directly along a 100 m fixed transect used by the Kelp Forest Monitoring Program for visual monitoring, using a GPS to ensure transects overlapped [23]. We collected three replicate 1L water samples from three locations on each transect, totaling 9 spatially structured replicates per site. Due to fieldwork logistical challenges, each site was sampled on a different day with a maximum of 72 hours between sampling events.

**Figure 1.** Map of Scorpion State Marine Reserve off Santa Cruz Island, CA, USA.

We collected seawater samples from 10 m below the surface and 1 m above the benthos using a 4 L Niskin bottle deployed from the UCLA RV Kodiak [25]. From each Niskin deployment, we transferred a single liter of seawater to an enteral feeding pouch and conducted gravity filtration through a sterile 0.22 µm Sterivex cartridge (MilliporeSigma, Burlington, MA, USA) in the field (Miya et al., 2016). Additionally, we processed three field blanks as a negative control that consisted of 1 L of distilled water following the method above. Finally, we dried Sterivex filters using a 3 mL syringe and then capped and stored the filters at -20˚C for DNA laboratory work back at UCLA (Miya et al., 2015).

**DNA extraction and library preparation**

We extracted eDNA from the Sterivex cartridge using the DNAeasy Tissue and Blood Kit (Qiagen Inc., Germantown, MD) following modifications of Spens et al. (2017). We PCR amplified the extracted eDNA using the MiFish Universal Teleost *12S* primer (Miya et al., 2015) with Nextera modifications following PCR and the library preparation methods of Curd et al. (2019) (See supplemental methods). All PCRs included a negative control where molecular grade water replaced the DNA extraction. For positive controls, we used DNA extractions of grass carp (*Ctenopharyngodon idella,* Cyprinidae*)* and Atlantic salmon (*Salmo salar,* Salmonidae), both non-native to California. Libraries were sequenced on a MiSeq PE 2x300bp at the Technology Center for Genomics & Bioinformatics (University of California- Los Angeles, CA, USA), using Reagent Kit V3 with 20% PhiX added to all sequencing runs.

**Bioinformatics**

To determine community composition, we used the *Anacapa Toolkit* (version: 1) to conduct quality control, amplicon sequence variant (ASV) parsing, and taxonomic assignment using user-generated custom reference databases [28]. The *Anacapa Toolkit* sequence QC and ASV parsing module relies on *cutadapt* (version: 1.16)[29], *FastX-toolkit* (version: 0.0.13) [30], and *DADA2* (version 1.6) [31] as dependencies and the *Anacapa classifier* modules relies on *Bowtie2* ( version 2.3.5)[32]and a modified version of *BLCA* [33] as dependencies. We processed sequences using the default parameters and assigned taxonomy using two *CRUX-generated* reference databases. We first assigned taxonomy using the FishCARD California fish specific reference database [34]. Second, we used the *CRUX*-generated *12S* reference database supplemented with FishCARD reference sequences to assign taxonomy using all available *12S* reference barcodes to identify any non-fish taxa. We note that *CRUX* relies on *ecoPCR* (version: 1.0.1)[35]*, blastn* (version: 2.6.0)[36]*,* and *Entrez-qiime* (version: 2.0)[37]asdependencies.

Raw ASV community table was decontaminated following Kelly et al. (2018) and McKnight et al. (2019) (See S1 Appendix). We chose a site occupancy cutoff score of 84% which corresponded with the minimum occupancy rate observed for 3 detections at a given location sampled.We then transformed all read counts into an eDNA index for beta-diversity statistics [15]. All non-fish species (mammals and birds) were removed prior to final analyses.

**eDNA data analysis**

To test for alpha diversity differences, total species richness for each site was then compared, using an Analysis of Variance (ANOVA) and subsequent Levine’s test for equality of variance [39].

To determine whether our eDNA sampling design was sufficient to fully capture fish community diversity, we created species rarefaction curves using the iNext package [40]. Species coverage estimates were then compared between each site, with and without site occupancy modeling, and using all three 1L replicates taken at three locations along a 100m transect (n=9) as well as only three 1L biological replicates (n=3). We ran a piecewise regression analysis to identify breakpoints in the rate of species diversity found per sample collected using the *R* packaged *segmented* [41].

To test for differences among fish communities, we calculated Bray-Curtis similarity distances on the eDNA index scores between all samples [22]. Specifically, we tested for the difference in community similarity variance between our three sites using an *adonis* PEMANOVA (*vegan* version: 2.4.2)[39], followed by a companion multivariate homogeneity of group dispersions test (BETADISPER) [39]. Both the PERMANOVA and BETADISPER were run using the following model: eDNA Index ~ Site + Location. We also visualized community beta diversity using non-metric multidimensional scaling (NMDS) [39]. To further investigate which species were driving eDNA community differences among sites, we conducted constrained analysis of principle components (CAP) [39].

