

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/361403653>

Seq' and ARMS shall find: DNA (meta)barcoding of Autonomous Reef Monitoring Structures across the tree of life uncovers hidden cryptobiome of tropical urban coral reefs

Article in Molecular Ecology · December 2023

DOI: 10.1111/mec.16568

CITATIONS

15

7 authors, including:



Yin Cheong Aden Ip
National Parks Board
21 PUBLICATIONS 321 CITATIONS

[SEE PROFILE](#)

READS

276



Jia Jin Marc Chang
National University of Singapore
19 PUBLICATIONS 236 CITATIONS

[SEE PROFILE](#)



Randolph Quek
DSO National Laboratories
31 PUBLICATIONS 323 CITATIONS

[SEE PROFILE](#)



Andrew G. Bauman
Nova Southeastern University
94 PUBLICATIONS 3,363 CITATIONS

[SEE PROFILE](#)

Seq' and ARMS shall find: DNA (meta)barcoding of Autonomous Reef Monitoring Structures across the tree of life uncovers hidden cryptobiome of tropical urban coral reefs

Yin Cheong Aden Ip¹  | Jia Jin Marc Chang¹  | Ren Min Oh¹  |
 Zheng Bin Randolph Quek^{1,2}  | Yong Kit Samuel Chan¹  |
 Andrew G. Bauman^{1,3}  | Danwei Huang^{1,4,5,6} 

¹Department of Biological Sciences,
 National University of Singapore,
 Singapore, Singapore

²Yale-NUS College, National University of
 Singapore, Singapore, Singapore

³Department of Marine and
 Environmental Sciences, Nova
 Southeastern University, Dania Beach,
 Florida, USA

⁴Centre for Nature-Based Climate
 Solutions, National University of
 Singapore, Singapore, Singapore

⁵Lee Kong Chian Natural History Museum,
 National University of Singapore,
 Singapore, Singapore

⁶Tropical Marine Science Institute,
 National University of Singapore,
 Singapore, Singapore

Correspondence

Yin Cheong Aden Ip, Department of
 Biological Sciences, National University of
 Singapore, 16 Science Drive 4, Singapore
 117558, Singapore.

Email: ycip@u.nus.edu

Funding information

AXA Postdoctoral Fellowship, Grant/
 Award Number: R-154-000-649-507;
 National Research Foundation Singapore,
 Grant/Award Number: MSRDP-P03;
 Singapore Ministry of Education Academic
 Research Fund Tier 1, Grant/Award
 Number: R-154-000-A63-114

Handling Editor: Henrik Krehenwinkel

Abstract

Coral reefs are among the richest marine ecosystems on Earth, but there remains much diversity hidden within cavities of complex reef structures awaiting discovery. While the abundance of corals and other macroinvertebrates are known to influence the diversity of other reef-associated organisms, much remains unknown on the drivers of cryptobenthic diversity. A combination of standardized sampling with 12 units of the Autonomous Reef Monitoring Structure (ARMS) and high-throughput sequencing was utilized to uncover reef cryptobiome diversity across the equatorial reefs in Singapore. DNA barcoding and metabarcoding of mitochondrial cytochrome c oxidase subunit I, nuclear 18S and bacterial 16S rRNA genes revealed the taxonomic composition of the reef cryptobiome, comprising 15,356 microbial ASVs from over 50 bacterial phyla, and 971 MOTUs across 15 metazoan and 19 non-metazoan eukaryote phyla. Environmental factors across different sites were tested for relationships with ARMS diversity. Differences among reefs in diversity patterns of metazoans and other eukaryotes, but not microbial communities, were associated with biotic (coral cover) and abiotic (distance, temperature and sediment) environmental variables. In particular, ARMS deployed at reefs with higher coral cover had greater metazoan diversity and encrusting plate cover, with larger-sized non-coral invertebrates influencing spatial patterns among sites. Our study showed that DNA barcoding and metabarcoding of ARMS constitute a valuable tool for quantifying cryptobenthic diversity patterns and can provide critical information for the effective management of coral reef ecosystems.

KEY WORDS

cryptobenthic diversity, eukaryotes, high-throughput sequencing, marine biomonitoring, metazoans, microbial

1 | INTRODUCTION

Coral reefs are some of the richest ecosystems on Earth (Fisher et al., 2011; Knowlton, 2001a, 2001b) and are home to more than a quarter of all marine eukaryotic species (Albano et al., 2011; Bouchet et al., 2009). The Central Indo-Pacific, for example, contains the highest species richness of multiple reef taxa (Roberts et al., 2002), including corals, molluscs and fish (Bellwood & Meyer, 2009; Hoeksema, 2007; Huang et al., 2018), making it a renowned marine biodiversity hotspot region. Such extraordinary biodiversity is attributed to the coral reef's structural complexity (Graham & Nash, 2013), for which stony corals (Cnidaria: Scleractinia) play an integral role. Individually, these sessile marine animal colonies produce calcium carbonate skeletons. Collectively, high coral cover comprising a wide variety of growth forms creates habitat complexity on the reef (Graham & Nash, 2013; Kovalenko et al., 2012) through the formation of cavities, crevices and additional area for organismal settlement and refugia (Fabricius et al., 2014; Idjadi & Edmunds, 2006; Rogers et al., 2018). These complex structural habitats support diverse, abundant and productive reef-associated organisms (Chong-Seng et al., 2012; Fabricius et al., 2014; Idjadi & Edmunds, 2006; Komyakova et al., 2013), encompassing a wide range of taxa from macrofauna to microbes (Garren & Azam, 2012; Glynn & Enochs, 2011; Kelly et al., 2014; Plaisance et al., 2011). Notably, the diversity of conspicuous taxa tends to covary in space and with environmental conditions, where these relationships have been rigorously examined with visually-reliant census methods (Danovaro et al., 2016; Idjadi & Edmunds, 2006; Komyakova et al., 2013).

