11-1-2021

DNA metabarcoding marker choice skews perception of marine eukaryotic biodiversity

Jordan M. Casey
The University of Texas at Austin

Emma Ransome
Smithsonian National Museum of Natural History

Allen G. Collins
Smithsonian National Museum of Natural History

Angka Mahardini
Korea Institute Of Ocean Science & Technology

Eka M. Kurniasih
Yayasan Biodiversitas Indonesia (BIONESIA)

See next page for additional authors

Recommended Citation

Jordan M. Casey, Emma Ransome, Allen G. Collins, Angka Mahardini, Eka M. Kurniasih, Andrianus Sembiring, Nina M.D. Schiettekatte, Ni Kadek Dita Cahyani, Aji Wahyu Anggoro, Mikaela Moore, Abby Uehling, Mahdi Belcaid, Paul H. Barber, Jonathan B. Geller, and Christopher P. Meyer. "DNA metabarcoding marker choice skews perception of marine eukaryotic biodiversity" *Environmental DNA* (2021): 1229-1246. https://doi.org/10.1002/edn3.245

Received: 5 April 2021 | Revised: 30 June 2021 | Accepted: 27 July 2021

DOI: 10.1002/edn3.245

Environmental DNA
Open Access
Dedicated to the study and use of environmental DNA for basic and applied sciences

WILEY

ORIGINAL ARTICLE

DNA metabarcoding marker choice skews perception of marine eukaryotic biodiversity

Jordan M. Casey[1,2] | Emma Ransome[2,3] | Allen G. Collins[2,4] | Angka Mahardini[5,6] | Eka M. Kurniasih[7] | Andrianus Sembiring[7] | Nina M. D. Schiettekatte[8,9] | Ni Kadek Dita Cahyani[6,7,10] | Aji Wahyu Anggoro[7,10] | Mikaela Moore[2,11] | Abby Uehling[2] | Mahdi Belcaid[12] | Paul H. Barber[10] | Jonathan B. Geller[13] | Christopher P. Meyer[2]

[1]Department of Marine Science, Marine Science Institute, University of Texas at Austin, Port Aransas, Texas, USA

[2]Department of Invertebrate Zoology, National Museum of Natural History, Smithsonian Institution, Washington, District of Columbia, USA

[3]Department of Life Sciences, Imperial College, London, UK

[4]National Systematics Laboratory, NOAA Fisheries, Washington, District of Columbia, USA

[5]Jeju Marine Research Center, Korea Institute of Ocean Science and Technology, Jeju, Republic of Korea

[6]Indonesian Biodiversity Research Center, Denpasar, Bali, Indonesia

[7]Yayasan Biodiversitas Indonesia, Denpasar, Bali, Indonesia

[8]PSL Université Paris: EPHE-UPVD-CNRS, USR 3278 CRIOBE, Université de Perpignan, Perpignan, France

[9]Laboratoire d'Excellence "CORAIL", Perpignan, France

[10]Department of Ecology and Evolutionary Biology, University of California – Los Angeles, Los Angeles, California, USA

[11]Department of Geological Sciences and Geological Engineering, Queen's University, Kingston, ON, Canada

[12]Department of Information and Computer Sciences, University of Hawaii at Mānoa, Honolulu, Hawaii, USA

[13]Moss Landing Marine Laboratories, San José State University, Moss Landing, California, USA

Correspondence
Jordan M. Casey, Department of Marine Science, Marine Science Institute, University of Texas at Austin, Port Aransas, TX, USA.
Email: jcasey508@gmail.com

Funding information
National Science Foundation, Grant/Award Number: OISE 1243541; United States Agency for International Development, Grant/Award Number: 497-A-00-10-00008-00

Abstract

DNA metabarcoding is an increasingly popular technique to investigate biodiversity; however, many methodological unknowns remain, especially concerning the biases resulting from marker choice. Regions of the cytochrome *c* oxidase subunit I (COI) and 18S rDNA (18S) genes are commonly employed "universal" markers for eukaryotes, but the extent of taxonomic biases introduced by these markers and how such biases may impact metabarcoding performance is not well quantified. Here, focusing on macroeukaryotes, we use standardized sampling from autonomous reef monitoring structures (ARMS) deployed in the world's most biodiverse marine ecosystem, the Coral Triangle, to compare the performance of COI and 18S markers. We then compared metabarcoding data to image-based annotations of ARMS plates. Although both markers provided similar estimates of taxonomic richness and total sequence reads, marker choice skewed estimates of eukaryotic diversity. The COI marker recovered relative abundances of the dominant sessile phyla consistent with image annotations. Both COI and the image annotations provided higher relative abundance estimates of

Bryozoa and Porifera and lower estimates of Chordata as compared to 18S, but 18S recovered 25% more phyla than COI. Thus, while COI more reliably reflects the occurrence of dominant sessile phyla, 18S provides a more holistic representation of overall taxonomic diversity. Ideal marker choice is, therefore, contingent on study system and research question, especially in relation to desired taxonomic resolution, and a multi-marker approach provides the greatest application across a broad range of research objectives. As metabarcoding becomes an essential tool to monitor biodiversity in our changing world, it is critical to evaluate biases associated with marker choice.

KEYWORDS
18S rDNA, coral reefs, cryptic diversity, cytochrome *c* oxidase subunit I, high-throughput sequencing, taxonomy

1 | INTRODUCTION

Biodiversity loss is one of the greatest threats to ecosystems worldwide (Butchart et al., 2010; McCauley et al., 2015), with species loss inextricably linked to a reduction of ecosystem services, such as storm protection and food provisioning for humankind (Díaz et al., 2006; Worm et al., 2006). To inform ecosystem management, the accurate quantification of biodiversity is of utmost importance in today's changing world. Traditional taxonomic methods can be tedious and require high levels of expertise, especially for small, cryptic organisms. A powerful alternative is DNA metabarcoding, which consists of the bulk extraction of DNA from environmental samples, followed by the mass amplification and identification of a multitude of taxa based on DNA sequences of universal markers (Baird & Hajibabaei, 2012; Hajibabaei et al., 2011; Taberlet et al., 2012). This molecular technique has been widely and successfully employed to quantify biodiversity (Elbrecht et al., 2017; Ficetola et al., 2008; Gibson et al., 2014; Ji et al., 2013; Leray & Knowlton, 2015). Applications are diverse and include microbiome studies (Pollock et al., 2018), food web reconstruction (Casey et al., 2019), and water-based environmental DNA (eDNA; DiBattista et al., 2020).

However, there are many known biases associated with DNA metabarcoding. In an effort to minimize sampling biases when estimating cryptic biodiversity with DNA metabarcoding on coral reefs, researchers have developed autonomous reef monitoring structures (ARMS; oceanarms.org), standardized collection devices that are designed to mimic the structural complexity of coral reefs (Zimmerman & Martin, 2004). ARMS consist of nine stacked 225 × 225 mm PVC plates, which are separated by spacers and mounted on a base plate. Alternating layers between the plates contain crossed PVC bars, which create crevice-like formations that are ideal for the colonization of benthic organisms and motile meiofauna (Menge et al., 1983). ARMS have been deployed across many marine environments, and they are especially useful in hyperdiverse systems, such as coral reefs, where estimating cryptic biodiversity is challenging without the use of standardized collection structures (Al-Rshaidat et al., 2016; Leray & Knowlton, 2015; Pearman et al., 2018, 2019; Plaisance et al., 2011).

However, even with such standardized sampling methods, DNA metabarcoding has methodological biases (Alberdi et al., 2018). Sources of biases include sampling methods (Elbrecht et al., 2017; Ransome et al., 2017), DNA extraction method (Deiner et al., 2015), PCR amplification protocol and primer choice (Clarke et al., 2014, 2017; Cowart et al., 2015; Piñol et al., 2019; Tragin et al., 2018; Zhan et al., 2014), library preparation techniques (Braukmann et al., 2019; Zizka et al., 2019), sequencing error (Deagle et al., 2019; Elbrecht & Leese, 2015), bioinformatics biases (Brannock & Halanych, 2015; Brown et al., 2015; Flynn et al., 2015), and the unavailability of an adequate reference database (Cristescu, 2014; DiBattista et al., 2020; Gold et al., 2020; Ransome et al., 2017). One the greatest sources of bias in metabarcoding studies is associated with marker choice. Notably, a marker may fail to identify an organism or group of organisms due to issues with primer universality or reference database gaps. In studies focused on eukaryotic biodiversity (e.g., ARMS sampling), the mitochondrial cytochrome *c* oxidase subunit I (COI) and small subunit (SSU) 18S ribosomal RNA (rRNA) gene regions are two common targets that are considered "universal markers" that capture a broad range of eukaryotic taxa (Hebert et al., 2003; Leray & Knowlton, 2016; Tragin et al., 2018). Of these, 18S is the more commonly selected universal marker (59% and 28% of studies target the 18S and COI gene, respectively; van der Loos & Nijland, 2020). While both are considered universal markers that perform well across many eukaryotic taxa, they have known limitations that impact their ability to faithfully detect taxonomic diversity.

The COI marker is often selected due to the purported availability of well-curated reference databases (Hebert et al., 2003) and higher taxonomic precision; yet, it is ineffective across certain taxa. For example, while the high substitution rate in COI can provide excellent resolution in terms of taxonomic assignments (Andújar et al., 2018; Leray & Knowlton, 2016), species level discrimination is poor for some early diverging metazoan groups, such as Porifera and Anthozoa (Hebert et al., 2003; Huang et al., 2008; Shearer et al., 2002). The comparatively fast mutation rate also makes it difficult

to design primers that are truly universal across eukaryotic taxonomic groups (Deagle et al., 2014; Geller et al., 2013). For instance, many "universal" COI primers do not amplify well for Calcarea, Nematoda, or Platyhelminthes (Andújar et al., 2018; Bhadury et al., 2006; Prosser et al., 2013; Voigt & Wörheide, 2016), although recent primer development has shown slightly more success with these groups (Wangensteen et al., 2018). Another issue is that elevated rates of mutation cause saturation at high taxonomic levels, inhibiting accurate taxonomic assignments when closely related sequences at low taxonomic levels are unavailable (Deagle et al., 2014; Leray & Knowlton, 2016; Ransome et al., 2017).

Conversely, due to the relatively conserved nature of 18S, it provides noteworthy taxonomic coverage across a broad range of eukaryotes (Creer et al., 2010; Fonseca et al., 2010; Hadziavdic et al., 2014; López-García et al., 2001; Moon-van der Staay et al., 2001; Zhan et al., 2013), but this comes at the cost of inferior taxonomic resolution compared to COI (Leray & Knowlton, 2016; Wangensteen et al., 2018). While broad-range biodiversity assessments (currently dominated by plankton studies) commonly target the highly diverse V4 and V9 regions of 18S to maximize coverage across a wide assortment of eukaryotic groups (Casas et al., 2017; DiBattista et al., 2020; Pearman et al., 2014; Pearman & Irigoien, 2015; Tragin et al., 2018; de Vargas et al., 2015), the 18S gene region considerably underestimates eukaryotic species, especially cryptic microbial metazoans that are unaccounted for in DNA reference databases (Piganeau et al., 2011; Tang et al., 2012; Wangensteen et al., 2018; Wu et al., 2015). Furthermore, the 18S rRNA gene does not amplify well for some clades of mollusks, such as Solenogastres, due to elevated substitution rates (Meyer et al., 2010).

Although the limitations and biases inherent in the COI and 18S markers are foundational to the interpretation of metabarcoding data, the only known performance comparisons of these markers were conducted in the temperate zone (Clarke et al., 2017; Wangensteen et al., 2018; Zhan et al., 2014). No study has systematically quantified the performance of these two markers across the wide array of eukaryotic taxa present in hyperdiverse marine ecosystems, such as the Coral Triangle. Furthermore, no current study has compared marker performance to the visual annotation of standardized coral reef images, which remains the leading method to quantify sessile marine organisms (Beijbom et al., 2015; David et al., 2019; Williams et al., 2019).

In this study, we compare phylum-level biodiversity recovered through COI *versus* 18S metabarcoding (assessed via the performance of one primer set for each marker) based on ARMS deployed on a coral reef off the coast of Bali, Indonesia. Then, we compare marker performance to the visual annotation of sessile organisms from ARMS plates. In addition, we examine the relative impact of small-scale temporal and spatial variation on the recovery of marine eukaryotic diversity through metabarcoding. Combined, these comparisons provide a baseline of methodological biases and how they may impact our perception of marine eukaryotic biodiversity.

2 | MATERIALS AND METHODS

2.1 | ARMS deployment

We deployed all ARMS at a depth of approximately 10 m on the forereef of Close Encounters Reef near Pemuteran, Bali, Indonesia (8°07'40.5"S, 114°40'05.1"E; Figure 1a) in July of 2011. We deployed a total of six ARMS in two groups of three. ARMS within each group were placed approximately 2 m apart from each other, and the groups of three (Site 1 and Site 2) were separated by approximately 100 m. In June of 2012, we retrieved those six ARMS. To test for temporal variation, we deployed three new ARMS in June of 2012 at Site 1 and collected those ARMS in June of 2013 (Figure 1b). Thus, 1 year was allotted for the colonization of sessile and motile organisms on the artificial reef structures during both deployment periods.