**Visual underwater census methods**

To assess fish communities using underwater visual census techniques, SCUBA divers from the Kelp Forest Monitoring Program followed standard survey protocols following Kushner et al*.* (2013). These protocols include survey types: visual fish transects, roving diver fish counts, and 1m quadrats. The visual fish transects targeted 13 indicator species of fish on visual fish transects recording the counts of adults and juveniles. This protocol consists of performing 2m x 3m x 50 m transects along the 100 m permanent transect. During roving diver fish count surveys all positively identified species are recorded. This protocol consists of 3-6 divers counting all fish species observed during a 30 minute time period, covering as much of the 2000 m2 of bottom and entire water column as possible. The 1 m2 quadrat records three small demersal species of fish*.* All visual surveys occurred along a permanent 100 m transect at each site and were conducted within 2 weeks of eDNA sampling (See S1 Appendix).

**Comparison of eDNA and visual underwater census methods**

We compared species detected by eDNA and underwater visual census approaches across corresponding transects at each site. We identified core taxa that were shared across all sites for a given method. In addition, we identified species that eDNA methods failed to detect but were observed in visual census surveys and vice versa. We note that given the few numbers of sites (n=3) we were unable to robustly compare abundance estimates between methods.

Results

**eDNA results**

We generated over 4 million reads that passed quality control. The *Anacapa Toolkit* identified 2,906 ASVs from 3,091,063 reads representing 27 samples and 8 controls. After the second decontamination step, however, totals reduced to 441 ASVs and 2.23 million reads (Tables S1-S3).

Combined, eDNA metabarcoding successfully detected 42 fish taxa, representing 40 unique species, 39 genera, 28 families, and two classes (Tables S1-S3). eDNA detected 31 species within the MPA, 36 at the edge, and 38 species outside the MPA. The three sites shared a core group of 29 taxa including bony fish and one species of ray (Fig 2) (Table S4). Of these taxa, 18 species are associated with rocky reef habitat, five species are associated with sandy bottom habitat, four species are pelagic-neritic, and two species are pelagic-oceanic.

**Figure 2.** Venn Diagram of Fish Species Detected with eDNA

Species rarefaction curves showed that sampling at each site (n=9) was insufficient to capture all species diversity (Fig 3). Sample coverage estimates from eDNA results before filtering by site occupancy modeling filters were 94.4%, 88.9%, and 93.0% for the MPA, edge, and outside sites, respectively. Coverage estimates dropped to 80.2%, 80.0%, 82.0% for the MPA, edge, and outside sites, respectively, when only three 1L samples were used. Piecewise regression analysis showed a transition from exponential to linear increase in species detected per replicate between three and four replicate water samples per site (3.35-3.47) with subsequent diminishing sample coverage returns with the addition of more samples. In contrast, species diversity was near saturated (96.7%, 96.4%, and 99.8% for the MPA, edge, and outside sites, respectively) when applying a site occupancy rate above 84% and using three 1 L replicates taken at three locations along a 100 m transect. However, using only three samples, sample coverage dropped to 87.1%, 90.3%, 88.9% for the MPA, edge, and outside sites, respectively.

**Figure 3.** Species Rarefaction Curves. **a**) Species rarefaction curves for all fishes found at each site across 3 1L replicates taken at 3 locations along a 100m transect. **b**) Species rarefaction curves for fish species with occupancy rates above 84% found at each site across 3 1L replicates taken at 3 locations along a 100m transect. Sample coverage estimates were higher for species with occupancy rates above 84% (96.7-99.8%) than for all species (88.9%-94.4%). For species with occupancy rates above 84% sample coverage estimates ranged from 87.3-90.0% for only 3 1L replicates.

Analyses showed a significant difference in the total number of observed species across sites, with the site outside the MPA having significantly higher diversity than both the edge and MPA sites (ANOVA, p<0.001, Levine’s test p> 0.5). Observed species differences between sites were partially driven by the presence of non-rocky reef taxa (48.4%, 5/13), primarily pelagic, mobile, sandy bottom, and intertidal species. Moreover, there were also significant differences in fish communities among the three sites as well as among the three sampling locations along each of the three transects (PERMANOVA p<0.001, betadisper p>0.05). Location along the transect explained 27.5% of the total variance while site (e.g. inside, edge and outside MPA) explained 22.5% of the total variance; 50.0% of the total variance was unexplained.

NMDS ordination showed weak clustering of samples by both location and site (NMDS, Stress 0.21; Fig 4). Constrained analysis of principle components (CAP) found significant differences in species assemblages between samples collected at different sites and locations (CAP, p<0.001) (Fig 5), further indicating difference in eDNA signatures across sites and locations. CAP analysis identified seven taxa with the strongest differences between sites. The MPA site had higher eDNA index scores of opaleye (*Girella nigricans,* Kyphosidae*)* and kelp bass (*Paralabrax clathratus,* Serranidae)*.* The edge site had higher index scores of blacksmith (*Chromis punctipinnis,* Pomacentridae) and fantail flounder (*Xystreurys liolepis,* Paralychthyidae).The site outside the MPA had higher index scores of giant black sea bass (*Stereolepis gigas,* Polyprionidae), Pacific barracuda (*Sphyraena argentea*, Sphyraenidae), and topsmelt (*Atherinops affinis,* Atherinopsidae)*.*

**Figure 4.** NMDS of Bray-Curtis Dissimilarities. Bray-Curtis dissimilarities were calculated between all samples using only species with occupancy rates over 84%. Samples from Sites (colors) and locations (shapes) are similar to each other (NMDS, Stress = 0.21).