However, most survey studies have typically focused on conspicuous reef taxa (e.g., corals, fish and macroalgae; Huang et al., 2009; Leray et al., 2012; Wells et al., 2019), with more attention paid to microbes only in recent years (Afiq-Rosli et al., 2021; Oh et al., 2021; Pearman et al., 2019; Wainwright, Bauman et al., 2019). The paucity of empirical data on drivers of reef cryptobiome diversity—comprising organisms colonizing hidden areas within the reef structural complex—is evident (Carvalho et al., 2019). Although positive relationships between the physical heterogeneity of reef ecosystems, invertebrate cryptofaunal diversity and organism body sizes have been demonstrated (Fraser et al., 2021; Valencia & Giraldo, 2021), such patterns have mostly been inferred from habitat-degradation experiments conducted on dead corals and rubble that are relatively easy to sample or collect with invasive methods (Biondi et al., 2020; Enochs & Manzello, 2012; Fraser et al., 2021; Valencia & Giraldo, 2021).

To date, only a few studies have documented the association of microeukaryotic diversity gradients with environmental factors across continental shelves (Pearman et al., 2018), and also at small spatial scales within tropical and naturally-acidified reef systems in the Coral Triangle (Casey et al., 2021; Plaisance et al., 2021). This is not surprising given that visual reef surveys would overlook nocturnal, small and hidden organisms (Brandl et al., 2018, 2019; Enochs et al., 2011; Pearman et al., 2016), and sampling the reef cryptobiome can be destructive and hence not preferred. This is further

exacerbated by difficulties in species identification, and skilled taxonomic expertise for accurate identification is often needed but is also costly (i.e., manpower hours), curtailing its utility for large-scale studies. Yet, the reef cryptobiota makes up more than 60% of the biomass on coral reefs (Glynn & Enochs, 2011; Plaisance et al., 2011). These cryptic but hyperdiverse taxa are functionally important (Peixoto et al., 2017), contribute to the persistence of coral reef ecosystems (Gibson et al., 2011), and are critical for the health and resilience of coral reef ecosystems (Charpy et al., 2012; Fraser et al., 2021; Glynn & Enochs, 2011; Wang et al., 2012). For example, they are responsible for nutrient and carbon cycling, nitrogen fixation, detrital pathways, and are integral components of reef productivity and food sources (Charpy et al., 2012; Glynn & Enochs, 2011; Jones et al., 2015; Kramer et al., 2017; Wang et al., 2012). Fundamentally, we lack understanding on whether factors like coral cover, which influence the diversity of conspicuous taxa (Biondi et al., 2020; Fraser et al., 2021; Valencia & Giraldo, 2021), are also relevant to the structuring of cryptobenthic communities.

In this study, we combined standardized sampling using Autonomous Reef Monitoring Structures (ARMS) with high-throughput sequencing to test whether the well-known relationship between coral cover and reef-associated fauna also drives cryptobenthic diversity patterns. Given that the majority of the cryptofauna are specialized to live in cracks, crevices, and hidden spaces within the reef structural complex (Carvalho et al., 2019; Fraser et al., 2021; Rogers et al., 2018), we hypothesised that spatial patterns of cryptobenthic diversity based on standardized sampling and coral and invertebrate cover would be congruent—ARMS deployed at sites with higher coral cover will recruit or be encrusted with a greater diversity of organisms than sites with lower coral cover. Furthermore, we also hypothesised that abiotic environmental variables (distance from mainland, sea surface temperature and total suspended solids) were important in structuring the cryptobenthic communities (Jones et al., 2015; Kling et al., 2020; Rusch et al., 2009). We tested these hypotheses by deploying ARMS units at four reef sites, applying DNA barcoding and metabarcoding to characterize eukaryotic and bacterial ARMS-recruited organisms, and analysing these biotic patterns alongside abiotic parameters. An improved understanding of these cryptobenthic communities can provide insights into the underlying biodiversity and biogeochemical processes of coral reef ecosystems (Jones et al., 2015; Rusch et al., 2009; Wang et al., 2012).

2 | MATERIALS AND METHODS

2.1 | ARMS deployment, retrieval and processing

The Autonomous Reef Monitoring Structure (ARMS) is a cost-effective and manpower efficient long-term sampling device (Leray & Knowlton, 2015) designed to mimic the three-dimensional coral reef matrix by providing finely-structured spaces for colonization and settlement of a variety of reef organisms (Zimmerman & Martin, 2004). The recruited cryptobiota can be processed

with either morphological (David et al., 2019; Palomino-Alvarez et al., 2021) and/or molecular tools (Leray & Knowlton, 2015; Ransome et al., 2017). This device is not without limitations, as it might exclude pioneer species or induce colonization trends that are distinct from those on natural substrates (Pearman et al., 2020). Nevertheless, the key advantages are its relatively nondestructive nature and, more importantly, its broad applicability across multiple habitat types and taxa, permitting the consistent and equitable comparison and quantification of benthic communities across diverse groups (Carvalho et al., 2019; Pearman et al., 2019, 2020; Vicente et al., 2021). Here, triplicate ARMS units were deployed at four reef sites (total 12 units) at depths of 1.7–6.3 m via SCUBA at the Southern Islands of Singapore—Kusu Island (K), Raffles Lighthouse (R), Terumbu Pempang Tengah (T) and Pulau Semakau (S) (Figure 1; see also Appendix S1 for GPS coordinates and depths). All units were anchored into loose coral rubble between live coral colonies, spaced about 10 m apart, and left in situ for approximately 2 years from July 2016 to June/July 2018 (Appendix S1).