2.2 | ARMS collection and processing

To collect the ARMS, we secured a fitted 100 μm Nitex-lined crate over the ARMS with elastic cords to prevent the loss of any organisms, except, potentially, some microscopic organisms. Upon retrieval, we immediately submerged the ARMS in a container of 40 μm filtered, aerated seawater and transported them to a laboratory in Pemuteran, where we maintained them in filtered, aerated seawater until disassembly. All bins, trays, and other tools were sterilized with bleach and rinsed or soaked in sterile water prior to processing. We processed the ARMS following a standardized disassembly and sampling protocol (Leray & Knowlton, 2015; Ransome et al., 2017).

After we disassembled the ARMS plate by plate, we vigorously shook the plates, one by one, to remove motile organisms, and then we moved each plate into a shallow tray of filtered seawater to photograph the top and bottom of each plate. Using a tripod and a strobe system, we took high resolution photographs of each plate (Figure S1). Next, the predominately motile organisms dislodged from the ARMS plates were separated into three fractions using a series of sieves (2 mm, 500 μm, and 106 μm). Stacking the 2 mm sieve on top of the 500 μm sieve, all water, debris, and organisms from the ARMS disassembly bin was passed through both sieves and into a new bin. The larger, motile organisms from the 2 mm sieve were put aside in a tray for voucher-based DNA barcoding. We repeatedly sieved the remaining sediment and associated organisms to create two fractions: a 500 μm to 2 mm size fraction (hereafter referred to as the 500 μm fraction) and 106 μm to 500 μm size fraction (hereafter referred to as the 100 μm fraction). Subsequent to this fractionation, we concentrated these fractions using a 40 μm Nitex mesh stretched between fitted PVC pipes and then rinsed them with 95% ethanol. Finally, we transferred the concentrated fraction samples to falcon tubes and preserved them with 95% ethanol.

To collect the sessile fraction, we scraped everything off the surface of the ARMS plates with paint scrapers, using filtered

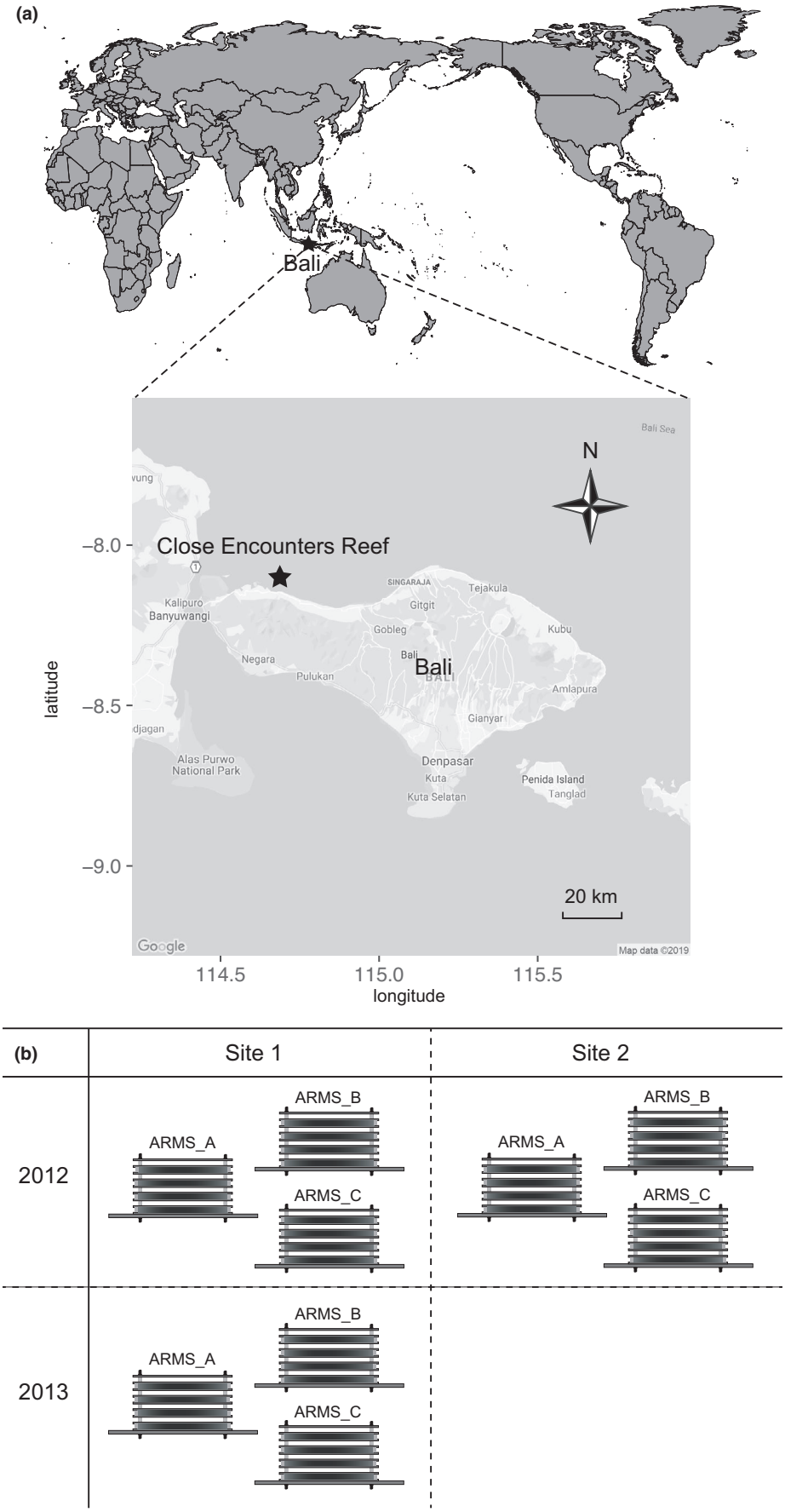

FIGURE 1 (a) Map marking the study reef, Close Encounters Reef, near Bali, Indonesia, and (b) autonomous reef monitoring structure (ARMS) sampling schematic across year and site

seawater to rinse the plates during the process. We then transferred all scraped contents to a blender for homogenization. The blended material was then poured into a 40 μm Nitex mesh net, washed with filtered seawater until it ran clear, and squeezed through the net to dry the fraction. This "sessile fraction" was then transferred into 50 ml falcon tubes (10 g per falcon tube) and preserved with 95% ethanol. Ultimately, we had a total of 27 samples (3 fractions from each of the 9 ARMS).

2.3 | DNA extractions

To prepare the fractions for DNA extraction, the 500 μm and 100 μm fractions were decanted to separate calcium carbonate and terrigenous sediment that could impact the pH and efficiency of the DNA extraction kits. Briefly, we used a 1 L Erlenmeyer flask and sterile water, suspended the entire fraction, allowed the sediment to settle, and decanted the suspended content through a geological sieve to recover the organic contents. We used a 106 μm sieve to decant the 500 μm fraction and a 45 μm sieve to decant the 100 μm fraction. Post-decantation, the sample was halved by volume; one half was transferred to a mortar and pestle for homogenization, and the second half was frozen as back-up material. For the sessile fraction, the homogenized material was vigorously mixed before subsampling to ensure that the subsample was representative.

DNA extraction and sequencing were performed at the National Museum of Natural History (NMNH) in Washington, DC, USA. We used up to 10 g of material from the sessile fraction and half of the material from the 500 μm and 100 μm fractions for the DNA extraction. After the addition of 400 μg/ml of Proteinase K and an overnight incubation at 56°C at 200 rpm, we extracted DNA with a PowerMax Soil DNA Isolation Kit (MoBio, Carlsbad, CA, USA) following the manufacturer's protocols. To remove potential PCR inhibiters from these community DNA extractions, we cleaned the DNA extractions using a PowerClean DNA CleanUp Kit (MoBio) following the manufacturer's protocols.

2.4 | Library preparation and sequencing

COI and 18S library preparation were conducted following published library preparation protocols successfully executed at the NMNH (COI: Leray & Knowlton, 2015; 18S: 2015 Ocean Sampling Day Protocol, Kopf et al., 2015).

We amplified a 313 bp fragment from the COI gene using seven-tailed primer pairs of mICOIintF and jgHCO2190 (Geller et al., 2013; Leray et al., 2013). These primers included six base pair tags on the 5′ end of each primer (Table S1). Our selected COI primers are widely employed across metabarcoding studies and cover 262,144 variations (i.e., unique sequence combinations) of the two COI binding sites. Thus, these primer pairs provide a broad representation of the COI gene region. PCR amplification, sample pooling, and amplicon library preparation methods follow Casey et al. (2019). Briefly, PCR reactions had a total volume of 20 μl: 1 μl of 10 μM forward and reverse primer, 1.4 μl dNTP, 2 μl Advantage 2 DNA Buffer (Takara Bio USA, Mountain View, CA, USA), 0.4 μl Advantage 2 Polymerase (Takara Bio USA), 13.2 μl of distilled water, and 1 μl of 10 ng DNA template. We performed a two-step touchdown PCR. The initial denaturation was at 95°C for 10 min, proceeded by the first step for 16 cycles: 95°C for 10 s, 62°C (−1°C per cycle) for 30 s, and 72°C for 60 s. The second step was run for 20 cycles: 95°C for 10 s, 46°C for 30s, and 72°C for 7 min, followed by a final extension at 72°C for 7 min. We included negative controls in our PCR reactions (no contamination was detected), ran PCR reactions in triplicate, and verified success on 1.2% agarose gels, then all successful reactions were pooled into a single product.

We quantified and pooled PCR products using a Qubit Fluorometer (Invitrogen, Carlsbad, CA, USA) with a Qubit dsDNA HS Assay Kit (Thermo Fisher Scientific, Waltham, MA, USA). We pooled tailed primer pairs in equimolar concentrations, then used bead cleaning at a concentration of 0.8x vol/vol with Agencourt AMPure XP beads (Beckman Coulter, Brea, CA, USA). Library preparation was performed with a TruSeq DNA PCR-Free IT Library Prep Kit (Illumina, San Diego, CA, USA) following the manufacturer's protocols. Following ligation of the adapter sequences, we quantified the indexed samples using a Qubit Fluorometer, pooled 50 ng of each library, then normalized the single library-prepped sample to 4nM prior to sequencing. Sequencing was conducted on an Illumina MiSeq with a MiSeq Reagent Kit v3, 600-cycle (Illumina), using a 1% PhiX spike.

For 18S, we amplified and sequenced a 536 bp fragment from the V4 region using the V4_18SNext.For and V4_18SNext.Rev primers (Piredda et al., 2017; Tragin et al., 2018) following the 2015 Ocean Sampling Day Protocol (Kopf et al., 2015). Again, we selected 18S primers that are widely employed across metabarcoding studies. Briefly, we ran each PCR reaction in a volume of 20 μl: 1.25 μl of 0.5 μM forward and reverse primer, 0.5 μl dNTP, 5 μl 5x High-Fidelity DNA Buffer (Thermo Fisher Scientific, Waltham, MA, USA), 0.5 μl of 1 U Phusion High-Fidelity DNA Polymerase (Thermo Fisher Scientific), 16.0 μl distilled water, and 0.5 μl of 20 ng DNA template. Thermocycling employed a two-step PCR protocol. The initial denaturation was at 98°C for 30 s, then the first step included ten cycles: 98°C for 10 s, 44°C for 30 s, and 72°C for 15 s. The second step included 15 cycles: 98°C for 10 s, 62°C for 30 s, and 72°C for 15 s, followed by a final extension at 72°C for 7 min. We included negative controls in our PCR reactions (no contamination was detected) and verified amplification success on 1.2% agarose gels. Then, we bead cleaned PCR products with Agencourt AMPure XP Beads (Beckman Coulter) at a concentration of 1.2× vol/vol. We then quantified PCR concentrations using a Qubit Fluorometer (Invitrogen) to measure all PCR products with a Qubit dsDNA HS Assay Kit (Thermo Fisher Scientific) to calculate the appropriate DNA concentration for the second round of PCR.

For library preparation, we used a dual index approach with the Nextera DNA Library Prep Kit (Illumina) and the Nextera Index Kit (Illumina). We ran each indexing PCR reaction in a volume of 50 μl:

5 μl of Index 1 (i7), 5 μl of Index 2 (i5), 1 μl dNTP, 10 μl 5x High-Fidelity DNA Buffer (Thermo Fisher Scientific), 0.5 μl of 1 U Phusion High-Fidelity DNA Polymerase (Thermo Fisher Scientific), a volume of DNA template required to yield 40 ng, and the remaining volume of distilled water. The PCR amplification included an initial denaturation at 98°C for 30 s, followed by five cycles as follows: 98°C for 10 s, 65°C for 30 s, and 72°C for 3 min. We used Agencourt AMPure XP Beads (Beckman Coulter) at a concentration of 0.6x vol/vol to clean the PCR products. Prior to sequencing, we quantified indexed samples using a Qubit Fluorometer, then we pooled samples in equimolar ratios prior to normalizing libraries to 2 nM and sequencing them on an Illumina MiSeq with a MiSeq Reagent Kit v2, 500-cycle (Illumina).

2.5 | Sequence bioinformatics

Unlike COI bioinformatic pipelines, 18S pipelines are well-established in the metabarcoding literature. Given the intrinsically distinct nature of these two primer sets, different bioinformatic pipelines were used to process the COI and 18S sequences.