**Figure 5.** Constrained Analysis of Principle Components (CAP) Ordination. Bray-Curtis dissimilarities were calculated between all samples using only species with occupancy rates over 84%. Samples from sites and locations within sites were used as independent variables. Site and locations within sites are significantly more similar to each other (CAP, p<0.001). Sites (shapes) and Locations (colors) are plotted against CAP1 and CAP2 axes. Arrows correspond to direction and strength (length) of each species. Only the top 7 species with CAP distances greater than 0.40 were plotted.

**Visual census surveys results**

Across all three sites, 25 bony fish species were recorded using underwater visual censuses, representing 20 genera, 13 families, and one class (Fig 6) (Table S5), 11 of which were shared across all three sites (Table S6). Within the MPA site, visual census methods detected 21 unique species, 18 genera, and 11 families. At the edge site visual census methods detected 18 species, 16 genera, 11 families, and four classes. Lastly, at the outside site visual census methods detected 13 species, 13 genera, 10 families, and four classes. Of all taxa observed in visual census methods, 24 species were associated with rocky reef habitat and one species was pelagic-neritic. The pelagic-neritic species, top smelt (*Atherinops affinis,* Atherinopsidae), was only found in the MPA site.

On average, roving diver fish counts recorded 17.6 species per replicate survey (Range: 10-22). Visual fish counts recorded an average 7.8 species per replicate survey out of the 13 indicator species (Range: 5-10). 1m quadrats recorded an average 2.3 species of 3 target species (Range: 1-3).

**Figure 6**. Venn Diagram of Species Observed from Visual SCUBA Surveys.

**Comparison of eDNA and visual census surveys**

eDNA detected 76% (19 out of 25) of species observed during all combined National Park Service transect surveys (Table S5-S6). eDNA failed to resolve *Lythrypnus dalli* (Gobiidae), *L. zebra, Sebastes atrovirens* (Sebastidae)*, S. auriculatus, S. chrysomelas,* and *S. serranoides* to species level. At the genus level, eDNA performed markedly better recovering 95% (19 out of 20) of genera observed during under water censuses. The remaining genus *Lythrypnus* was detected prior to site occupancy modeling, but occurred in only one replicate at two separate sites.

In addition to the above, eDNA recovered 23 species that were not recorded during the visual censuses conducted by the National Park Service. Of these, all were California fish species previously recorded in Kelp Forest Monitoring Program surveys (Table S7), but not observed during our paired surveys.

There were few conspicuous differences in species observed across sites, with visual census results identifying 11 common taxa across all sites (Table S6). Of these, 10 were also found to be common across all sites using eDNA methods with one species (*Lythrypnus dalli*) not detected by eDNA. Species richness from visual census data showed that fish diversity was highest within MPA (n=21), lowest outside the MPA (n=13) and intermediate (n=18) on the edge of the MPA, while eDNA had the opposite pattern.

Discussion

Results demonstrate the power of for detecting a broad range of fish biodiversity in California kelp forest ecosystems, providing more detailed species inventories needed for marine ecosystem monitoring [9].eDNA was able to detect significant differences in fish communities inside, on the edge of, and outside of the Scorpion State Marine Reserve, even though the closest sites were no more than 500 m apart. Even within each of these sampling sites, eDNA distinguished among sample locations separated by only 50 m, highlighting the sensitivity of eDNA in capturing local fish communities, and matching previous studies showing fine-scale spatial resolution of eDNA signatures [12].

Importantly, eDNA captured 76% of fish diversity observed during visual surveys, despite species rarefaction indicating insufficient sampling. In total, eDNA only failed to identify six of 25 fish species observed during visual surveys, four of these being rockfish (*Sebastes,* Sebastidae), a taxon that *12S* barcoding cannot distinguish to species [34]. This small deficiency was offset by detecting an 23 additional fish taxa not recorded during paired Kelp Forest Monitoring Program visual monitoring, representing an important advantage of eDNA. Because sampling can be obtained easily and processed economically, eDNA could allow for more frequent monitoring, expanding the scope of MPA monitoring programs while providing greater personnel safety.

**The utility of eDNA for MPA monitoring**

Despite the limited sampling design and the inability of our *12S* barcode to distinguish species of rockfish and gobies, eDNA largely recovered the same taxa observed in visual census surveys. This strong concordance likely stems from high eDNA detection probabilities lasting only a few hours [20], such that eDNA captures marine communities that were recently present [21]. The similarity of eDNA and visual surveys is even more remarkable given that eDNA and visual surveys were taken two weeks apart, a result that strongly suggests that fish diversity captured by eDNA is truly representative of fish communities and their differences inside and outside the Scorpion State Marine Reserve [22].

In addition, eDNA recorded an additional 23 species not recorded from visual surveys, but have been previously reported in other Kelp Forest Monitoring Program surveys (Table S7). Importantly, these taxa included species of significant management concern such as the critically endangered (IUCN) giant black seabass (*Stereolepis gigas,* Polyprionidae) and important fished species like yellowtail amberjack (*Seriola lalandi,* Carangidae). Additionally, although we focused on teleost fishes, our eDNA data included elasmobranchs, marine mammals, and marine birds, taxa that play important roles in nearshore rocky reef ecosystems, but that can be difficult to survey and monitor [9]. The expanded taxonomic coverage and the ability to detect rare, or cryptic taxa is a significant advantage of eDNA over traditional visual surveys, expanding the scope of MPA monitoring by capturing entire communities rather than a selected subset of taxa.