The retrieval, transport of units and sample processing followed protocols from Leray and Knowlton (2015). Following retrieval in 2018, the ARMS units were transported back to the laboratory and submerged in UV-sterilized 0.45 µm-filtered seawater for disassembly and sorting into two main fractions (Motile or Sessile), with each fraction further subdivided into two or three subcategories (Table 1). Disassembled plates were first imaged on both sides before ≥2 mm-sized organisms were picked out and vouchered individually (Figure S1). The remaining sessile organisms were scraped off the plates and homogenized with either mortar and pestle or kitchen blenders (i.e., blended sessile section, BSS). Water used to house the ARMS units during disassembly were then filtered through 2 mm,

500 and 106 µm sieves, forming the respectively-named fractions of the motile section (Table 1). Motile specimens retained by the 2 mm sieve were also vouchered individually, while the 500 and 106 µm fractions were similarly homogenized like BSS. All sample types were preserved in 100% ethanol and stored at -80°C until further processing (Ransome et al., 2017). All vouchers were deposited in Lee Kong Chian Natural History Museum's Cryogenic Collection (see Data Availability section, Appendix S1).

2.2 | DNA extraction

DNA extraction protocols used depended on sample processing method (Table 1). Genomic DNA for ≥2 mm voucher specimens (motile and sessile) was extracted individually using the abGenix Animal Tissue Genomic DNA Extraction kit according to the manufacturer's instructions (AITbiotech Pte Ltd; Chang, Ip, Bauman, & Huang, 2020). As the bulk of motile ≥2 mm samples were already barcoded in Chang, Ip, Bauman, and Huang (2020), we only performed re-extractions for 188 samples for which we previously did not obtain DNA barcodes. The homogenized sample fractions (106, 500 µm and BSS) were subjected to cell dislodgement protocol by Duhamel and Jacquet (2006) that involved bead-beater treatment (Omni Bead Ruptor 24) of 2 g (w/w) of each sample immersed in a mix of UV-sterilized 0.45 µm-filtered seawater and Tween80 (Sigma-Aldrich). Thereafter, tissue digestion and DNA isolation with phenol:chloroform:isoamyl-alcohol (25:24:1) followed the Ip et al. (2019) protocol. Blank extraction negative controls for homogenized samples were conducted for each extraction round and followed the same downstream protocols as sample extracts (Ip, Chang, et al., 2021). Consumables used

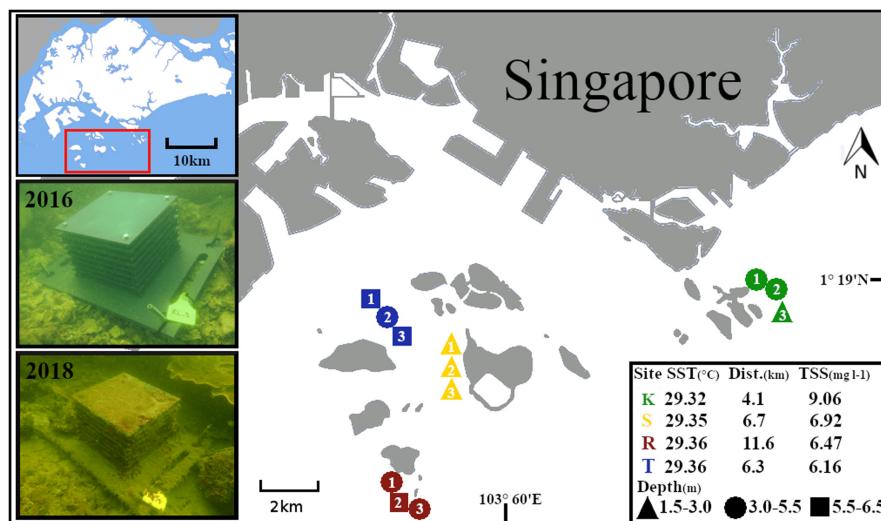


FIGURE 1 Map of 12 ARMS units deployed across four reef sites at the Southern Islands of Singapore (main figure, and top left panel): Kusu Island (K; green), Raffles Lighthouse (R; red), Pulau Semakau (S; yellow), and Terumbu Pempang Tengah (T; blue). Numbers (1–3) indicate the replicate unit deployed at each locality. Photographs showing unencrusted ARMS unit (R3) deployed in July 2016 (middle left panel), and the same unit during retrieval in July 2018 after 2 years of settlement and recruitment with reef organisms (bottom left panel). Environmental data from each locality are shown, with deployment depths indicated by different shapes, while distance from mainland (Dist. [km]), mean sea surface temperature (SST [°C]) and mean total suspended solids (TSS [mg/L]) in 2017 are compiled in the legend (bottom right)