To process the COI sequences, we first merged paired-end reads using PEAR (Zhang et al., 2013), demultiplexed sequences using the FASTX-Toolkit FASTQ Barcode splitter (http://hannonlab.cshl/fastx_toolkit/), and removed tags/indexes with FLEXBAR (Dodt et al., 2012). We then conducted quality filtering with Trimmomatic (removing sequences <200 bp; (Bolger et al., 2014), and additional sequence cleaning was performed via the Multiple Alignment of Coding Sequences (MACSE; Ranwez et al., 2011), which explicitly accounts for the underlying codon structure to filter out inconsistent sequences. Specifically, we aligned sequences to a high-quality library of COI barcodes from the Mo'orea BIOCODE database (Meyer, 2016) to remove sequences with stop codons, frameshifts, insertions, and more than three deletions (Leray et al., 2012). We then clustered dereplicated sequences into operational taxonomic units (OTUs) at 97% similarity using the *uclust* command in USEARCH (Edgar, 2010), discarding all singleton OTUs (OTUs that occur only once across the dataset). We assigned OTUs to phylum using the basic local alignment search tool (BLASTn; Altschul et al., 1990), comparing OTUs against two databases: a local download (on April 14, 2020) of COI data from NCBI excluding environmental samples (but including prokaryotes) and a local BIOCODE database (Meyer, 2016) using the Smithsonian Institution High Performance Computing Cluster (https://doi.org/10.25572/SIHPC). We assigned phylum when sequence similarity was ≥85% (following Ransome et al., 2017), and OTUs with a similarity threshold below that threshold were designated as "unknown." In cases where a different result was obtained from the NCBI and BIOCODE databases, we selected the result from the BIOCODE database because all BIOCODE samples have associated vouchers. Phylum names were based on a recently created hierarchical classification of life (Ruggiero et al., 2015).

For 18S sequences, we used USEARCH for bioinformatic processing. First, we merged pair-end reads and performed quality filtering. We then removed reads shorter than 380 bp and longer than 440 bp. After primer removal, sequences were clustered at 99% similarity. We used VSEARCH for reference-based and *de novo* detection of chimeras, which were subsequently removed, with the Protist Ribosomal Reference (PR^2) database (v4.5; Guillou et al., 2013). After discarding all singleton OTUs, we assigned taxonomy at ≥90% sequence identity with BLASTn using the PR^2 database and the SILVA database (release 128; Pruesse et al., 2007). Phylum-level taxonomy was accepted when sequence similarity was ≥90%, and OTUs below that threshold were designated as "unknown." Again, we named phyla based on a recently created hierarchical classification of life (Ruggiero et al., 2015).

2.6 | Visual annotation of ARMS plates

To visually analyze the taxa on each ARMS plate, we used CoralNet, a semi-automated, online resource that annotates benthic images of coral reefs (Beijbom et al., 2012, 2015; Williams et al., 2019). We initially scored images manually to train the program, then we used semi-automation for the remaining assignments, but we visually verified automatically scored plates to ensure accuracy. Following standard ARMS protocols (oceanarms.org), we used the following categories for ARMS annotations: bryozoan, sponge, colonial tunicate, red encrusting algae, green algae, foraminifera, calcareous worm tube, soft worm tube, bivalve, hard coral, solitary tunicate, soft coral, no recruitment, and unknown. For each plate, we employed a 15 × 15 point matrix, for a total of 225 annotated points. To facilitate comparison with the COI and 18S metabarcoding datasets, we summarized all annotations at the phylum level (Annelida, Bryozoa, Chlorophyta, Chordata, Cnidaria, Mollusca, Porifera, Protista, and Rhodophyta), and we calculated averages for each phylum across all plates (including tops and bottoms) from a single ARMS.

2.7 | Data analysis

We ran all analyses with the statistical software R (version 3.6.1; R Core Team, 2019). Data wrangling was performed with the *tidyverse* package (Wickham, 2017), and all visualizations were made with the packages *ggmap* (Kahle & Wickham, 2013) and *ggplot2* (Wickham, 2016).

From the raw OTU tables, we produced rarefaction curves for the COI and 18S datasets to visualize sampling effort and OTU richness across sequences, as well as to simulate extrapolated projections with additional sampling with the *iNEXT* package (Chao et al., 2014). For both markers, we generated rarefaction curves for each fraction (500 μm, 100 μm, and sessile fractions) of each ARMS, as well as for each ARMS as a singular unit after summing across fractions.

Next, we converted the raw sequence read data to relative read abundances (RRA; in accordance with Deagle et al., 2019), which we used for the majority of the analyses and is a fitting metric for the analysis of hyperdiverse metabarcoding data (Casey et al., 2019).

Using the RRA data, we generated non-metric multidimensional scaling (nMDS) ordinations based on Bray–Curtis dissimilarity matrices with the *vegan* package (Oksanen et al., 2018) to examine the clustering relationships between ARMS and fractions for the COI and 18S markers. The nMDS ordinations were also used to visualize the potential impact of spatiotemporal variables (year and site) on these assemblages.

All taxonomy-based analyses were conducted at the level of phylum. We found a higher percent of unidentified sequences in the COI dataset as compared to the 18S dataset (COI = 41.1%; 18S = 0.6%), which substantially impacted RRAs across identified phyla. Thus, for all phylum-based analyses, we removed all unidentified OTUs from the raw dataset and recalculated RRAs using only OTUs with phylum-level taxonomic assignments. Furthermore, we maintained the separation between spatiotemporal variables across all analyses to avoid confounding factors. To visualize the performance of each marker, we calculated and ranked the 10 phyla with the highest average relative abundances across the ARMS from each spatiotemporal treatment (2012_Site1, 2012_Site2, and 2013_Site1), as well as creating an "Other" category to summarize all remaining phyla not included in the top 10 (see list of phyla in the "Other" category; Table S2).

To analyze the taxonomic differences between the COI *versus* 18S taxonomic data from the ARMS summed across fractions, we ran a Bayesian mixed model with the packages *brms* (Bürkner, 2017), *tidybayes* (Kay, 2019), and *rstan* (Stan Development Team, 2019). We examined the 20 co-occurring phyla that occurred on each ARMS across both metabarcoding datasets. The model assessed the effect of marker choice (COI *versus* 18S) on the RRAs of each phylum. To tease apart the effects of the spatiotemporal variables, we incorporated treatment (year and site) as a random effect.

We then ran a Bayesian mixed model to compare the taxonomic differences of the COI and 18S data from the sessile fraction of each ARMS to the visually annotated ARMS plate data, again using the packages *brms* (Bürkner, 2017), *tidybayes* (Kay, 2019), and *rstan* (Stan Development Team, 2019). This analysis included the seven co-occurring phyla on each ARMS across all datasets: Annelida, Bryozoa, Chlorophyta, Chordata, Mollusca, Porifera, and Rhodophyta. We recalculated RRAs after subsetting each dataset to include these seven phyla. The model examined the impact of quantification technique (COI, 18S, or visual annotation) on the RRAs of each phylum. We included treatment (year and site) as a random effect.

For both models, we applied a log-transformation (log(x)) on the OTU relative abundances and fitted the model to a student *t* distribution to avoid outliers that would heavily impact the model (Gelman et al., 2013). The models were run with four chains for 2000 iterations, each containing 1000 warm-up samples. We specified weakly informative priors with a uniform distribution (b ~ uniform (−100,0)) given that each value, as a predicted relative abundance, should not exceed one, and our data were log-transformed. For each model, the posterior predictive distributions reflected the observed data, and all model parameters had stable trace plots.

3 | RESULTS

Following quality filtering and the removal of singletons, the final COI OTU table included 3,964,674 sequences and 31,900 OTUs; for 18S, the final OTU table included 3,696,915 sequences and 25,994 OTUs. In the COI dataset, an average of 41.1% of the sequences across samples were unidentified to any taxonomic level, representing 49.44% (15,770) of all OTUs. In contrast, in the 18S dataset, only an average of 0.62% of the sequences across samples were unidentified to any taxonomic level, representing 0.8% (209) of all OTUs. There was a higher phylum-level taxonomic diversity represented in the 18S dataset, which recovered 51 phyla as compared to 38 phyla in the COI dataset. Of these phyla, 20 phyla co-occurred across each ARMS in both datasets, and 65% of these co-occurring phyla were metazoans.

For COI, the number of OTUs per ARMS ranged from 6,580 to 14,237 OTUs, and the sequencing depth ranged from 214,705 to 735,839 sequences. For 18S, the number of OTUs per ARMS ranged from 7,113 to 11,237 OTUs, and the sequencing depth ranged from 286,645 to 527,256 sequences. Within each ARMS, rarefaction curves did not reach saturation, but OTU accumulation slowed (Figure 2). However, within each fraction (i.e., 500 μm fraction, 100 μm fraction, and sessile fraction) of each ARMS, rarefactions rarely approached saturation (Figure S2). Thus, summing the three fractions within an ARMS provides a more complete representation of sequence coverage and taxonomic diversity due to the substantial taxonomic overlap among fractions. Consequently, most downstream analyses focused on pooled fractions to represent the entire ARMS rather than individual fractions.

For both the COI and 18S markers, the nMDS ordination analysis of the ARMS samples revealed high levels of clustering among individual ARMS (shaded triangles) and fractions from different ARMS (dotted polygons) (Figure 3). There was a high level of overlap between the ARMS samples collected from Site 1 and Site 2 in 2012 (i.e., spatial overlap). Furthermore, the ARMS samples from Site 1 in 2012 had partial overlap with the ARMS samples collected from Site 1 in 2013 (i.e., temporal overlap). We also detected fine-scale clustering: the 100 μm and sessile fractions clustered more tightly among themselves than the 500 μm fraction. Overall, intrinsic ARMS-related drivers (ARMS and fraction) exhibit higher levels of clustering than spatial and temporal variables, and these patterns do not differ according to marker.

Of the 10 phyla with the highest average relative abundances across ARMS (across years and sites) in each dataset, 9 out of 10 of the phyla were co-occurring (Figure 4). Only the phyla with the 10th highest average relative abundance in the COI dataset (Ochrophyta) and 18S dataset (Platyhelminthes) did not occur in the top 10 phyla of the other dataset. Notably, the average relative abundance of Platyhelminthes was extremely low across the COI dataset as compared to the 18S dataset (COI = <0.01%, 18S = 1.47%). Overall, Porifera had the highest average relative abundance in the COI dataset, which was considerably higher than in the 18S dataset

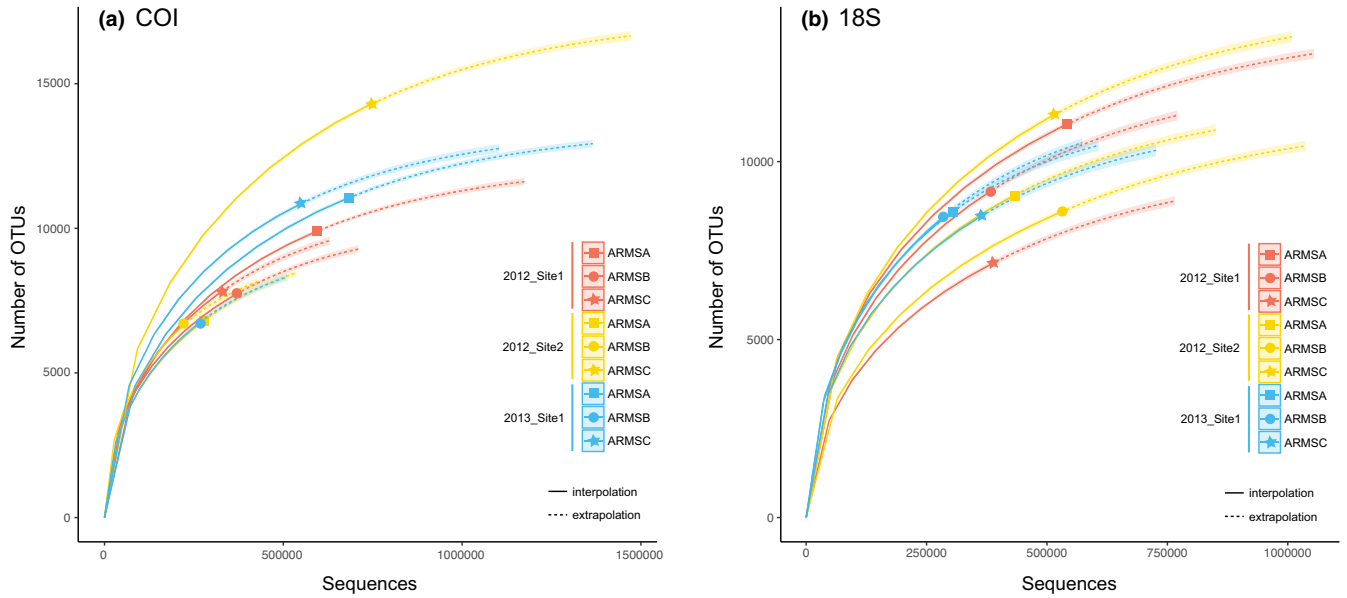

FIGURE 2 Rarefaction curves for the (a) COI and (b) 18S markers showing operational taxonomic unit (OTU) richness according to total sequences for each autonomous reef monitoring structure (ARMS) in each year and site. Solid lines indicate interpolated values, and dotted lines indicate predicted extrapolated values