Key to MPA monitoring is the ability to distinguish among communities inside and outside of the MPA. Not only did eDNA detect significant differences inside and outside the MPA, it could also differentiate among samples taken 50 m apart. This result adds to a growing literature that shows the fate and transport of eDNA in marine environments is relatively limited in space and time [20,42,43], and highlights the suitability of eDNA for comparing inside and outside of even relatively small MPAs [12].

While eDNA found significant differences inside and outside of the MPA and provided data on more taxa than visual survey methods, it did it for a fraction of the cost and effort. Roving fish diver counts, the most similar visual survey to eDNA monitoring methods, costs the Kelp Forest Monitoring Program ~$1,200 per site (Table S8). In contrast, the eDNA sampling design employed in this study including materials, labor, and transportation was ~$600 per site (Table S9)—and 25% of this total was just transportation. Moreover, total costs could have been significantly reduced by sampling in one day, which was not possible due to vessel logistics. Further cost efficiencies can come from automating lab methods and conducting sequencing in house [8].

In addition to the above, eDNA has other significant advantages. It can potentially detect invasive species, even when rare [8]. Sequence data from eDNA provides an annual snapshot of standing genetic diversity, providing the ability to monitor changes over time [9]. Similarly, in species with population structure, eDNA could provide evidence of range shifts associated with climate change [44]. Importantly, given eDNA metabarcoding samples can be preserved and archived, eDNA samples can be reanalyzed in the future with improved metabarcoding methods to answer additional hypotheses and environmental monitoring goals [8]. Combined, the above advantages of eDNA suggest that even if eDNA metabarcoding is not viewed as a full replacement for visual surveys, the power of this method, and it’s ease of sampling and affordability argue for using eDNA as a critically important complementary tool to greatly expand current monitoring activities.

**Limitations and caveats of eDNA**

Although this and other studies highlight the promise of eDNA for monitoring marine ecosystems, there are also important limitations. One key limitation is the lack of universal barcode loci. Four of the six undetected species in this study were rockfish in the genus *Sebastes.* While the MiFish *12S* metabarcoding primers have broad utility in vertebrates, rockfishesarea recent adaptive radiation [45] with a highly conserved 12S sequence, resulting in the inability to distinguish among rockfish ASVs. Identifyingrockfish tospecies using eDNA approaches is critical for MPA monitoring efforts in California as *Sebastes* are important for commercial and recreational fisheries [46] and play a wide array of functional and ecological roles in nearshore ecosystems [45].

In addition, eDNA failed to detect two gobies, *Lythrypnus dalli* and *L. zebra* (Gobiidae). Previous efforts to barcode *L. dalli* for the FishCARD reference database found two insertions not found in any other California goby, including the sister species *L. zebra* [34]. Thus, primer mismatch may have limited the amplification and detection of some *L. dalli* in our eDNA samples. Alternatively, the eDNA methods employed here may not be suited for small, crevice-dwelling fish species such as gobies. Species of *Lythrypnus* rarely leave the reef boundary layer [47]. As such their eDNA maybe entrained close to the reef, resulting in hyper-spatial variability of eDNA signatures [12]. More work is necessary to determine whether eDNA can reliably detect species living in interstitial reef habitat. This limitation, however, is not unique to eDNA as the Kelp Forest Monitoring Program employs 1 m2 quadrat surveys, specifically designed to capture these taxa. Likewise eDNA surveys that specifically sample within the boundary layer may be needed to survey benthic cryptic species.

Another limitation of eDNA is standardizing processing techniques, including the spatial design of field sampling, number of replicates, and sequencing depth [8,15,19]. The three replicate water samples taken from a single location and time recovered only 88.3% of the species present based on modeled species coverage estimates of species with at least 84% occupancy. This value increased to near saturation (>97.6%) by sampling three replicate water samples from three locations along a 100 m transect. That said, rarefaction curves indicated that additional sampling would have recovered additional species. These results provide important benchmarks for replication and sampling efficiency within nearshore marine environments and highlight the need to adjust sampling intensity and replicates, depending on the questions to be addressed with eDNA.

Despite not achieving saturation with our sampling design, we did observe a transition from exponential to linear addition of species detections with additional sampling similar to that previously demonstrated in mesocosm experiments [19]. This shift likely reflects the biological reality of eDNA within marine ecosystems, with a few taxa being abundant and a long tail of low abundant species [15]. As such, while only a few replicates are needed to capture local core species diversity, high technical (PCR) and biological (bottle) replication may be required to saturate species detection [19]. Thus, if the goal is to detect rare species, it is imperative to increase sample number, an unsurprising result given the reality of detection probabilities of rare taxa [18]. Despite this caveat and our relatively limited number of sample replicates, we still detected rare species such as giant black seabass (*Stereolepis gigas,* Polyprionidae) suggesting that eDNA is likely still superior to visual techniques at rare species detection [10].