TABLE 1 ARMS fractions analysed in this study and corresponding experimental details. The number of PCR samples and negatives marked with asterisk (*) includes data from Chang, Ip, Bauman, & Huang (2020)

Fraction	Subfraction	No. of samples	DNA extraction	Gene analysed	DNA sequencing
Motile	≥2 mm	725 samples 913 PCR samples* 46 negatives*	abGenix animal tissue DNA extraction kit (Chang, Ip, Bauman, and Huang, 2020)	COI	ONT MinION (FLO-MIN106D)
Motile	500 µm	12 samples 108 PCR samples 45 negatives	Cell dislodgement (Duhamel & Jacquet, 2006) followed by phenol-chloroform extraction (Ip, Tay, et al., 2021)	COI, 16S, 18S	Illumina MiSeq (300-bp PE) Illumina HiSeq2500 (250-bp PE)
Motile	106 µm	12 samples 108 PCR samples 45 negatives	Cell dislodgement (Duhamel & Jacquet, 2006) followed by phenol-chloroform extraction (Ip, Tay, et al., 2021)	COI, 16S, 18S	Illumina MiSeq (300-bp PE) Illumina HiSeq2500 (250-bp PE)
Sessile	Blended sessile section (BSS)	12 samples 108 PCR samples 45 negatives	Cell dislodgement (Duhamel & Jacquet, 2006) followed by phenol-chloroform extraction (Ip, Tay, et al., 2021)	COI, 16S, 18S	Illumina MiSeq (300-bp PE) Illumina HiSeq2500 (250-bp PE)
Sessile	≥2 mm	203 samples 203 PCR samples 30 negatives	abGenix animal tissue DNA extraction kit (Chang, Ip, Bauman, and Huang, 2020)	COI	Illumina MiSeq (250-bp PE)

were autoclaved, bleached, and cleaned with 70% ethanol and UV-sterilized. All reagents were also filtered through a sterile 0.2 µm Corning syringe filter before use.

2.3 | Gene amplification, library preparation and sequencing

We amplified up to three gene markers depending on the ARMS fraction studied: bacterial 16S for marine microbes (Caporaso et al., 2012), 18S rRNA for marine eukaryotes (Lane, 1991; Medlin et al., 1988; Stoeck et al., 2010), and COI for marine metazoans (Leray et al., 2013; Lobo et al., 2013) (Table 1). Primers, PCR mastermixes, thermocycling conditions for each gene, as well as multiplexing strategies and sequencing platforms are found in Table S1. For the metabarcoding assay, triplicate PCR reactions (all uniquely tagged) were conducted for each of the 36 samples (3 homogenized fractions × 12 ARMS units) and their associated extraction negatives, for each of the three gene markers. Negative PCR controls were also included for each batch of PCRs. Amplification success was verified on 2% agarose gels stained with GelRed (Biotium Inc.). In total, we performed 1651 PCR reactions (1440 PCR samples and 211 negative controls) (Table 1).

All tagged PCR samples were pooled and normalized within each gene marker while ensuring that tags were nonoverlapping. These PCR sample pools were then cleaned separately using AMPure XP (Beckman Coulter). Library preparation was only performed for COI and 18S markers as Illumina adapters were already added at the PCR step for 16S. Illumina PCR-free libraries were prepared with NEBNext Ultra II DNA Library Prep Kit (New England Biolabs) following the manufacturer's protocol up to the adapter ligation step (i.e., no PCR enrichment), and each of

these library pools were further multiplexed with unique Illumina adapters (TruSeq Single Indexes, Set B). All Illumina libraries were outsourced to Genome Institute of Singapore for sequencing (Table 1). Nanopore sequencing was conducted in-house, and the library was prepared using the Ligation Sequencing Kit (SQK-LSK109), following modifications from Chang, Ip, Bauman, and Huang (2020). MinION sequencing was run for 24 h with a fresh R9.4.1. flow cell on MinKNOW version 20.06.5 for Ubuntu 18 (version 4.0.5). Combined wet-laboratory processing and sequencing costs are compiled in Table S1.

2.4 | Bioinformatics

An overview of the bioinformatic processes involved for each gene and sequencing type is provided in Table 2.

2.4.1 | COI barcoding of ≥2 mm specimens

For nanopore barcoding, we standardized basecalling of fast5 files to Guppy version 4.2.2+effbaf8 for the new set of raw fast5 reads generated, along with the two Chang, Ip, Bauman, and Huang (2020) data sets. MinION barcodes were called using the improved mini-Barcoder pipeline (Chang, Ip, et al., 2020; Srivathsan et al., 2019). Finally, MinION barcodes were subjected to contamination checks, merged and selected for the best performing data set following the criteria set by Chang, Ip, Bauman, and Huang (2020).

For Illumina barcoding, we followed the pipeline in Chang, Ip, Bauman, and Huang (2020). A translation check was performed with Seqkit version 2.1 (Shen et al., 2016) to ensure our Illumina sequences had no internal stop codons.