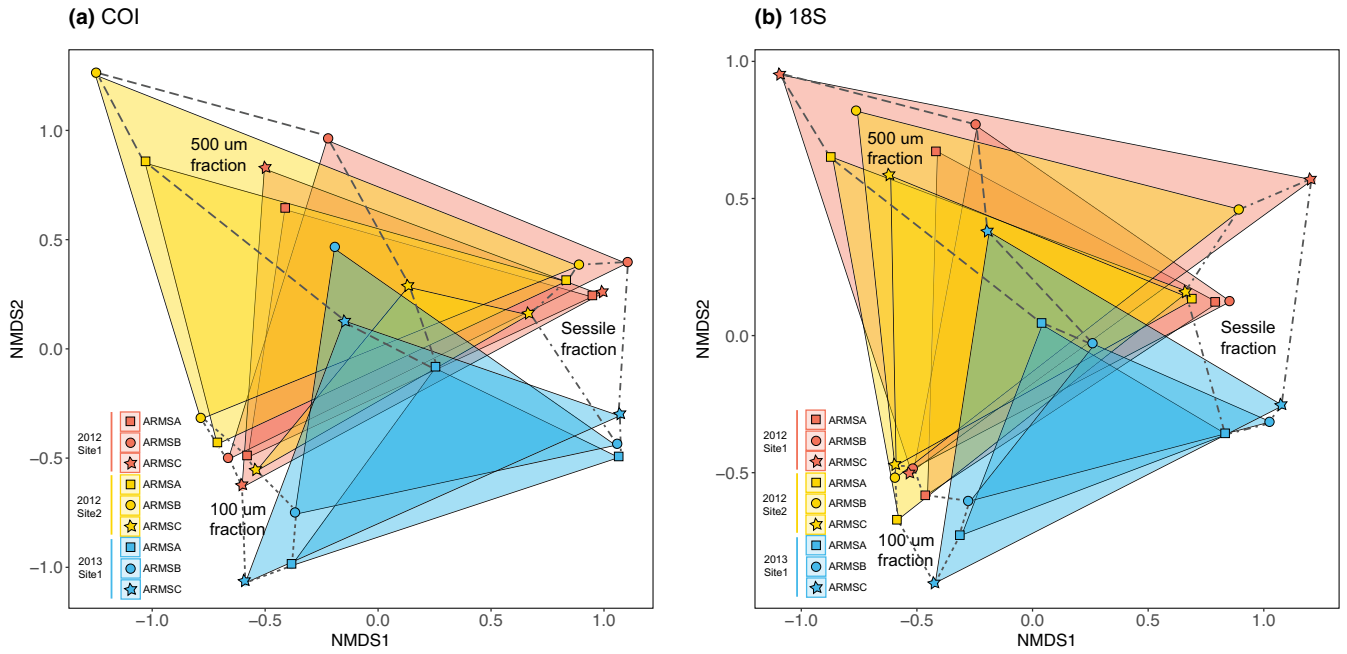

FIGURE 3 Non-metric multidimensional scaling (nMDS) ordination plots for the (a) COI and (b) 18S markers showing the association of autonomous reef monitoring structure (ARMS; colored triangles) and fractions (dotted polygons), as well as the impact of spatiotemporal variables (year and site). Each point indicates a single ARMS, either from Site 1 in 2012, Site 2 in 2012, or Site 1 in 2013. Fractions include the 500 μm–2 mm fraction (500 μm fraction), the 500 μm–106 μm fraction (100 μm fraction), and the sessile fraction

(COI = 17.22%, 18S = 8.98%), and Chordata had the highest average relative abundance in the 18S dataset, which was substantially higher than in the COI dataset (COI = 3.94%, 18S = 18.31%). Arthropoda had the second highest average relative abundance in both datasets and occurred at a similar frequency (COI = 16.33%, 18S = 18.19%); however, the rest of the phyla differed in rankings between the two markers. A higher proportion of the 18S data (6.71%) as compared to the COI data (1.61%) comprised of the "Other" category (i.e., sum of all remaining taxa that were identified to phylum but did not rank in the top 10 phyla).

Several strong differences emerged between the two markers upon modeling the relative abundances of the 20 co-occurring

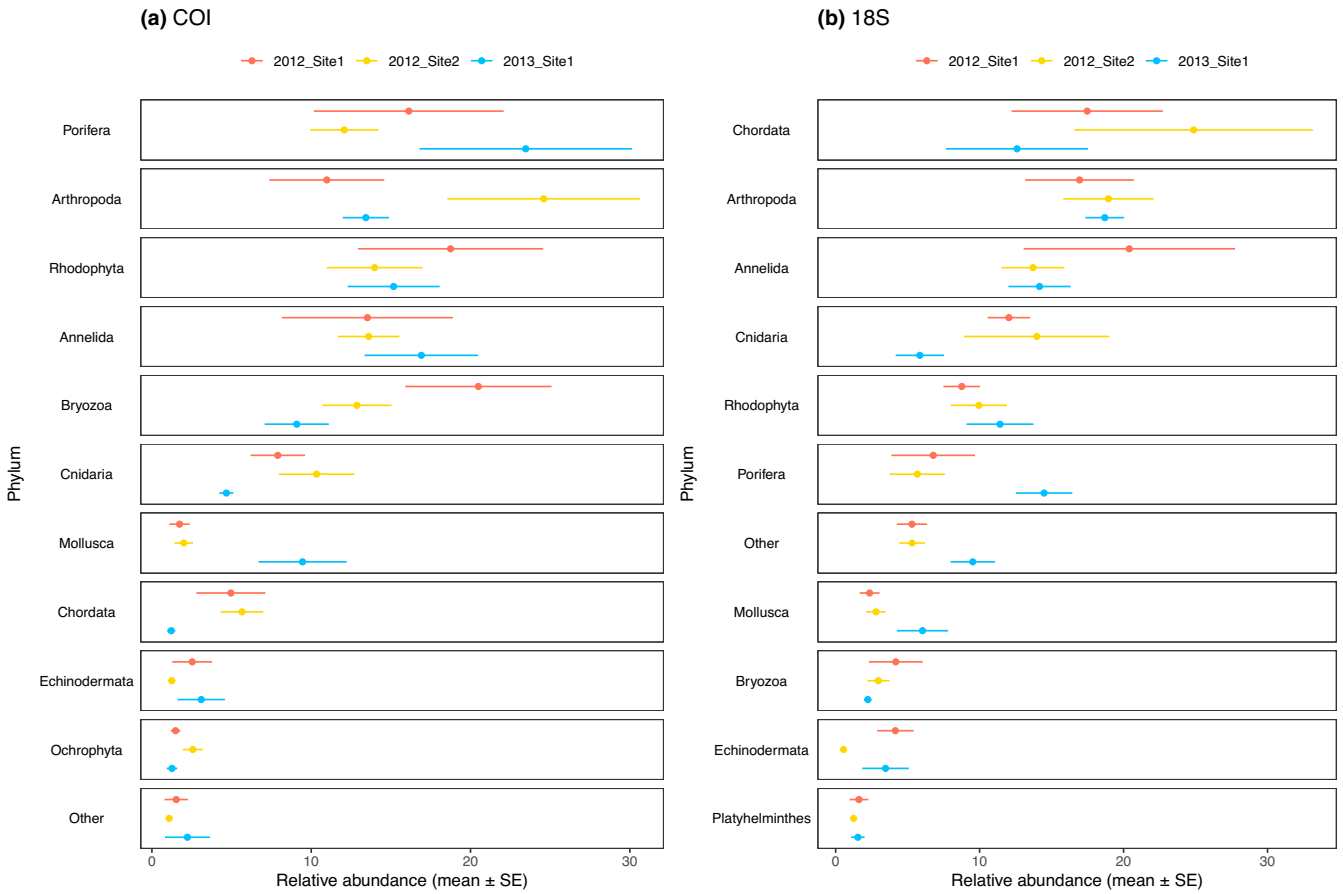

FIGURE 4 The 10 phyla with the highest average relative abundances (mean ±SE) for the (a) COI and (b) 18S markers. Relative abundances were averaged across autonomous reef monitoring structures (ARMS; $n = 3$) and separated by year and site. The "Other" category sums across all remaining phyla not included in the top 10

phyla (Figure 5). The most notable differences occurred in the predicted relative abundances of Amoebozoa, Ascomycota, Bryozoa, Chlorophyta, Chordata, Entoprocta, and Nematoda. Of these seven phyla, five of them were rare across both datasets (see Table S3 for means ±95% credible intervals), so the discordant results between the markers for these phyla are less relevant for community wide patterns. However, the predicted relative abundances of Bryozoa and Chordata were comparatively high, with both ranking in the top 10 most abundant phyla in the COI and 18S datasets. Consistently across all spatiotemporal variables, Bryozoa exhibited a markedly higher predicted relative abundance in the COI *versus* 18S dataset, whereas Chordata had a substantially lower predicted relative abundance in the COI *versus* 18S dataset. The COI marker also predicted a slightly higher relative abundance for Porifera as compared to the 18S marker.

The spatiotemporal variables had a minimal impact on the distribution of phyla for both the COI and 18S markers (Figure S3). Year had a larger effect size than site, as was also showcased in the nMDS ordination (Figure 3). The only notable trends were the slight increase of Mollusca and Entoprocta in 2013 as compared to 2012. Importantly, time and space resulted in comparatively minimal differences in taxonomic composition between the two markers.

Finally, notable differences emerged in the comparison between the molecular markers and the image annotation from ARMS plates with CoralNet. When limiting the analysis to the sessile fraction in the molecular datasets, the 18S marker continued to represent the highest phylum-level diversity, with 44 phyla appearing in the 18S dataset as compared to 32 phyla in the COI dataset. Of these phyla, 14 phyla co-occurred across each ARMS in both datasets. We only scored nine phyla across the annotated ARMS plates; this number was limited by the standardized, predefined categories used for ARMS annotations. Modeling the relative abundances of the seven co-occurring phyla among the COI dataset, 18S dataset, and the annotation of ARMS plates revealed marked differences across three phyla: Bryozoa, Chlorophyta, and Chordata (Figure 6). The COI marker predicted similar relative abundances as the annotated ARMS plates for Bryozoa and Chordata, whereas the 18S marker predicted substantially lower relative abundances for Bryozoa and higher relative abundances for Chordata (see Table S4 for means ±95% credible intervals). For Chlorophyta, the COI and 18S markers predicted markedly lower relative abundances than the annotated ARMS plates, but this phylum was comparatively rare across all datasets. The COI marker and the annotated ARMS plates also predicted similar relative abundances for Porifera, but the predictions were

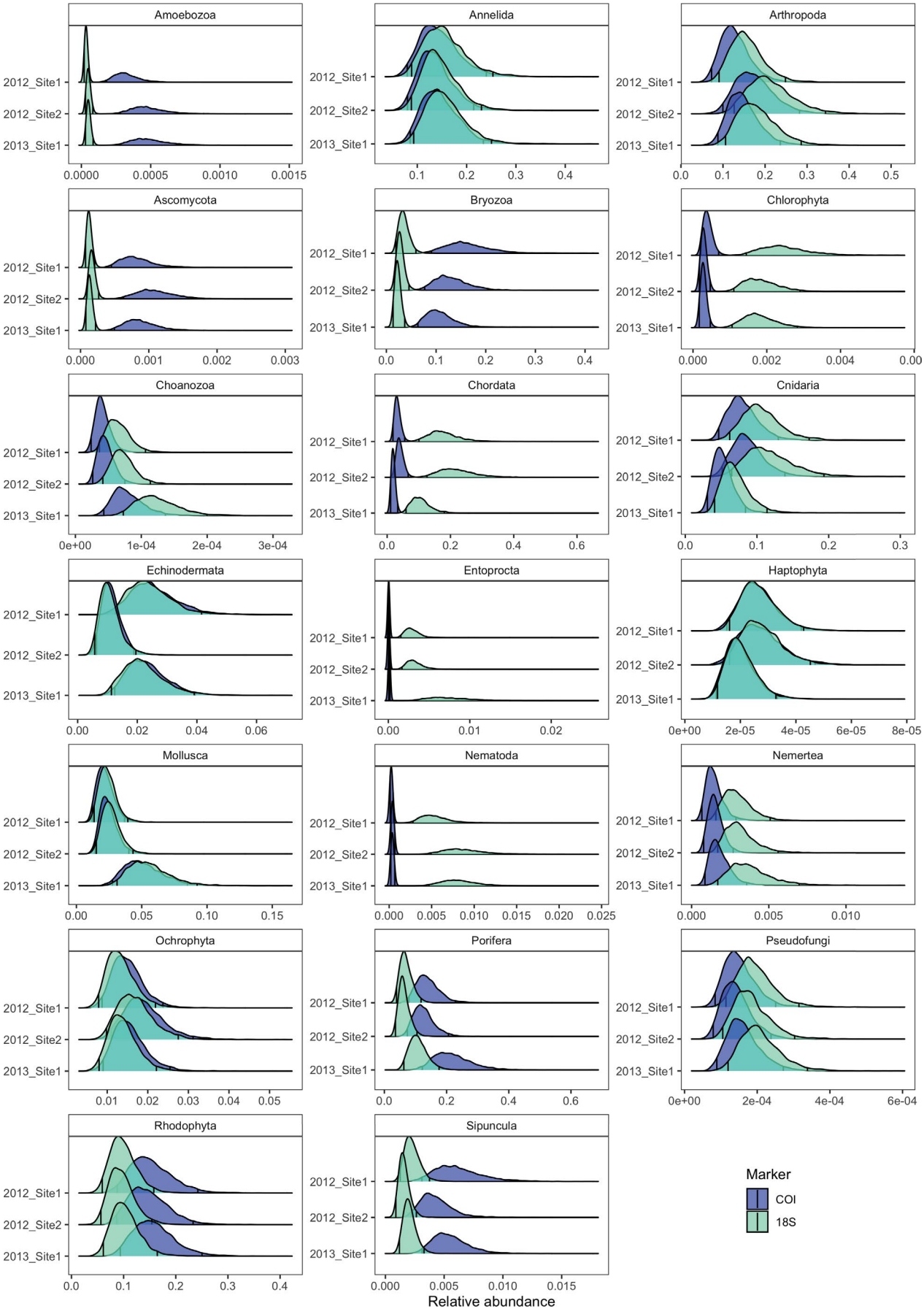

FIGURE 5 The impact of COI *versus* 18S marker choice on the relative abundances of the 20 co-occurring phyla, separated by year and site. The mean posterior distributions (±95% highest density intervals (HDIs) designated by vertical black bars within each curve) are from a Bayesian mixed model. When the HDIs between the two curves do not overlap, this indicates strong differences between quantification techniques

slightly lower for the 18S marker. The molecular markers and annotations provided remarkably similar predicted relative abundances for Mollusca and Rhodophyta.