**Importance of site occupancy modelling**

Site occupancy modeling showed that almost all species with occupancy rates higher than 84% were common Southern California kelp forest species [48]. In contrast, almost all pelagic and intertidal species that should not be present in a kelp forest had low occupancy rates and were detected only in a single bottle replicate (Tables S1-S2). These low occupancy detections cannot be contamination because they did not occur in field or laboratory controls; instead, they likely represent eDNA transported between habitats [16]. Regardless, site occupancy modeling removed the vast majority of unexpected kelp forest fishes, highlighting its value for determining true species detections in a rigorous and repeatable way [18,19], aiding in the interpretation and comparison of eDNA results.

While site occupancy modelling removed non-kelp forest taxa (e.g. Blue whale; *Balaenoptera musculus* (Cetacea); California sea lion, *Zalophus californianus* (Otariidae); pelagic cormorant *Urile pelagicus* (Phalacrocoracidae); Table S10), it also removed some kelp forest species (e.g. zebra goby, *L. dalli,* Gobiidae; swell shark, *Cephaloscyllium ventriosum,* Scyliorhinidae; zebra-perch *Hermosilla azurea,* Embiotocidae; California angel shark, *Squatina californica,* Squatinidae). These results highlight the need for increased replication depending on the management question, just as it may require more visual surveys to observe numerically rarer taxa, such as sharks. Although the ability of eDNA to detect marine mammals and birds is useful, visual observations maybe more effective depending on the taxa, suggesting that complementary methods may yield the most effective sampling regime [49].

**Diversity inside and outside MPAs**

Traditional visual surveys most often report higher biodiversity and biomass inside MPAs [50], including Scorpion State Marine Reserve [6]. However, our results surprisingly indicate lower diversity inside the MPA. This paradoxical result is partially explained by the inability of eDNA to resolve *Sebastes* species that were visually observed inside (n=3) and on the edge of the MPA (n=1), but not outside. Furthermore, species rarefaction curves also suggest that all samples had sufficient read depth to saturate species richness following site occupancy modelling (Figure S1).

Instead, species richness differences observed here must be due to another unaccounted for factor. First, the observed pattern of species richness from eDNA may be the true pattern of species richness and that SCUBA surveys did not capture the true pattern of species richness. Second, both SCUBA and eDNA methods may have under-sampled diversity across these three reefs requiring higher sampling effort to capture the true species richness patterns. These two explanations highlight the difficulty of comparing two surveying methods with imperfect detection rates [49,51]. For visual transect surveys this could be overcome with additional survey types that focus on identifying more species as well as surveying a greater number of transects within each site [52,53]. For eDNA surveys this could require an increased number of replicate locations within a site or replicate water bottles at each location to avoid potential patchiness of eDNA dispersion across a reef [8]. In addition, a greater number of technical PCR replicates may be required to reach full saturation as has been observed in controlled mesocosm experiments[19].

Lastly, another potential explanation for this unexpected result is that low density of kelp forest fishes outside the MPA increased the detection of non-kelp forest taxa advected from elsewhere. In total, 38.5% of taxa detected outside the MPA were non-rocky reef species such as Yellowtail amberjack (*Seriola lalandi,* Carangidae), California clingfish (*Gobiesox rhessodon,* Gobiesocidae), and Fantail flounder (*Xystreurys liolepis,* Paralichthyidae)*.* Although these species occasionally pass through rocky reef environments, another potential explanation is that eDNA from these species were transported from nearby pelagic, intertidal, and sandy bottom ecosystems [16,43] or perhaps that they rapidly swim through these areas bud do not reside such as is likely with Seriola lalandri. While such transport would be expected at all sites, high fish abundance inside the MPA could result in a strongly skewed ratio of endogenous kelp forest eDNA to exogenous pelagic eDNA [16], with the signal of kelp forest taxa dominating that of pelagic species.

This paradoxical pattern of species richness highlights that eDNA data must be interpreted with caution [54]. Metabarcoding methods often perform unexpectedly when DNA concentrations are low, increasing the probability of sequencing rare species [8]. Thus additional ecological metrics to species richness, ones that are more representative of ecological patterns and processes, are needed to optimally interpret eDNA results [15]. These results ultimately highlight the value of ground truthing eDNA results with visual surveys in novel applications to ensure proper interpretation of results [49]. However, they also highlight that comparing survey methods with imperfect detection efforts in the field is challenging as there is no “gold standard” for ecosystem biomonitoring [49,51].

**Conclusion**

Marine protected areas are indispensable tools for protecting marine ecosystems and effective monitoring is paramount to their success [1]. Our results demonstrate that eDNA can distinguish fish assemblages inside and outside MPAs, and can detect other vertebrates, like marine mammals and birds, of special conservation concern.

Given its power, ease of sampling and relative affordability, eDNA could provide critical cost-added-benefits of repeated temporal or expanded spatial sampling of marine protected areas. In particular, eDNA metabarcoding can overcome many of the current limitations of visual monitoring, increasing sampling frequency and expanding monitoring beyond a small subset of “important” focal taxa. Such expanded monitoring would improve our ability to understand the ecological processes, human impacts, and management strategies that affect marine community communities that MPAs are designed to protect.