TABLE 2 Sequencing results and bioinformatics pipelines used for each gene and sequencing type

Gene	Sequencing	Num. Of reads	Quality filtering	Taxonomic classification	ASVs/MOTUs
COI	Nanopore	1,748,677	Basecalling with Guppy version 4.2.2+effba8 Barcode calling and amino acid error correction with miniBarcode, see Chang, Ip, Bauman, and Huang (2020)	MOTU clustering with objective clustering (3% threshold) Match top hits with BLAST+ version 2.8.1; parse BLAST results with reidsidentifier RDP Classifier at 0.8 confidence cutoff on MIDORI Reference 2 web server	410 MOTUs (COI combined)
COI	Illumina	575,549	Assembly with PEAR version 0.9.10: -n 280 -m 320 -q 30 Demultiplexing and quality filtering with OBITools version 1.2.11 Barcode, see Chang, Ip, Bauman, and Huang (2020) Metabarcoding, see Ip, Chang, et al. (2021); Ip, Tay, et al. (2021) Pseudogene and chimera filtering with SeqKit version 2.1 translate function	MOTU clustering with objective clustering (3% threshold) Match top hits with BLAST+ version 2.8.1; parse BLAST results with reidsidentifier RDP Classifier at 0.8 confidence cutoff on MIDORI Reference 2 web server	410 MOTUs (COI combined)
18S	Illumina	3,242,890	Assembly with PEAR version 0.9.10: -n 160 -m 220 -q 20 Demultiplexing and quality filtering with OBITools version 1.2.11, see Ip, Chang, et al. (2021) Ip, Tay, et al. (2021). Chimera filtering with DECIPHER version 2.20 web server	MOTU clustering with objective clustering (1% threshold) Match top hits with BLAST+ version 2.8.1; parse BLAST results with reidsidentifier IDTAXA classifier using PR2 18S version 4.13 and SILVA SSU version 138 training sets, at 60% confidence threshold	561 MOTUs (combined), 375 non-metazoans, 186 metazoans
16S	Illumina	888,997	Remove adapter, barcode and primer sequence with Cutadapt version 2.1 Quality trimming with DADA2 version 1.16 with default parameters Remove reads in negative controls with decontam version 1.12 Remove organelle reads with phyloseq version 1.18	RDP classifier, trained against SILVA project version 138, at default 50% bootstrap confidence.	15,356 ASVs

2.4.2 | Illumina 18S and COI metabarcoding

Paired-end reads were merged using PEAR version 0.9.10 (Zhang et al., 2014; see Table 2). OBITools version 1.2.11 (Boyer et al., 2016) was used for demultiplexing and further downstream processing of assembled reads, following Ip, Tay, et al. (2021). The 18S data set was filtered for metazoan and non-metazoan sequences separately, while the COI data set was filtered only for metazoan sequences, using reidsidentifier version 1.0 (Srivathsan et al., 2015) to obtain preliminary taxonomic identities of each sequence. Briefly, the top 10 matches to GenBank database of each sequence were parsed by identity with the reidsidentifier.py script, to select the best match with the highest identity score at a minimum 80% length overlap of each sequence (250/313 bp for COI; 160/200 bp for 18S). Manual

inspection of the top 10 matches was also conducted to examine sequences with several highly similar matches and to downgrade the identities to the next taxonomic level if required. This was done after BLASTn implemented on BLAST+ version 2.8.1 (Camacho et al., 2009) to match against the NCBI nt database (downloaded 2 September 2019; a more recent download on 9 February 2022 did not improve sequence identities significantly—with <1% of sequences upgraded to species-level identity), retaining only sequence reads with ≥90% sequence similarity for 18S and ≥80% sequence similarity for COI.

We used the obistat module to summarize the total read count per replicate file, followed by a read count filtering with obigrep for each PCR replicate based on a relative threshold. Only sequences whose abundance exceeded 0.00001 ($1e^{-5}$) for 18S and 0.0001 ($1e^{-4}$) for COI of the total read count (excluding singleton

reads) for the PCR replicate were used in the analyses. We then used obiclean to collapse sequences with potential PCR sequencing errors into respective unique sequence reads. Sequences from sample PCR replicates were also matched against sequences found in the negative PCRs and removed from downstream analyses. Objective clustering (Meier et al., 2016) was carried out following Wang et al. (2018) to group sequence reads into molecular operational taxonomic units (MOTUs) based on a 1% distance threshold for 18S (Stoeck et al., 2010) and 3% for COI (Ip et al., 2019; Ip, Tay, et al., 2021). We screened for potential chimeras in 18S MOTUs with DECIPHER v2.20 web server (Wright et al., 2012). For COI MOTUs, we screened for pseudogenes and chimeras through translation checks with Seqkit version 2.1 (Shen et al., 2016). MOTUs that failed these checks were excluded from the overall data set. Lastly, to increase confidence of taxa detection, only MOTUs that were present in at least two out of three PCR replicates were retained for downstream analyses.

Final taxonomic identities of MOTUs were obtained using GenBank BLASTn and readsidentifier described above. In addition, we also parsed 18S MOTU data set through IDTAXA web classifier (Murali et al., 2018), at the recommended 60% confidence threshold for taxonomic assignment. We used both PR2 18S version 4.13 (Guillou et al., 2012) and SILVA SSU version 138 training sets for IDTAXA. Similarly, COI MOTUs were queried against the MIDORI Reference 2 server (Leray et al., 2018; Machida et al., 2017) with RDP Classifier (Wang et al., 2007) at the default 0.8 confidence cutoff to obtain taxonomic assignments. To aid the standardized reporting of identities across diverse taxa, MOTU identities were presented using GenBank output to species and genus level, while each sequence's general categorization of the remaining taxonomic levels up to phylum level were presented in the MOTU lists in Appendix S1. Additionally, all other taxonomic assignments from PR2 and MIDORI databases were not used for standardized reporting but shown in Appendix S1.