4 | DISCUSSION

The standardized sampling provided by ARMS creates a unique opportunity to rigorously assess potential sources of biases in DNA metabarcoding of marine eukaryotic biodiversity. Although marker choice did not impact estimates of OTU richness and overall diversity patterns, the COI and 18S datasets provided distinct pictures of eukaryotic taxonomic diversity. 18S metabarcoding identified 25% more phyla and resulted in vastly higher identified OTUs than COI metabarcoding. However, COI more faithfully recovered diversity patterns of sessile organisms on ARMS plates, as supported by visual image analysis. In contrast to marker choice, year and location of ARMS deployment had a minimal impact on eukaryotic communities. Thus, consistently across small-scale spatial and temporal variations, the taxonomic representation of different "universal" metabarcoding markers provided skewed estimates of biodiversity patterns.

4.1 | Markers matter

One of the major dissimilarities between the two markers was the more than 50-fold difference in unidentified OTUs in the COI (15,770 OTUs; 49.44% of all OTUs) compared to the 18S (209 OTUs; 0.8% of all OTUs) datasets. Importantly, COI failed to significantly represent (<0.01% across the dataset) several taxa that had considerable occurrences across the 18S dataset, such as Entoprocta (1.17%), Platyhelminthes (1.47%), and Nematoda (0.82%). Previous studies highlight that COI does not perform well in Platyhelminthes or Nematoda (Andújar et al., 2018; Leray & Knowlton, 2017; Prosser et al., 2013). However, Platyhelminthes and Nematoda only represented ~2.5% of the total OTUs recovered by 18S, so the poor performance of COI for these two phyla does not account for the nearly 50% of unidentified OTUs in the COI dataset. Inadequate COI reference library coverage may also hinder our ability to assign taxonomy to OTUs (Cristescu, 2014; DiBattista et al., 2020; Gold et al., 2020; Ransome et al., 2017). In addition, although fast mutation rates in COI often permit species-level identification (Hebert et al., 2003), they may impede phylum-level assignment for certain taxa, even when employing a relatively conservative 85% similarity threshold.

When examining only the co-occurring phyla between the two markers, COI provided a markedly higher estimate of Bryozoa, while 18S provided a higher estimate of Chordata. Furthermore, based on comparisons to visual annotations, the COI marker appears to more faithfully predict the relative abundances of several of the more dominant sessile taxa on ARMS plates (e.g., Bryozoa, Chordata, and Porifera). As such, 18S is not universally superior to COI. The differential performance of these markers highlights the importance of assessing the intrinsic taxonomic biases associated with amplicon sequencing so that we may select appropriate universal markers and adopt a more nuanced interpretation of metabarcoding data. Such work is vital as we expand our reliance on high-throughput sequencing for biodiversity assessment and management (Zinger et al., 2019).

4.2 | Metabarcoding versus visual assessment

While it is essential to understand how metabarcoding results vary based on marker choice, it is equally important to assess how these markers reflect actual patterns of diversity. We identified several strong differences between metabarcoding *versus* visual image analysis when estimating the dominant sessile phyla on ARMS plates. Both COI and image analysis recovered a higher average relative abundance of Bryozoa and Porifera than the 18S marker. In contrast, the 18S marker predicted a higher relative abundance of Chordata, most likely dominated by tunicates, than COI and image analysis.

Although the COI and 18S rDNA markers are both commonly utilized to examine phylogenetic relationships within Bryozoa, Porifera, and Chordata (Erpenbeck & Wörheide, 2007; Fuchs et al., 2009; Holland, 2016), specific primer sets have highly variable amplification success rates across these taxa based on the original intent of the primer design. For example, the selected COI primer set used in this study was designed to optimally amplify macroscopic metazoans across a broad taxonomic range from coral reef environments (Geller et al., 2013; Leray et al., 2013). In contrast, the 18S primer set was originally designed to amplify protist assemblages in the Mediterranean (Piredda et al., 2017; Tragin et al., 2018). As such, our marker comparison is limited to the performance of these particular primer sets, and variation in marker performance across taxa should be expected and carefully considered in metabarcoding study design.

Importantly, while the 18S rDNA gene region is broadly informative across metazoan phylogeny and was even used in the first molecular phylogenetic analysis of the metazoan tree of life (Field et al., 1988), current phylogenies often use a broad range of longer nuclear and mitochondrial genes to account for the variability in gene regions across metazoans (Bourlat et al., 2008). Such variation cannot be accounted for in a single marker with a short enough length to be compatible with current sequencing technologies. The higher rate of consensus between the COI marker and the annotated ARMS plates suggests that at least in hyperdiverse coral reef environments, targeting the hypervariable COI region may provide a more reliable characterization of the relative abundances of dominant eukaryote phyla than the 18S region. Moreover, the COI profiles yield more

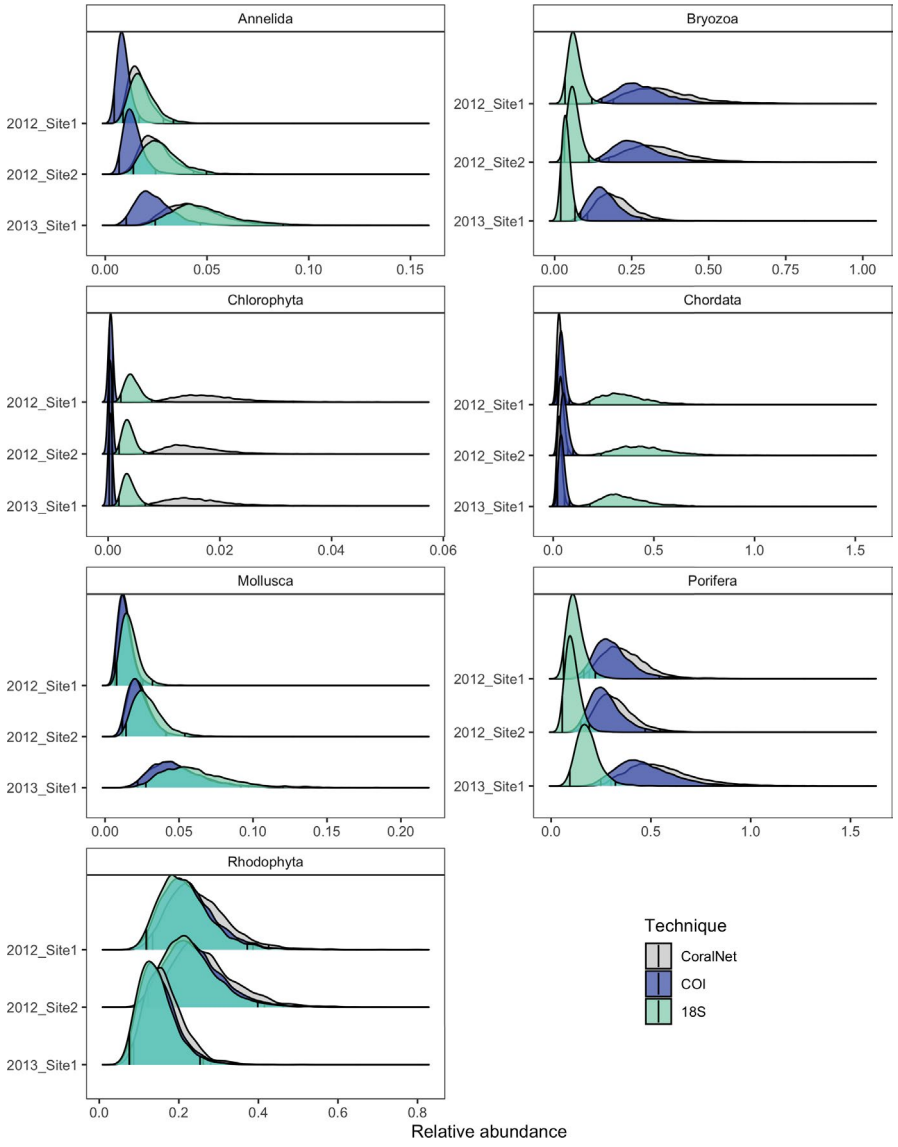

FIGURE 6 The effect of COI marker choice, 18S marker choice, and visual image analysis of ARMS plates on the relative abundances of the seven co-occurring phyla, separated by year and site. The mean posterior distributions (±95% highest density intervals (HDIs) designated by vertical black bars within each curve) are from a Bayesian mixed model. When the HDIs between the two curves do not overlap, this indicates strong differences among quantification techniques

taxonomic precision, especially as sampling expands across biogeographic regions and reference databases become more complete (DiBattista et al., 2020; Gold et al., 2020).

An important caveat, however, is that the apparent consistency in the proportions of dominant sessile phyla between the COI marker and the annotated ARMS plates required the removal of many rare phyla from the metabarcoding datasets to perform this cross-technique comparison. The removal of rare taxa disproportionately masked the performance of the 18S marker, which recovered the highest phylum-level diversity with 45 phyla for 18S *versus* 33 phyla for COI. Only nine phyla were recognized across the annotated ARMS plates, but this characterization was limited by the standardized categories employed in ARMS visual annotations. Even among these limited categories, visual estimates can be compromised due to the competitive nature of dominant benthic organisms, such as algae, sponges, bryozoans, and tunicates, which may overgrow other encrusting marine taxa (Chadwick & Morrow, 2011) and make them undetectable via visual analysis. As such, further studies are required to refine our understanding of these potential sources of bias in metabarcoding studies.

Although COI was superior to 18S in our visual comparisons, key components of marine communities, such as microbial eukaryotic organisms (i.e., protists), are not visible to the human eye and cannot be detected using current methods of automated image analysis. Instead, amplicon sequencing, predominantly 18S rDNA sequencing, is frequently employed to delineate the molecular diversity of these important cryptic organisms (Epstein & López-García, 2008). For these taxa, the 18S gene region will likely continue to provide significantly better estimates of cryptic taxonomic diversity than the COI marker or annotated image analysis, highlighting the utility of DNA metabarcoding, especially with certain markers, to reveal cryptic marine taxa (DiBattista et al., 2020; Ji et al., 2013; Leray & Knowlton, 2015; Ransome et al., 2017).

4.3 | Impact of space and time

While comparisons of COI and 18S data revealed significant marker-based biases, spatiotemporal variables only had a minor impact on taxonomic distributions, with the impact of time having a slightly

stronger effect than site. Understanding spatial and temporal variation in marine biodiversity is critical to monitoring marine ecosystems (Elahi et al., 2015). In terms of micro-eukaryotes (e.g., meiofauna or plankton), such comparisons are, in fact, one of the main applications of DNA metabarcoding. While spatial variation may impact marine eukaryotes in sediment (Leasi et al., 2016) or plankton (Pearman et al., 2014) samples, temporal factors and associated environmental variables (e.g., salinity, temperature) are the primary drivers that shape these eukaryotic assemblages (Brannock et al., 2016; Chain et al., 2016; DiBattista et al., 2020; Salonen et al., 2019).

In our study, the relatively minor temporal effects may be attributed to our study system, tropical coral reefs, which do not have strong seasonal variation compared to temperate marine systems. The lack of temporal variation may also be attributed to sample collection occurring in June in both 2012 and 2013. A seasonal effect may have been detected if samples were taken over the course of an entire year. Nonetheless, we detected a greater impact of temporal variation as compared to spatial variation (across the limited spatial range of approximately 100 m), suggesting that ARMS may be sensitive to shifts in reef systems over longer time periods. As such, the ability to detect changes in cryptic eukaryotic communities associated with ARMS using DNA metabarcoding may provide a powerful approach to monitor environmental change on coral reefs across space and time (Ransome et al., 2017). While our data show that ARMS are sensitive to minor shifts in community structure from year to year, further studies across wider spatial and temporal gradients will be required to comprehensively examine the impact of space and time on eukaryotic coral reef assemblages.

4.4 | Moving forward with metabarcoding

One of the main limiting factors in our comparison of the COI and 18S markers was the comparatively high number of unidentified OTUs in the COI dataset (49.44% of OTUs and 41.1% of sequences) as compared to the 18S (0.8% of OTUs and 0.62% of sequence). High rates of homoplasy in the hypervariable COI gene region make the assignment of OTUs difficult without the presence of reference sequences from closely related organisms (Deagle et al., 2014; Leray & Knowlton, 2015, 2016). Despite the increasing number of COI records available in public databases (Porter & Hajibabaei, 2018), the lack of comprehensive COI reference barcode libraries across the wide taxonomic scope of hyperdiverse assemblages (such as coral reefs) remains the greatest challenge to harnessing the full capacity of COI metabarcoding, and the contrasting ability of the more slowly evolving, yet better sampled, 18S marker to identify the same community is striking.