However, important aspects of eDNA remain unresolved, notably determining abundance and biomass via eDNA [55]. Given the mixed results on the efficacy of biomass estimates from eDNA in marine systems [8,55], eDNA will not be viewed yet as a wholesale replacement for visual monitoring, but instead as a powerful complementary tool. There will always be value in the direct observation by divers in particular for biomass estimates, informing size class distributions and sex ratios [24], but eDNA provides an important way to simultaneously make surveys more comprehensive and efficient. By replacing aspects of underwater visual surveys, eDNA could reduce the dive time per site, allowing more sites to be surveyed more frequently or improve overall biodiversity estimates. Additionally, whereas it takes 3 months for the Kelp Forest Monitoring Program to complete diver-based surveys, field collection of eDNA could be completed in a week, allowing for surveys to occur during short periods of good weather in the winter when full visual surveys would be impossible. As such, eDNA could greatly expand current monitoring activities across space, time, and depth, providing resource managers critical information on the response of MPAs to changing environments and management practices, contributing greatly to marine ecosystem understanding.

**Acknowledgments**

We thank the Channel Islands National Park Service and Kelp Forest Monitoring Program divers for conducting underwater visual censuses. We gratefully acknowledge Tiara Moore, Nick Schooler, Rachel Turba, and Lucia Bertero for their tireless help in the field collection of eDNA samples. We also acknowledge Taylor Ely and Onny Marwayana for assistance with laboratory work.

Data Availability

We intend to archive all data on the Dryad Digital Repository and a link will be provided upon acceptance.

References

1. Edgar GJ, Stuart-Smith RD, Willis TJ, Kininmonth S, Baker SC, Banks S, et al. Global conservation outcomes depend on marine protected areas with five key features. Nature. 2014;506. doi:10.1038/nature13022

2. Nickols KJ, White JW, Malone D, Carr MH, Starr RM, Baskett ML, et al. Setting ecological expectations for adaptive management of marine protected areas. J Appl Ecol. 2019;56: 2376–2385. doi:https://doi.org/10.1111/1365-2664.13463

3. Usseglio P. Ecology of Fishes on Coral Reefs. Cambridge University Press; 2015. doi:https://doi.org/10.1017/CBO9781316105412

4. Sprague J. Personal Observation Joshua Sprague. 2020.

5. Lindfield SJ, Harvey ES, McIlwain JL, Halford AR. Silent fish surveys: bubble-free diving highlights inaccuracies associated with SCUBA-based surveys in heavily fished areas. Börger L, editor. Methods Ecol Evol. 2014;5: 1061–1069. doi:10.1111/2041-210X.12262

6. Pondella II DJ, Caselle JE, Claisse JT, Williams JP, Davis K, Williams CM, et al. Baseline Characterization of the Shallow Rocky Reef and Kelp Forest Ecosystems of the South Coast Study Region. 2015. Available: https://caseagrant.ucsd.edu/news/summaries-of-projects-selected-for-funding-through-the-south-coast-mpa-baseline-program

7. Bernard ATF, Götz A, Kerwath SE, Wilke CG. Observer bias and detection probability in underwater visual census of fish assemblages measured with independent double-observers. J Exp Mar Bio Ecol. 2013;443: 75–84. doi:10.1016/j.jembe.2013.02.039

8. Deiner K, Bik HM, Mächler E, Seymour M, Lacoursière-Roussel A, Altermatt F, et al. Environmental DNA metabarcoding: Transforming how we survey animal and plant communities. Mol Ecol. 2017;26: 5872–5895. doi:10.1111/mec.14350

9. Kelly RP, Port J a., Yamahara KM, Martone RG, Lowell N, Thomsen PF, et al. Harnessing DNA to improve environmental management. Science (80- ). 2014;344. doi:10.1126/science.1251156

10. Thomsen PF, Kielgast J, Iversen LL, Møller PR, Rasmussen M, Willerslev E. Detection of a diverse marine fish fauna using environmental DNA from seawater samples. PLoS One. 2012;7: e41732. doi:10.1371/journal.pone.0041732

11. Willis TJ. Visual census methods underestimate density and diversity of cryptic reef fishes. J Fish Biol. 2001;59: 1408–1411. doi:https://doi.org/10.1111/j.1095-8649.2001.tb00202.x

12. Port JA, O’Donnell JL, Romero‐Maraccini OC, Leary PR, Litvin SY, Nickols KJ, et al. Assessing vertebrate biodiversity in a kelp forest ecosystem using environmental DNA. Mol Ecol. 2015. doi:https://doi.org/10.1111/mec.13481

13. Miya M, Minamoto T, Yamanaka H, Oka S, Sato K, Yamamoto S, et al. Use of a Filter Cartridge for Filtration of Water Samples and Extraction of Environmental DNA. J Vis Exp. 2016; e54741–e54741. doi:10.3791/54741

14. Yamahara KM, Preston CM, Birch JM, Walz KR, Marin III R, Jensen S, et al. In-situ Autonomous Acquisition and Preservation of Marine Environmental DNA Using an Autonomous Underwater Vehicle. Front Mar Sci. 2019;6: 373. doi:https://doi.org/10.3389/fmars.2019.00373

15. Kelly RP, Shelton AO, Gallego R. Understanding PCR processes to Draw Meaningful conclusions from environmental DNA Studies. Sci Rep. 2019;9: 1–14. doi:https://doi.org/10.1038/s41598-019-48546-x