2.4.3 | Illumina 16S metabarcoding

Samples were first demultiplexed based on barcode indexes. Adapter, barcode and primer sequences were then removed using Cutadapt version 2.1 (Martin, 2011). Forward and reverse reads were trimmed to 240- and 160-bp, and then filtered with standard parameters using DADA2 version 1.16 (Callahan et al., 2016) in R version 3.5.2 (R Core Team, 2021) before inference of amplicon sequence variants (ASV). Chimeric reads were also de novo filtered. Reads found in blank extractions and PCR negatives were removed using the decontam version 1.12 package (Davis et al., 2018). Taxonomic assignment of ASVs was done with the RDP classifier method (Wang et al., 2007) against a reference training set from the SILVA version 138 database, using the default minimum 50% bootstrap cutoff recommended for 250 bp fragments (Wang et al., 2007). ASVs of organelle origin and those not assigned to a phylum were removed using phyloseq version 1.18 (McMurdie & Holmes, 2013).

2.5 | Data analysis

All ARMS samples were grouped by site and sample size fractions to examine differences in communities across all three genes via permutational multivariate analysis of variance (PERMANOVA; Anderson, 2001). These sample categories included the four deployment sites K, R, S, T; the three sample size fractions and sessile section were $\geq 2\text{ mm}$, $500\text{ }\mu\text{m}$, $106\text{ }\mu\text{m}$, and BSS. Following verification of homogeneity of dispersion using betadisper in vegan version 2.5 (Oksanen et al., 2013) on RStudio (R Core Team, 2021), we performed PERMANOVA using the adonis function to test whether sites and size fractions were associated with MOTU/ASV compositional differences. Similarities among 16S, 18S and COI communities across sites were then evaluated using the Jaccard similarity coefficient (Jaccard, 1901) and visualized using nonmetric multidimensional scaling (nMDS) plots.

To assess whether 12 ARMS units were sufficient for estimating overall putative species richness, we also plotted six accumulation curves for all MOTUs/ASVs across three gene (16S, 18S, COI) markers, PCR methods (barcoding and metabarcoding) and three broad taxa (bacterial microbes, non-metazoan eukaryotes, metazoans) using specaccum in vegan v2.5. Following which, the iNEXT package version 2.0.6 was used to plot rarefaction curves for MOTUs only, so as to extrapolate and predict the extent of additional sampling required for estimating overall reef cryptobiome richness (Hsieh et al., 2016). Taxonomic composition and relative taxon abundances for each ARMS unit were determined separately between barcoding and metabarcoding samples, since the diversity from individual $\geq 2\text{ mm}$ vouchers and homogenate samples were not directly comparable. As such, a combination of sequence read counts normalized by sample read depth within each metabarcoded size fraction (106, $500\text{ }\mu\text{m}$ and BSS), and the raw numbers of $\geq 2\text{ mm}$ specimens catalogued in the barcoding section, were used to estimate the overall relative abundances of cryptobiome taxa (Laporte et al., 2021).

Six photos from each ARMS unit were analysed with ImageJ version 2.0 (<https://imagej.net/Fiji/Downloads>) to quantify the plate area encrusted by sessile organisms by dividing the plate area encrusted over total plate area (expressed in percentage). For this purpose, we used images taken of the top and bottom surfaces of plates 1, 4 and 8 of every unit (Figure S1; top = a, bottom = b), following David et al. (2019). These plates represented microhabitat diversity within the ARMS units, where the top surface of plate 1 was fully exposed to the external environment while the remaining five surfaces were generally more sheltered. The bottom of plate 1, 4 and top surfaces of plate 8 were more open to current flows, while the latter two plates' bottom surfaces were more sheltered against currents.

Using geosphere and vegan, Mantel tests with the Spearman's rank method were conducted to determine correlations between communities (Jaccard similarity), environmental variables (Euclidean distance from mainland (km), deployment depth (m), total suspended solids (TSS; mg/L), mean sea surface temperature (SST; °C)) and geographic distance (distance among deployment sites; Haversine distance). Depth was recorded during deployment

in 2016, while TSS, SST, distance from a site to the nearest mainland Singapore shore, for years 2016–2017, were extracted from Ng et al. (2021). We also used one-way analysis of variance (ANOVA) and Tukey's post hoc test to investigate differences in ARMS plate encrustation, coral and invertebrate cover among sites. Coral and invertebrate cover data based on line intercept transect (LIT) surveys in 2016 were obtained from Wong et al. (2018). Diversity associated with the ARMS units at each site was quantified with the Shannon diversity index (H'). Relationships between each site's taxon diversity for all three genes and ARMS plate encrustation cover were investigated using linear regression models. One-way ANOVAs and Tukey's post-hoc tests were used to examine differences in Shannon diversity for all three genes among sites of varying coral and invertebrate cover.