To surmount these limitations, we encourage employing a multimarker approach to metabarcoding across a variety of marine habitats and regions (see also van der Loos & Nijland, 2020). Such studies are essential to identify "dark" taxa that remain undescribed, even morphologically, and will help prioritize voucher-based sampling to advance catalogues of marine biodiversity and hasten the completion of reference databases. Currently, the scientific community considers taxonomic descriptions as low-impact science, impeding taxonomists' access to research funding and stalling this fundamental characterization of biodiversity (Agnarsson & Kuntner, 2007). Although taxonomy-free metabarcoding methods are an alternative, taxonomy often plays a critical role in metabarcoding marine communities, and increasing DNA barcoding library coverage is essential to accurately monitor biodiversity patterns with advanced molecular techniques. Highlighting the gaps that currently limit the power of metabarcoding should help re-prioritize taxonomic investigations, descriptions, and biodiversity inventories.

5 | CONCLUSIONS

The two most commonly employed metabarcoding markers to estimate eukaryotic biodiversity, COI and 18S rDNA, recovered distinct patterns of eukaryotes in a hyperdiverse coral reef environment. COI more accurately reflected the relative abundances of dominant taxa while 18S provided a deeper, and likely more complete, representation of cryptic taxonomic diversity. Thus, ideal primer selection for metabarcoding is largely dependent on study system, research question, and desired taxonomic resolution. To bridge the limitations of these universal markers, we encourage a multimarker approach to DNA metabarcoding (e.g., Günther et al., 2018; Stefanni et al., 2018; West et al., 2020; G. K. Zhang et al., 2018).

As metabarcoding quickly becomes an indispensable tool for biodiversity monitoring, it is essential to expand efforts to ground truth methodological biases related to amplicon sequencing. While understanding methodological biases, such as marker selection, are important, only through a combination of focused biodiversity inventories and a concerted effort to expand taxonomic coverage of reference libraries will DNA metabarcoding achieve its full potential for biodiversity monitoring. Given the environmental challenges and unprecedented rate of change facing biomes across the planet, such work is increasingly urgent.

FUNDING INFORMATION

This work was funded by the National Science Foundation (NSF) Partnerships for International Research and Education (PIRE) program (OISE 1243541) and the United States Agency for International Development (497-A-00-10-00008-00).

ACKNOWLEDGMENTS

We thank the Indonesian Government, the Ministry for Research and Technology (RISTEK), the Indonesian Institute of Sciences (LIPI), and the Indonesian Biodiversity Research Center (IBRC) for supporting this work. We also thank Dr. Forest Rohwer for his support as a principle investigator and Dr. Irma Arlyza and Dirhamsyah from the LIPI Research Center for Oceanography. This work was conducted under permit numbers 160/SIP/FRP/SM/VI/2012 (P.H.B.), 185/SIP/FRP/SM/VI/2013 (P.H.B.), and 221/SIP/FRP/SM/VI/2013 (C.P.M.).

CONFLICT OF INTEREST

None declared.

AUTHOR CONTRIBUTIONS

J.M.C., E.R., A.G.C., P.H.B., J.B.G., and C.P.M. contributed to the study design. A.M., E.M.K., A.S., N.K.D.C., A.W.A., P.H.B., J.B.G., and C.P.M. deployed, collected, and processed the ARMS. E.R., A.M., and E.M.K. performed molecular lab work. E.R., A.G.C., M.B., and C.P.M. conducted pipeline bioinformatics. J.M.C. and N.M.D.S. analyzed the data. J.M.C. wrote the first draft of the paper, and all authors contributed to writing and approved the final draft for publication.

DATA AVAILABILITY STATEMENT

Illumina MiSeq FASTA files and high-quality images of the ARMS plates are available on the California Digital Library (https://n2t.net/ark:/21547/DXy2). The final OTU tables, CoralNet data, metadata, and R scripts are available on Dryad (https://doi.org/10.5061/dryad.rr4xgxd8r).

ORCID

Jordan M. Casey https://orcid.org/0000-0002-2434-7207
Paul H. Barber https://orcid.org/0000-0002-1486-8404

REFERENCES

Agnarsson, I., & Kuntner, M. (2007). Taxonomy in a changing world: Seeking solutions for a science in crisis. *Systematic Biology*, *56*(3), 531–539. https://doi.org/10.1080/10635150701424546

Alberdi, A., Aizpurua, O., Gilbert, M. T. P., & Bohmann, K. (2018). Scrutinizing key steps for reliable metabarcoding of environmental samples. *Methods in Ecology and Evolution*, *9*(1), 134–147. https://doi.org/10.1111/2041-210X.12849

Al-Rshaidat, M. M. D., Snider, A., Rosebraugh, S., Devine, A. M., Devine, T. D., Plaisance, L., Knowlton, N., & Leray, M. (2016). Deep COI sequencing of standardized benthic samples unveils overlooked diversity of Jordanian coral reefs in the northern Red Sea. *Genome*, *59*(9), 724–737. https://doi.org/10.1139/gen-2015-0208

Altschul, S. F., Gish, W., Miller, W., Myers, E. W., & Lipman, D. J. (1990). Basic local alignment search tool. *Journal of Molecular Biology*, 403–410. https://doi.org/10.1016/S0022-2836(05)80360-2

Andújar, C., Arribas, P., Yu, D. W., Vogler, A. P., & Emerson, B. C. (2018). Why the COI barcode should be the community DNA metabarcode for the metazoa. *Molecular Ecology*, *27*(20), 3968–3975. https://doi.org/10.1111/mec.14844

Baird, D. J., & Hajibabaei, M. (2012). Biomonitoring 2.0: A new paradigm in ecosystem assessment made possible by next-generation DNA sequencing. *Molecular Ecology*, *21*(8), 2039–2044. https://doi.org/10.1111/j.1365-294X.2012.05519.x

Beijbom, O., Edmunds, P. J., Kline, D. I., Mitchell, B. G., & Kriegman, D. (2012). Automated annotation of coral reef survey images. *IEEE Conference on Computer Vision and Pattern Recognition*, *2012*, 1170–1177. https://doi.org/10.1109/CVPR.2012.6247798

Beijbom, O., Edmunds, P. J., Roelfsema, C., Smith, J., Kline, D. I., Neal, B. P., Dunlap, M. J., Moriarty, V., Fan, T.-Y., Tan, C.-J., Chan, S., Treibitz, T., Gamst, A., Mitchell, B. G., & Kriegman, D. (2015). Towards automated annotation of benthic survey images: Variability of human experts and operational modes of automation. *PLoS One*, *10*(7), e0130312. https://doi.org/10.1371/journal.pone.0130312

Bhadury, P., Austen, M., Bilton, D., Lambshead, P., Rogers, A., & Smerdon, G. (2006). Development and evaluation of a DNA-barcoding approach for the rapid identification of nematodes. *Marine Ecology Progress Series*, *320*, 1–9. https://doi.org/10.3354/meps320001

Bolger, A. M., Lohse, M., & Usadel, B. (2014). Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics*, *30*(15), 2114–2120. https://doi.org/10.1093/bioinformatics/btu170

Bourlat, S., Nielsen, C., Economou, A., & Telford, M. (2008). Testing the new animal phylogeny: A phylum level molecular analysis of the animal kingdom. *Molecular Phylogenetics and Evolution*, *49*(1), 23–31. https://doi.org/10.1016/j.ympev.2008.07.008

Brannock, P. M., & Halanych, K. M. (2015). Meiofaunal community analysis by high-throughput sequencing: Comparison of extraction, quality filtering, and clustering methods. *Marine Genomics*, *23*, 67–75. https://doi.org/10.1016/j.margen.2015.05.007

Brannock, P. M., Ortmann, A. C., Moss, A. G., & Halanych, K. M. (2016). Metabarcoding reveals environmental factors influencing spatio-temporal variation in pelagic micro-eukaryotes. *Molecular Ecology*, *25*(15), 3593–3604. https://doi.org/10.1111/mec.13709

Braukmann, T. W. A., Ivanova, N. V., Prosser, S. W. J., Elbrecht, V., Steinke, D., Ratnasingham, S., de Waard, J. R., Sones, J. E., Zakharov, E. V., & Hebert, P. D. N. (2019). Metabarcoding a diverse arthropod mock community. *Molecular Ecology Resources*, *19*(3), 711–727. https://doi.org/10.1111/1755-0998.13008

Brown, E. A., Chain, F. J. J., Crease, T. J., MacIsaac, H. J., & Cristescu, M. E. (2015). Divergence thresholds and divergent biodiversity estimates: Can metabarcoding reliably describe zooplankton communities? *Ecology and Evolution*, *5*(11), 2234–2251. https://doi.org/10.1002/ece3.1485

Bürkner, P.-C. (2017). brms: An R Package for Bayesian Multilevel Models Using Stan. *Journal of Statistical Software*, *1*(1), 2017. https://www.jstatsoft.org/v080/i01

Butchart, S. H. M., Walpole, M., Collen, B., van Strien, A., Scharlemann, J. P. W., Almond, R. E. A., Baillie, J. E. M., Bomhard, B., Brown, C., Bruno, J., Carpenter, K. E., Carr, G. M., Chanson, J., Chenery, A. M., Csirke, J., Davidson, N. C., Dentener, F., Foster, M., Galli, A., … Watson, R. (2010). Global biodiversity: Indicators of recent declines. *Science*, *328*(5982), 1164. https://doi.org/10.1126/science.1187512

Casas, L., Pearman, J. K., & Irigoien, X. (2017). Metabarcoding reveals seasonal and temperature-dependent succession of zooplankton communities in the red sea. *Frontiers in Marine Science*, *4*, https://doi.org/10.3389/fmars.2017.00241

Casey, J. M., Meyer, C. P., Morat, F., Brandl, S. J., Planes, S., & Parravicini, V. (2019). Reconstructing hyperdiverse food webs: Gut content metabarcoding as a tool to disentangle trophic interactions on coral reefs. *Methods in Ecology and Evolution*, *10*(8), 1157–1170. https://doi.org/10.1111/2041-210X.13206

Chadwick, N. E., & Morrow, K. M. (2011). Competition among sessile organisms on coral reefs. In Z. Dubinsky, & N. Stambler (Eds.), *Coral reefs: An ecosystem in transition* (pp. 347–371). Springer Netherlands. https://doi.org/10.1007/978-94-007-0114-4_20

Chain, F. J. J., Brown, E. A., MacIsaac, H. J., & Cristescu, M. E. (2016). Metabarcoding reveals strong spatial structure and temporal turnover of zooplankton communities among marine and freshwater ports. *Diversity and Distributions*, *22*(5), 493–504. https://doi.org/10.1111/ddi.12427

Chao, A., Gotelli, N. J., Hsieh, T. C., Sander, E. L., Ma, K. H., Colwell, R. K., & Ellison, A. M. (2014). Rarefaction and extrapolation with Hill numbers: A framework for sampling and estimation in species diversity studies. *Ecological Monographs*, *84*(1), 45–67. https://doi.org/10.1890/13-0133.1

Clarke, L. J., Beard, J. M., Swadling, K. M., & Deagle, B. E. (2017). Effect of marker choice and thermal cycling protocol on zooplankton DNA metabarcoding studies. *Ecology and Evolution*, *7*(3), 873–883. https://doi.org/10.1002/ece3.2667

Clarke, L. J., Soubrier, J., Weyrich, L. S., & Cooper, A. (2014). Environmental metabarcodes for insects: In silico PCR reveals potential for

taxonomic bias. *Molecular Ecology Resources*, *14*(6), 1160–1170. https://doi.org/10.1111/1755-0998.12265

Cowart, D. A., Pinheiro, M., Mouchel, O., Maguer, M., Grall, J., Miné, J., & Arnaud-Haond, S. (2015). Metabarcoding is powerful yet still blind: a comparative analysis of morphological and molecular surveys of seagrass communities. *PLoS One*, *10*(2), e0117562. https://doi.org/10.1371/journal.pone.0117562

Creer, S., Fonseca, V. G., Porazinska, D. L., Giblin-Davis, R. M., Sung, W., Power, D. M., Packer, M., Carvalho, G. R., Blaxter, M. L., Lambshead, P. J. D., & Thomas, W. K. (2010). Ultrasequencing of the meiofaunal biosphere: Practice, pitfalls and promises. *Molecular Ecology*, *19*(s1), 4–20. https://doi.org/10.1111/j.1365-294X.2009.04473.x

Cristescu, M. E. (2014). From barcoding single individuals to metabarcoding biological communities: Towards an integrative approach to the study of global biodiversity. *Trends in Ecology & Evolution*, *29*(10), 566–571. https://doi.org/10.1016/j.tree.2014.08.001

David, R., Uyarra, M. C., Carvalho, S., Anlauf, H., Borja, A., Cahill, A. E., Carugati, L., Danovaro, R., De Jode, A., Feral, J.-P., Guillemain, D., Martire, M. L., D'Avray, L. T. D. V., Pearman, J. K., & Chenuil, A. (2019). Lessons from photo analyses of Autonomous Reef Monitoring Structures as tools to detect (bio-)geographical, spatial, and environmental effects. *Marine Pollution Bulletin*, *141*, 420–429. https://doi.org/10.1016/j.marpolbul.2019.02.066

de Vargas, C., Audic, S., Henry, N., Decelle, J., Mahé, F., Logares, R., Lara, E., Berney, C., Le Bescot, N., Probert, I., Carmichael, M., Poulain, J., Romac, S., Colin, S., Aury, J.-M., Bittner, L., Chaffron, S., Dunthorn, M., Engelen, S., … Karsenti, E. (2015). Eukaryotic plankton diversity in the sunlit ocean. *Science*, *348*(6237), 1261605. https://doi.org/10.1126/science.1261605