16. Andruszkiewicz EA, Koseff JR, Fringer OB, Ouellette NT, Lowe AB, Edwards CA, et al. Modeling environmental DNA transport in the coastal ocean using Lagrangian particle tracking. Front Mar Sci. 2019;6: 477. doi:10.3389/fmars.2019.00477

17. Alberdi A, Aizpurua O, Gilbert MTP, Bohmann K. Scrutinizing key steps for reliable metabarcoding of environmental samples. Methods Ecol Evol. 2018;9: 134–147. doi:doi:10.1111/2041-210X.12849

18. Chambert T, Pilliod DS, Goldberg CS, Doi H, Takahara T. An analytical framework for estimating aquatic species density from environmental DNA. Ecol Evol. 2018;8: 3468–3477. doi:10.1002/ece3.3764

19. Doi H, Fukaya K, Oka S, Sato K, Kondoh M, Miya M. Evaluation of detection probabilities at the water-filtering and initial PCR steps in environmental DNA metabarcoding using a multispecies site occupancy model. Sci Rep. 2019;9: 3581. doi:https://doi.org/10.1038/s41598-019-40233-1

20. Murakami H, Yoon S, Kasai A, Minamoto T, Yamamoto S, Sakata MK, et al. Dispersion and degradation of environmental DNA from caged fish in a marine environment. Fish Sci. 2019;85: 327–337. doi:https://doi.org/10.1007/s12562-019-01341-z

21. Collins RA, Wangensteen OS, O’Gorman EJ, Mariani S, Sims DW, Genner MJ. Persistence of environmental DNA in marine systems. Commun Biol. 2018;1: 1–11.

22. Kelly RP, Gallego R, Jacobs-Palmer E. The effect of tides on nearshore environmental DNA. PeerJ. 2018;6: e4521. doi:https://doi.org/10.7717/peerj.4521

23. Kushner DJ, Rassweiler A, McLaughlin JP, Lafferty KD. A multi-decade time series of kelp forest community structure at the California Channel Islands. Ecology. 2013;94: 2655. doi:https://doi.org/10.1890/13-0562R.1

24. Caselle JE, Rassweiler A, Hamilton SL, Warner RR. Recovery trajectories of kelp forest animals are rapid yet spatially variable across a network of temperate marine protected areas. Sci Rep. 2015;5: 14102. doi:10.1038/srep14102

25. Thomsen PF, Møller PR, Sigsgaard EE, Knudsen SW, Jørgensen OA, Willerslev E. Environmental DNA from seawater samples correlate with trawl catches of subarctic, deepwater fishes. PLoS One. 2016;11. doi:10.1371/journal.pone.0165252

26. Miya M, Sato Y, Fukunaga T, Sado T, Poulsen JY, Sato K, 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. 2015;2: 150088. doi:10.1098/rsos.150088

27. Spens J, Evans AR, Halfmaerten D, Knudsen SW, Sengupta ME, Mak SST, et al. Comparison of capture and storage methods for aqueous macrobial eDNA using an optimized extraction protocol: advantage of enclosed filter. Yu D, editor. Methods Ecol Evol. 2017;8: 635–645. doi:10.1111/2041-210X.12683

28. Curd EE, Gold Z, Kandlikar GS, Gomer J, Ogden M, O’Connell T, et al. Anacapa: an environmental DNA toolkit for processing multilocus metabarcode datasets. Methods Ecol Evol. 2019;10: 1469– 1475. doi:https://doi.org/10.1111/2041-210X.13214

29. Martin M. Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet J. 2011;17: 10–12. doi:10.14806/ej.17.1.200

30. Gordon A, Hannon GJ. Fastx-toolkit. FASTQ/A short-reads preprocessing tools (unpublished) http://hannonlab cshl edu/fastx\_toolkit. 2010;5.

31. Callahan BJ, McMurdie PJ, Rosen MJ, Han AW, Johnson AJA, Holmes SP. DADA2: high-resolution sample inference from Illumina amplicon data. Nat Methods. 2016;13: 581.

32. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. Nat Methods. 2012;9: 357.

33. Gao X, Lin H, Revanna K, Dong Q. A Bayesian taxonomic classification method for 16S rRNA gene sequences with improved species-level accuracy. BMC Bioinformatics. 2017;18: 247.

34. Gold ZJ. Design and Implementation of Environmental DNA Metabarcoding Methods for Monitoring the Southern California Marine Protected Area Network. UCLA. 2020. doi:ark:/13030/m5j44187

35. Ficetola GF, Coissac E, Zundel S, Riaz T, Shehzad W, Bessière J, et al. An in silico approach for the evaluation of DNA barcodes. BMC Genomics. 2010;11: 434.

36. Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, et al. BLAST+: architecture and applications. BMC Bioinformatics. 2009;10: 421.

37. Baker CCM. entrez qiime: a utility for generating QIIME input files from the NCBI databases. 2016.

38. McKnight DT, Huerlimann R, Bower DS, Schwarzkopf L, Alford RA, Zenger KR. microDecon: A highly accurate read‐subtraction tool for the post‐sequencing removal of contamination in metabarcoding studies. Environ DNA. 2019. doi:https://doi.org/10.1002/edn3.11

39. Oksanen J, Blanchet FG, Kindt R, Legendre P, Minchin PR, O’hara RB, et al. Package ‘vegan.’ Community Ecol Packag version. 2013;2.