3 | RESULTS

All 12 ARMS units were successfully retrieved. The $\geq 2\text{ mm}$ section yielded a total of 725 motile and 203 sessile specimens. MinION sequencing of the motile $\geq 2\text{ mm}$ specimens produced 1,748,677 demultiplexed reads (Table S1). Illumina sequencing generated the following amounts of paired-end reads after quality filtering: 888,997 16S microbial sequences; 2,507,734 non-metazoan eukaryote and 735,156 metazoan 18S sequences; as well as 157,941 and 420,608 metazoan COI sequences with metabarcoding and barcoding respectively. In summary, there were 15,356 microbial ASVs of 16S, 375 non-metazoan and 186 metazoan MOTUs with 18S, and 410 metazoan MOTUs with COI (Table 2). These comprise more than 50 microbial, 19 non-metazoan eukaryote and 15 metazoan phyla.

3.1 | Community assembly, richness and taxonomic composition

We found several key differences in MOTU and ASV richness patterns of the ARMS units among the four sites (K, R, S, T) (Figure 2a) for each major taxon group (microbial, non-metazoan, metazoan). For microbes (16S), site S had the highest richness with 4417 ASVs while site R had the lowest richness with 454 ASVs per unit. Site R had both the highest and lowest 18S richness per unit for non-metazoans (highest = 152, lowest = 44) and metazoans (highest = 71, lowest = 23). Contrastingly for the COI marker, we found site R to have the highest (107 MOTUs) and site S to have the lowest (47 MOTUs) MOTU richness per unit, while sites K and T had intermediate richness (COI: R > K/T > S).

MOTUs and ASVs overlapped more between sites than between size fractions (Figure S1). Comparing between sites, the highest number of unique non-metazoan MOTUs were found at site S (16S: 27.2%, 18S: 21.3%), and metazoan MOTUs at site S (18S: 16.1%) and R (COI: 22.4%). Taxon overlap between sites ranged from 21.8% for 16S microbial, 38.9% for 18S non-metazoans, 49.5% for 18S metazoans, to 33.8% for COI metazoans; while between size fractions, the

overlap ranged from 8.0% for 16S, 29.3% for 18S non-metazoans, 35.5% for 18S metazoans, to 8.3% for COI metazoans.

Relatedly, the overall reef cryptobiome showed similar taxonomic composition among sites but communities were more distinct across sample size fractions, especially for metazoan sequences. While each size fraction was dominated by vastly different metazoan phyla, they all generally comprised similar non-metazoan phyla (Figure 2b). Overall, microbial communities from the ARMS largely consisted of Proteobacteria (~70%; hereon, “~” represents “up to” the maximum percentage listed), Planctomycetes (~21%) and Actinobacteria (~19%) that dominated all ARMS units and fractions based on 16S metabarcoding, while Ascomycota (~66%), Basidiomycota (~28%), and Rhodophyta (~40%) were the dominant non-metazoan eukaryotes for 18S, with up to 66.7% of MOTUs that could not be assigned to any eukaryote phylum from GenBank records (<90% BLAST matches). For the metazoan component (18S and COI), Arthropoda and Mollusca dominated the motile fractions, while Porifera and Chordata (Asciidae) were abundant in the sessile section. The 18S metabarcodes showed that the 106 and 500 μm size fractions were dominated by Arthropoda (106 μm : ~75%, 500 μm : ~100%) and Annelida (106 μm : ~100%, 500 μm : ~100%), while Porifera (~45%) and Chordata (~33%) were prevalent in the BSS, with a small proportion of MOTUs that could not be assigned to any metazoan phyla (~5.6%). Rare metazoan phyla detected through 18S metabarcoding included Xenacelomorpha (~6.7%), Bryozoa (~8.7%) and Platyhelminthes (~3.2%) for 500 μm , 106 μm fraction and BSS sections, respectively. Metazoan composition based on COI metabarcoding corroborated the 18S results, with the 500 μm fraction dominated by Arthropoda (~100%) and Mollusca (~100%); the 106 μm fraction was dominated by Arthropoda (~100%); and the BSS section was dominated by Porifera (~85%) and Chordata (~33%). The rarest phyla detected were Annelida (~12.5%), Mollusca (~14.3%) and Nematoda (~3.0%) for 500 μm , 106 μm fraction and BSS sections, respectively. COI barcoding showed that the $\geq 2\text{ mm}$ fraction was dominated by Arthropoda (~55%), Mollusca (~45%) and Annelida (~30%), while Platyhelminthes was the rarest group (~3.3%). All COI MOTUs could be assigned to a metazoan phylum (~85% BLAST match) for both COI barcoding and metabarcoding.

The limited overlap of MOTUs/ASVs between size fractions was consistent among all three gene makers (16S: 1.4%, 18S non-metazoans: 12.0%, 18S metazoans: 12.9%, COI: 0.2%) (Figure S1). There was a higher degree of MOTU/ASV overlap between BSS and the two smaller-sized fractions (106 and 500 μm), while the degree of overlap between the two smaller-sized fractions (106 and 500 μm) was lower. For 16S and 18S, BSS had the most unique MOTUs/ASVs (16S microbial: 73.7%, 18S non-metazoans: 46.1%, 18S metazoans: 48.9%) and had more overlap in putative species with the larger 500 μm (16S microbial: 4.0%, 18S non-metazoans: 14.1%, 18S metazoans: 13.4%) than the smaller 106 μm fraction (16S microbial: 1.9%, 18S non-metazoans: 2.1%, 18S metazoans: 5.4%). Likewise, there were limited MOTUs/ASVs overlapping between 106 and 500 μm