Deagle, B. E., Jarman, S. N., Coissac, E., Pompanon, F., & Taberlet, P. (2014). DNA metabarcoding and the cytochrome c oxidase subunit I marker: Not a perfect match. *Biology Letters*, *10*(9), 20140562. https://doi.org/10.1098/rsbl.2014.0562

Deagle, B. E., Thomas, A. C., McInnes, J. C., Clarke, L. J., Vesterinen, E. J., Clare, E. L., Kartzinel, T. R., & Eveson, J. P. (2019). Counting with DNA in metabarcoding studies: How should we convert sequence reads to dietary data? *Molecular Ecology*, *28*(2), 391–406. https://doi.org/10.1111/mec.14734

Deiner, K., Walser, J.-C., Mächler, E., & Altermatt, F. (2015). Choice of capture and extraction methods affect detection of freshwater biodiversity from environmental DNA. *Special Issue: Environmental DNA: A Powerful New Tool for Biological Conservation*, *183*, 53–63. https://doi.org/10.1016/j.biocon.2014.11.018

Díaz, S., Fargione, J., Chapin, F. S. III, & Tilman, D. (2006). Biodiversity loss threatens human well-being. *PLOS Biology*, *4*(8), e277. https://doi.org/10.1371/journal.pbio.0040277

DiBattista, J. D., Reimer, J. D., Stat, M., Masucci, G. D., Biondi, P., De Brauwer, M., Wilkinson, S. P., Chariton, A. A., & Bunce, M. (2020). Environmental DNA can act as a biodiversity barometer of anthropogenic pressures in coastal ecosystems. *Scientific Reports*, *10*(1), https://doi.org/10.1038/s41598-020-64858-9

Dodt, M., Roehr, J., Ahmed, R., & Dieterich, C. (2012). FLEXBAR—flexible barcode and adapter processing for next-generation sequencing platforms. *Biology*, *1*(3), 895–905. https://doi.org/10.3390/biology1030895

Edgar, R. C. (2010). Search and clustering orders of magnitude faster than BLAST. *Bioinformatics*, *26*(19), 2460–2461. https://doi.org/10.1093/bioinformatics/btq461

Elahi, R., O'Connor, M. I., Byrnes, J. E. K., Dunic, J., Eriksson, B. K., Hensel, M. J. S., & Kearns, P. J. (2015). Recent trends in local-scale marine biodiversity reflect community structure and human impacts. *Current Biology*, *25*(14), 1938–1943. https://doi.org/10.1016/j.cub.2015.05.030

Elbrecht, V., & Leese, F. (2015). Can DNA-based ecosystem assessments quantify species abundance? Testing primer bias and biomass—sequence relationships with an innovative metabarcoding protocol. *PLoS One*, *10*(7), e0130324. https://doi.org/10.1371/journal.pone.0130324

Elbrecht, V., Peinert, B., & Leese, F. (2017). Sorting things out: Assessing effects of unequal specimen biomass on DNA metabarcoding. *Ecology and Evolution*, *7*(17), 6918–6926. https://doi.org/10.1002/ece3.3192

Elbrecht, V., Vamos, E. E., Meissner, K., Aroviita, J., & Leese, F. (2017). Assessing strengths and weaknesses of DNA metabarcoding-based macroinvertebrate identification for routine stream monitoring. *Methods in Ecology and Evolution*, *8*(10), 1265–1275. https://doi.org/10.1111/2041-210X.12789

Epstein, S., & López-García, P. (2008). "Missing" protists: A molecular prospective. *Biodiversity and Conservation*, *17*(2), 261–276. https://doi.org/10.1007/s10531-007-9250-y

Erpenbeck, D., & Wörheide, G. (2007). On the molecular phylogeny of sponges (Porifera)*. *Zootaxa*, *1668*(1), 107–126. https://doi.org/10.11646/zootaxa.1668.1.10

Ficetola, G. F., Miaud, C., Pompanon, F., & Taberlet, P. (2008). Species detection using environmental DNA from water samples. *Biology Letters*, *4*(4), 423–425. https://doi.org/10.1098/rsbl.2008.0118

Field, K., Olsen, G., Lane, D., Giovannoni, S., Ghiselin, M., Raff, E., Pace, N., & Raff, R. (1988). Molecular phylogeny of the animal kingdom. *Science*, *239*(4841), 748–753. https://doi.org/10.1126/science.3277277

Flynn, J. M., Brown, E. A., Chain, F. J. J., MacIsaac, H. J., & Cristescu, M. E. (2015). Toward accurate molecular identification of species in complex environmental samples: Testing the performance of sequence filtering and clustering methods. *Ecology and Evolution*, *5*(11), 2252–2266. https://doi.org/10.1002/ece3.1497

Fonseca, V. G., Carvalho, G. R., Sung, W., Johnson, H. F., Power, D. M., Neill, S. P., Packer, M., Blaxter, M. L., Lambshead, P. J. D., Thomas, W. K., & Creer, S. (2010). Second-generation environmental sequencing unmasks marine metazoan biodiversity. *Nature Communications*, *1*(1), 98. https://doi.org/10.1038/ncomms1095

Fuchs, J., Obst, M., & Sundberg, P. (2009). The first comprehensive molecular phylogeny of Bryozoa (Ectoprocta) based on combined analyses of nuclear and mitochondrial genes. *Molecular Phylogenetics and Evolution*, *52*(1), 225–233. https://doi.org/10.1016/j.ympev.2009.01.021

Geller, J., Meyer, C., Parker, M., & Hawk, H. (2013). Redesign of PCR primers for mitochondrial cytochrome c oxidase subunit I for marine invertebrates and application in all-taxa biotic surveys. *Molecular Ecology Resources*, *13*(5), 851–861. https://doi.org/10.1111/1755-0998.12138

Gelman, A., Carlin, J., Stern, H., Dunson, D., Vehtari, A., & Rubin, D. (2013). *Bayesian data analysis*. Chapman and Hall/CRC.

Gibson, J., Shokralla, S., Porter, T. M., King, I., van Konynenburg, S., Janzen, D. H., Hallwachs, W., & Hajibabaei, M. (2014). Simultaneous assessment of the macrobiome and microbiome in a bulk sample of tropical arthropods through DNA metasystematics. *Proceedings of the National Academy of Sciences*, *111*(22), 8007. https://doi.org/10.1073/pnas.1406468111

Gold, Z., Choi, E., Kacev, D., Frable, B., Burton, R., Goodwin, K., Thompson, A., & Barber, P. (2020). FishCARD: Fish 12S California current specific reference database for enhanced metabarcoding efforts [Preprint]. Preprints. https://doi.org/10.22541/au.159136805.55528691

Guillou, L., Bachar, D., Audic, S., Bass, D., Berney, C., Bittner, L., Boutte, C., Burgaud, G., de Vargas, C., Decelle, J., Del Campo, J., Dolan, J. R., Dunthorn, M., Edvardsen, B., Holzmann, M., Kooistra, W. H. C. F., Lara, E., Le Bescot, N., Logares, R., … Christen, R. (2013). The Protist Ribosomal Reference database (PR2): A catalog of unicellular eukaryote small sub-unit rRNA sequences with curated taxonomy. *Nucleic Acids Research*, *41*(Database issue), D597–D604. https://doi.org/10.1093/nar/gks1160

Günther, B., Knebelsberger, T., Neumann, H., Laakmann, S., & Martínez Arbizu, P. (2018). Metabarcoding of marine environmental DNA

based on mitochondrial and nuclear genes. *Scientific Reports*, *8*(1), https://doi.org/10.1038/s41598-018-32917-x

Hadziavdic, K., Lekang, K., Lanzen, A., Jonassen, I., Thompson, E. M., & Troedsson, C. (2014). Characterization of the 18S rRNA gene for designing universal eukaryote specific primers. *PLoS One*, *9*(2), e87624. https://doi.org/10.1371/journal.pone.0087624

Hajibabaei, M., Shokralla, S., Zhou, X., Singer, G. A. C., & Baird, D. J. (2011). Environmental barcoding: A next-generation sequencing approach for biomonitoring applications using river benthos. *PLoS One*, *6*(4), e17497. https://doi.org/10.1371/journal.pone.0017497

Hebert, P. D. N., Cywinska, A., Ball, S. L., & deWaard, J. R. (2003). Biological identifications through DNA barcodes. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, *270*(1512), 313–321. https://doi.org/10.1098/rspb.2002.2218

Hebert, P. D. N., Ratnasingham, S., & de Waard, J. R. (2003). Barcoding animal life: Cytochrome c oxidase subunit 1 divergences among closely related species. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, *270*(suppl_1). https://doi.org/10.1098/rsbl.2003.0025

Holland, L. Z. (2016). Tunicates. *Current Biology*, *26*(4), R146–R152. https://doi.org/10.1016/j.cub.2015.12.024

Huang, D., Meier, R., Todd, P. A., & Chou, L. M. (2008). Slow mitochondrial COI sequence evolution at the base of the metazoan tree and its implications for DNA barcoding. *Journal of Molecular Evolution*, *66*(2), 167–174. https://doi.org/10.1007/s00239-008-9069-5

Ji, Y., Ashton, L., Pedley, S. M., Edwards, D. P., Tang, Y., Nakamura, A., Kitching, R., Dolman, P. M., Woodcock, P., Edwards, F. A., Larsen, T. H., Hsu, W. W., Benedick, S., Hamer, K. C., Wilcove, D. S., Bruce, C., Wang, X., Levi, T., Lott, M., … Yu, D. W. (2013). Reliable, verifiable and efficient monitoring of biodiversity via metabarcoding. *Ecology Letters*, *16*(10), 1245–1257. https://doi.org/10.1111/ele.12162

Kahle, D., & Wickham, H. (2013). ggmap: Spatial Visualization with ggplot2. *The R Journal*, *5*, 144–161.

Kay, M. (2019). tidybayes: Tidy Data and Geoms for Bayesian Models (R package version 1.1.0) [Computer software]. https://doi.org/10.5281/zenodo.1308151

Kopf, A., Bicak, M., Kottmann, R., Schnetzer, J., Kostadinov, I., Lehmann, K., Fernandez-Guerra, A., Jeanthon, C., Rahav, E., Ullrich, M., Wichels, A., Gerdts, G., Polymenakou, P., Kotoulas, G., Siam, R., Abdallah, R. Z., Sonnenschein, E. C., Cariou, T., O'Gara, F., … Glöckner, F. O. (2015). The ocean sampling day consortium. *GigaScience*, *4*, 27. https://doi.org/10.1186/s13742-015-0066-5

Leasi, F., Gaynus, C., Mahardini, A., Moore, T. N., Norenburg, J. L., & Barber, P. H. (2016). Spatial and ecologic distribution of neglected microinvertebrate communities across endangered ecosystems: Meiofauna in Bali (Indonesia). *Marine Ecology*, *37*(5), 970–987. https://doi.org/10.1111/maec.12305

Leray, M., Boehm, J. T., Mills, S. C., & Meyer, C. P. (2012). Moorea BIOCODE barcode library as a tool for understanding predator–prey interactions: Insights into the diet of common predatory coral reef fishes. *Coral Reefs*, *31*(2), 383–388. https://doi.org/10.1007/s00338-011-0845-0

Leray, M., & Knowlton, N. (2015). DNA barcoding and metabarcoding of standardized samples reveal patterns of marine benthic diversity. *Proceedings of the National Academy of Sciences*, *112*(7), 2076–2081. https://doi.org/10.1073/pnas.1424997112

Leray, M., & Knowlton, N. (2016). Censusing marine eukaryotic diversity in the twenty-first century. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *371*(1702), 20150331. https://doi.org/10.1098/rstb.2015.0331

Leray, M., & Knowlton, N. (2017). Random sampling causes the low reproducibility of rare eukaryotic OTUs in Illumina COI metabarcoding. *PeerJ*, *5*, e3006. https://doi.org/10.7717/peerj.3006

Leray, M., Yang, J. Y., Meyer, C. P., Mills, S. C., Agudelo, N., Ranwez, V., Boehm, J. T., & Machida, R. J. (2013). A new versatile primer set targeting a short fragment of the mitochondrial COI region for metabarcoding metazoan diversity: Application for characterizing coral reef fish gut contents. *Frontiers in Zoology*, *10*(1), 34. https://doi.org/10.1186/1742-9994-10-34

López-García, P., Rodríguez-Valera, F., Pedrós-Alió, C., & Moreira, D. (2001). Unexpected diversity of small eukaryotes in deep-sea Antarctic plankton. *Nature*, *409*(6820), 603–607. https://doi.org/10.1038/35054537

McCauley, D. J., Pinsky, M. L., Palumbi, S. R., Estes, J. A., Joyce, F. H., & Warner, R. R. (2015). Marine defaunation: Animal loss in the global ocean. *Science*, *347*(6219), 1255641. https://doi.org/10.1126/science.1255641

Menge, B. A., Ashkenas, L. R., & Matson, A. (1983). Use of artificial holes in studying community development in cryptic marine habitats in a tropical rocky intertidal region. *Marine Biology*, *77*(2), 129–142. https://doi.org/10.1007/BF00396310

Meyer, A., Todt, C., Mikkelsen, N. T., & Lieb, B. (2010). Fast evolving 18S rRNA sequences from Solenogastres (Mollusca) resist standard PCR amplification and give new insights into mollusk substitution rate heterogeneity. *BMC Evolutionary Biology*, *10*(1), 70. https://doi.org/10.1186/1471-2148-10-70

Meyer, C. P. (2016). Moorea Biocode Project FASTA data. California Digital Library. http://n2t.net/ark:/13030/m5478zfg

Moon-van der Staay, S. Y., De Wachter, R., & Vaulot, D. (2001). Oceanic 18S rDNA sequences from picoplankton reveal unsuspected eukaryotic diversity. *Nature*, *409*(6820), 607–610. https://doi.org/10.1038/35054541

Oksanen, J., Blanchet, F. G., Friendly, M., Kindt, R., Legendre, P., McGlinn, D., Minchin, P. R., O'Hara, R. B., Simpson, G. L., Solymos, P., Stevens, M. H. H., Szoecs, E., & Wagner, H. (2018). vegan: Community Ecology Package. (R package version 2.5-2.) [Computer software]. https://CRAN.R-project.org/package=vegan

Pearman, J. K., Aylagas, E., Voolstra, C. R., Anlauf, H., Villalobos, R., & Carvalho, S. (2019). Disentangling the complex microbial community of coral reefs using standardized Autonomous Reef Monitoring Structures (ARMS). *Molecular Ecology*, *28*(15), 3496–3507. https://doi.org/10.1111/mec.15167.