40. Hsieh TC, Ma KH, Chao A. iNEXT: an R package for rarefaction and extrapolation of species diversity (Hill numbers). Methods Ecol Evol. 2016;7: 1451–1456. doi:https://doi.org/10.1111/2041-210X.12613

41. Muggeo VMR. Segmented: an R package to fit regression models with broken-line relationships. R news. 2008;8: 20–25. doi:https://doi.org/10.1002/sim.1545

42. Yamamoto S, Minami K, Fukaya K, Takahashi K, Sawada H, Murakami H, et al. Environmental DNA as a “snapshot” of fish distribution: A case study of Japanese jack mackerel in Maizuru Bay, Sea of Japan. PLoS One. 2016;11. doi:10.1371/journal.pone.0149786

43. O’Donnell JL, Kelly RP, Shelton AO, Samhouri JF, Lowell NC, Williams GD. Spatial distribution of environmental DNA in a nearshore marine habitat. PeerJ. 2017;5: e3044. doi:10.7717/peerj.3044

44. Sanford E, Sones JL, García-Reyes M, Goddard JHRR, Largier JL. Widespread shifts in the coastal biota of northern California during the 2014–2016 marine heatwaves. Sci Rep. 2019;9: 4216. doi:10.1038/s41598-019-40784-3

45. Hyde JR, Vetter RD. The origin, evolution, and diversification of rockfishes of the genus Sebastes (Cuvier). Mol Phylogenet Evol. 2007;44: 790–811. doi:10.1016/j.ympev.2006.12.026

46. Williams GD, Levin PS, Palsson WA. Rockfish in Puget Sound: An ecological history of exploitation. Mar Policy. 2010;34: 1010–1020. doi:https://doi.org/10.1016/j.marpol.2010.02.008

47. Behrents KC. The comparative ecology and interactions between two sympatric gobies (Lythrypnus dalli and Lythrypnus zebra). University of Southern California; 1984.

48. Love MS, Passarelli JK. Miller and Lea’s Guide to the Coastal Marine Fishes of California. 2nd. University of California Agriculture and Natural Resources; 2020.

49. Kelly RP, Closek CJ, O’Donnell JL, Kralj JE, Shelton AO, Samhouri JF. Genetic and manual survey methods yield different and complementary views of an ecosystem. Front Mar Sci. 2017;3: 283. doi:https://doi.org/10.3389/fmars.2016.00283

50. Lester SE, Halpern BS. Biological responses in marine no-take reserves versus partially protected areas. Mar Ecol Prog Ser. 2008;367: 49–56. doi:https://doi.org/10.3354/meps07599

51. Royle JA, Link WA. Generalized site occupancy models allowing for false positive and false negative errors. Ecology. 2006;87: 835–841.

52. Gray AE, Williams ID, Stamoulis KA, Boland RC, Lino KC, Hauk BB, et al. Comparison of Reef Fish Survey Data Gathered by Open and Closed Circuit SCUBA Divers Reveals Differences in Areas With Higher Fishing Pressure. Fulton CJ, editor. PLoS One. 2016;11: e0167724. doi:10.1371/journal.pone.0167724

53. Edgar GJ, Barrett NS, Morton AJ. Biases associated with the use of underwater visual census techniques to quantify the density and size-structure of fish populations. J Exp Mar Bio Ecol. 2004;308: 269–290. doi:10.1016/j.jembe.2004.03.004

54. Kelly RP, O’Donnell JL, Lowell NC, Shelton AO, Samhouri JF, Hennessey SM, et al. Genetic signatures of ecological diversity along an urbanization gradient. PeerJ. 2016;4: e2444. doi:https://doi.org/10.7717/peerj.2444

55. Shelton AO, Kelly RP, O’Donnell JL, Park L, Schwenke P, Greene C, et al. Environmental DNA provides quantitative estimates of a threatened salmon species. Biol Conserv. 2019;237: 383–391. doi:10.1016/j.biocon.2019.07.003

**Supplemental Tables**

Table S1: Taxonomy Table After Decontamination and Site Occupancy Modeling Using California Fish Reference Barcodes

Table S2: Taxonomy Table After Decontamination Using California Fish Reference Barcodes

Table S3: Metadata for eDNA Samples

Table S4: Summarized Species Table with Site Level Detections

Table S5: Species Detected with eDNA Metabarcoding and National Park Service Kelp Forest Monitoring Program Underwater Visual Surveys

Table S6: Underwater Visual Census Species Detections Across All Sites

Table S7. List of All Species Observed by the Channel Islands National Park Service Kelp Forest Monitoring Program

Table S8 Taxonomy Table After Decontamination Using All Vertebrate References

Table S9 Taxonomy Table After Decontamination and Site Occupancy Modeling Using All Vertebrate References

Table S10 Budget Estimate of eDNA Metabarcode Sampling

Table S11 Budget Estimate of eDNA Metabarcode Sampling

**Supplemental Figures**

FigS1: Species Richness Sequence Depth Rarefaction