Pearman, J. K., El-Sherbiny, M. M., Lanzén, A., Al-Aidaroos, A. M., & Irigoien, X. (2014). Zooplankton diversity across three Red Sea reefs using pyrosequencing. *Frontiers in Marine Science*, *1*, 27. https://doi.org/10.3389/fmars.2014.00027

Pearman, J. K., & Irigoien, X. (2015). Assessment of Zooplankton Community Composition along a Depth Profile in the Central Red Sea. *PLoS One*, *10*(7), e0133487. https://doi.org/10.1371/journal.pone.0133487

Pearman, J. K., Leray, M., Villalobos, R., Machida, R. J., Berumen, M. L., Knowlton, N., & Carvalho, S. (2018). Cross-shelf investigation of coral reef cryptic benthic organisms reveals diversity patterns of the hidden majority. *Scientific Reports*, *8*(1), https://doi.org/10.1038/s41598-018-26332-5

Piganeau, G., Eyre-Walker, A., Grimsley, N., & Moreau, H. (2011). How and why DNA barcodes underestimate the diversity of microbial eukaryotes. *PLoS One*, *6*(2), e16342. https://doi.org/10.1371/journal.pone.0016342

Piñol, J., Senar, M. A., & Symondson, W. O. C. (2019). The choice of universal primers and the characteristics of the species mixture determine when DNA metabarcoding can be quantitative. *Molecular Ecology*, *28*(2), 407–419. https://doi.org/10.1111/mec.14776

Piredda, R., Tomasino, M. P., D'Erchia, A. M., Manzari, C., Pesole, G., Montresor, M., Kooistra, W. H. C. F., Sarno, D., & Zingone, A. (2017). Diversity and temporal patterns of planktonic protist assemblages at a Mediterranean Long Term Ecological Research site. *FEMS*

Microbiology Ecology, *93*(fiw200), https://doi.org/10.1093/femsec/fiw200

Plaisance, L., Caley, M. J., Brainard, R. E., & Knowlton, N. (2011). The diversity of coral reefs: What are we missing? *PLoS One*, *6*(10), e25026. https://doi.org/10.1371/journal.pone.0025026

Pollock, J., Glendinning, L., Wisedchanwet, T., & Watson, M. (2018). The madness of microbiome: Attempting to find consensus "best practice" for 16s microbiome studies. *Applied and Environmental Microbiology*, *84*(7), e02627–e2717. https://doi.org/10.1128/AEM.02627-17

Porter, T. M., & Hajibabaei, M. (2018). Over 2.5 million COI sequences in GenBank and growing. *PLoS One*, *13*(9), e0200177. https://doi.org/10.1371/journal.pone.0200177

Prosser, S. W. J., Velarde-Aguilar, M. G., León-Règagnon, V., & Hebert, P. D. N. (2013). Advancing nematode barcoding: A primer cocktail for the cytochrome c oxidase subunit I gene from vertebrate parasitic nematodes. *Molecular Ecology Resources*, *13*(6), 1108–1115. https://doi.org/10.1111/1755-0998.12082

Pruesse, E., Quast, C., Knittel, K., Fuchs, B. M., Ludwig, W., Peplies, J., & Glöckner, F. O. (2007). SILVA: a comprehensive online resource for quality checked and aligned ribosomal RNA sequence data compatible with ARB. *Nucleic Acids Research*, *35*(21), 7188–7196. https://doi.org/10.1093/nar/gkm864

R Core Team (2019). R: A language and environment for statistical computing. R Foundation for Statistical Computing. https://www.R-project.org/

Ransome, E., Geller, J. B., Timmers, M., Leray, M., Mahardini, A., Sembiring, A., Collins, A. G., & Meyer, C. P. (2017). The importance of standardization for biodiversity comparisons: A case study using autonomous reef monitoring structures (ARMS) and metabarcoding to measure cryptic diversity on Mo'orea coral reefs, French Polynesia. *PLoS One*, *12*(4), e0175066. https://doi.org/10.1371/journal.pone.0175066

Ranwez, V., Harispe, S., Delsuc, F., & Douzery, E. J. P. (2011). MACSE: Multiple alignment of coding sequences accounting for frameshifts and stop codons. *PLoS One*, *6*(9), e22594. https://doi.org/10.1371/journal.pone.0022594

Ruggiero, M. A., Gordon, D. P., Orrell, T. M., Bailly, N., Bourgoin, T., Brusca, R. C., Cavalier-Smith, T., Guiry, M. D., & Kirk, P. M. (2015). A higher level classification of all living organisms. *PLoS One*, *10*(4), e0119248. https://doi.org/10.1371/journal.pone.0119248

Salonen, I. S., Chronopoulou, P.-M., Leskinen, E., & Koho, K. A. (2019). Metabarcoding successfully tracks temporal changes in eukaryotic communities in coastal sediments. *FEMS Microbiology Ecology*, *95*(1), https://doi.org/10.1093/femsec/fiy226

Shearer, T. L., van Oppen, M. J. H., Romano, S. L., & Wörheide, G. (2002). Slow mitochondrial DNA sequence evolution in the Anthozoa (Cnidaria). *Molecular Ecology*, *11*(12), 2475–2487. https://doi.org/10.1046/j.1365-294X.2002.01652.x

Stan Development Team. (2019). RStan: The R interface to Stan (R package version 2.19.2) [Computer software]. http://mc-stan.org/

Stefanni, S., Stanković, D., Borme, D., de Olazabal, A., Juretić, T., Pallavicini, A., & Tirelli, V. (2018). Multi-marker metabarcoding approach to study mesozooplankton at basin scale. *Scientific Reports*, *8*(1), https://doi.org/10.1038/s41598-018-30157-7

Taberlet, P., Coissac, E., Pompanon, F., Brochmann, C., & Willerslev, E. (2012). Towards next-generation biodiversity assessment using DNA metabarcoding. *Molecular Ecology*, *21*(8), 2045–2050. https://doi.org/10.1111/j.1365-294X.2012.05470.x

Tang, C. Q., Leasi, F., Obertegger, U., Kieneke, A., Barraclough, T. G., & Fontaneto, D. (2012). The widely used small subunit 18S rDNA molecule greatly underestimates true diversity in biodiversity surveys of the meiofauna. *Proceedings of the National Academy of Sciences*, *109*(40), 16208. https://doi.org/10.1073/pnas.1209160109

Tragin, M., Zingone, A., & Vaulot, D. (2018). Comparison of coastal phytoplankton composition estimated from the V4 and V9 regions of the 18S rRNA gene with a focus on photosynthetic groups and especially Chlorophyta: Comparison of V4 and V9 18S metabarcodes. *Environmental Microbiology*, *20*(2), 506–520. https://doi.org/10.1111/1462-2920.13952

van der Loos, L. M., & Nijland, R. (2020). Biases in bulk: DNA metabarcoding of marine communities and the methodology involved. *Molecular Ecology*, https://doi.org/10.1111/mec.15592

Voigt, O., & Wörheide, G. (2016). A short LSU rRNA fragment as a standard marker for integrative taxonomy in calcareous sponges (Porifera: Calcarea). *Organisms Diversity & Evolution*, *16*(1), 53–64. https://doi.org/10.1007/s13127-015-0247-1

Wangensteen, O. S., Palacín, C., Guardiola, M., & Turon, X. (2018). DNA metabarcoding of littoral hard-bottom communities: High diversity and database gaps revealed by two molecular markers. *PeerJ*, *6*, e4705. https://doi.org/10.7717/peerj.4705

West, K. M., Stat, M., Harvey, E. S., Skepper, C. L., DiBattista, J. D., Richards, Z. T., Travers, M. J., Newman, S. J., & Bunce, M. (2020). EDNA metabarcoding survey reveals fine-scale coral reef community variation across a remote, tropical island ecosystem. *Molecular Ecology*, *29*(6), 1069–1086. https://doi.org/10.1111/mec.15382

Wickham, H. (2016). *ggplot2: Elegant graphics for data analysis*. Springer-Verlag.

Wickham, H. (2017). tidyverse: Easily Install and Load the "Tidyverse". (R package version 1.2.1) [Computer software]. https://CRAN.R-project.org/package=tidyverse

Williams, I. D., Couch, C. S., Beijbom, O., Oliver, T. A., Vargas-Angel, B., Schumacher, B. D., & Brainard, R. E. (2019). Leveraging automated image analysis tools to transform our capacity to assess status and trends of coral reefs. *Frontiers in Marine Science*, *6*, 222. https://doi.org/10.3389/fmars.2019.00222

Worm, B., Barbier, E. B., Beaumont, N., Duffy, J. E., Folke, C., Halpern, B. S., Jackson, J. B. C., Lotze, H. K., Micheli, F., Palumbi, S. R., Sala, E., Selkoe, K. A., Stachowicz, J. J., & Watson, R. (2006). Impacts of biodiversity loss on ocean ecosystem services. *Science*, *314*(5800), 787–790. https://doi.org/10.1126/science.1132294

Wu, S., Xiong, J., & Yu, Y. (2015). Taxonomic resolutions based on 18S rRNA genes: A case study of subclass copepoda. *PLoS One*, *10*(6), e0131498. https://doi.org/10.1371/journal.pone.0131498

Zhan, A., Bailey, S. A., Heath, D. D., & MacIsaac, H. J. (2014). Performance comparison of genetic markers for high-throughput sequencing-based biodiversity assessment in complex communities. *Molecular Ecology Resources*. https://doi.org/10.1111/1755-0998.12254

Zhan, A., Hulák, M., Sylvester, F., Huang, X., Adebayo, A. A., Abbott, C. L., Adamowicz, S. J., Heath, D. D., Cristescu, M. E., & MacIsaac, H. J. (2013). High sensitivity of 454 pyrosequencing for detection of rare species in aquatic communities. *Methods in Ecology and Evolution*, *4*(6), 558–565. https://doi.org/10.1111/2041-210X.12037

Zhang, G. K., Chain, F. J. J., Abbott, C. L., & Cristescu, M. E. (2018). Metabarcoding using multiplexed markers increases species detection in complex zooplankton communities. *Evolutionary Applications*, *11*(10), 1901–1914. https://doi.org/10.1111/eva.12694

Zhang, J., Kobert, K., Flouri, T., & Stamatakis, A. (2013). PEAR: A fast and accurate Illumina Paired-End reAd mergeR. *Bioinformatics*, *30*(5), 614–620. https://doi.org/10.1093/bioinformatics/btt593

Zimmerman, T. L., & Martin, J. W. (2004). Artificial reef matrix structures (arms): An inexpensive and effective method for collecting coral reef-associated invertebrates. *Gulf and Caribbean Research*, *16*. https://doi.org/10.18785/gcr.1601.08

Zinger, L., Bonin, A., Alsos, I. G., Bálint, M., Bik, H., Boyer, F., Chariton, A. A., Creer, S., Coissac, E., Deagle, B. E., De Barba, M., Dickie, I. A., Dumbrell, A. J., Ficetola, G. F., Fierer, N., Fumagalli, L., Gilbert, M. T. P., Jarman, S., Jumpponen, A., … Taberlet, P. (2019). DNA

metabarcoding—Need for robust experimental designs to draw sound ecological conclusions. *Molecular Ecology*. https://doi.org/10.1111/mec.15060

Zizka, V. M. A., Elbrecht, V., Macher, J., & Leese, F. (2019). Assessing the influence of sample tagging and library preparation on DNA metabarcoding. *Molecular Ecology Resources*. https://doi.org/10.1111/1755-0998.13018

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.

How to cite this article: Casey, J. M., Ransome, E., Collins, A. G., Mahardini, A., Kurniasih, E. M., Sembiring, A., Schiettekatte, N. M. D., Cahyani, N. K. D., Wahyu Anggoro, A., Moore, M., Uehling, A., Belcaid, M., Barber, P. H., Geller, J. B., & Meyer, C. P. (2021). DNA metabarcoding marker choice skews perception of marine eukaryotic biodiversity. *Environmental DNA*, 3, 1229–1246. https://doi.org/10.1002/edn3.245