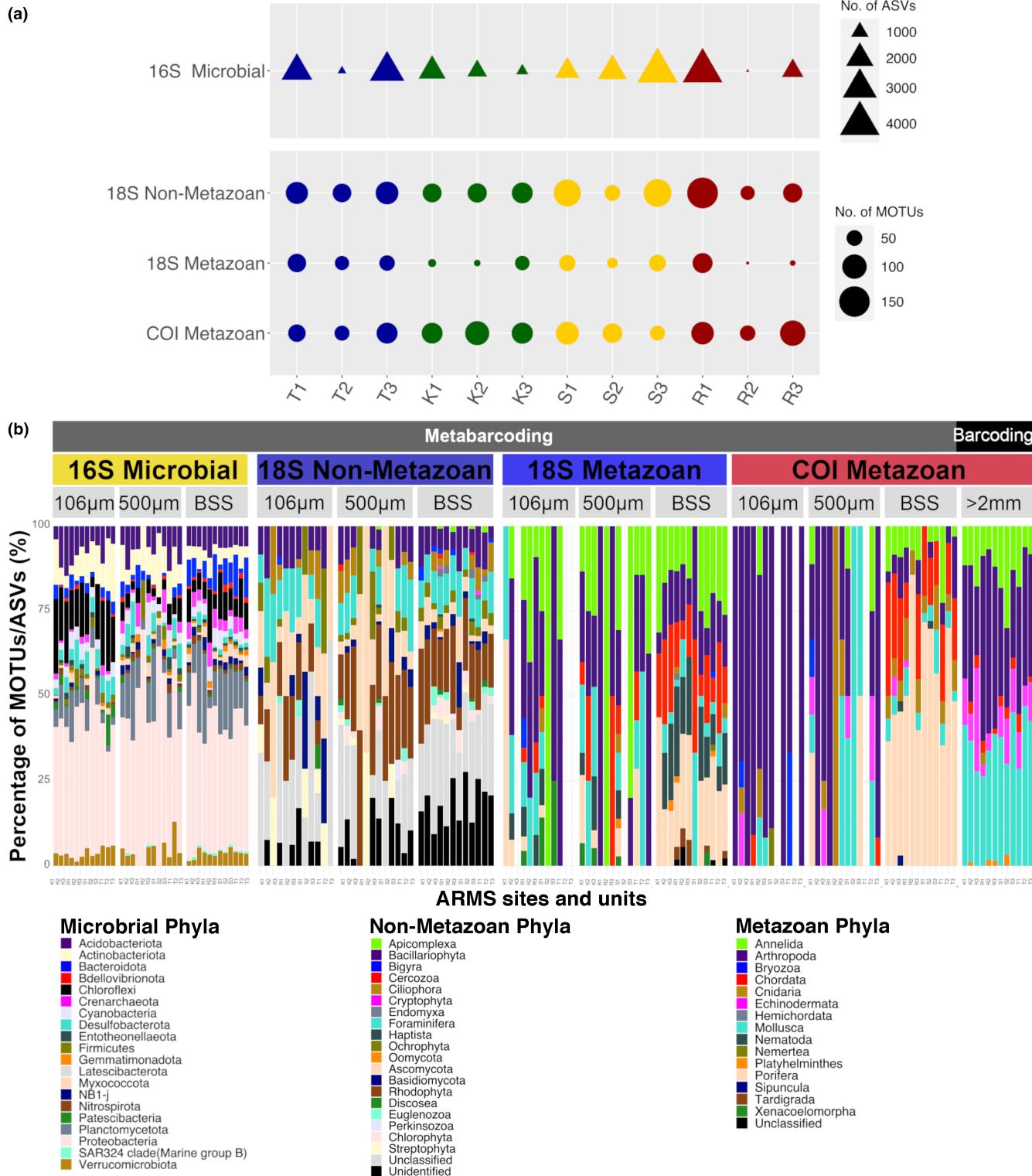


FIGURE 2 (a) Putative species (molecular operational taxonomic unit [MOTU] and amplicon sequence variant [ASV]) richness of each locality and Autonomous Reef Monitoring Structure (ARMS) unit (to scale), derived from three gene markers 16S, 18S and cytochrome oxidase I (COI). Microbial ASVs were of the 16S marker, non-metazoan eukaryote MOTUs the 18S marker, while metazoan MOTUs were derived from both COI and 18S markers. (b) Percentage composition (y-axis) of the microbial ASVs, non-metazoan and metazoan eukaryote MOTUs at phylum level for all 12 ARMS units (x-axis), for 16S (yellow; metabarcoding only; top 20 most abundant phyla), 18S (blue; metabarcoding only) and COI (red; both metabarcoding and barcoding) gene markers across the three size fractions (≥ 2 mm, 500 µm, 106 µm) and bulk sessile (BSS) section. The relative abundances of the barcoding section (top most, right-side black coloured strip) were inferred directly from number of collected ≥ 2 mm specimens, while metabarcoding section's (top most, left-side grey coloured strip) was generated from the normalization of sequence counts within each size fraction of the ARMS units. Refer to legends below for microbial, non-metazoan or metazoan eukaryote phylum colour codes, respectively

fractions (16S microbial: 0.7%, 18S non-metazoans: 1.1%, 18S metazoans: 3.8%). Similar patterns were observed with COI, where unique MOTUs from barcoding and metabarcoding were largely present in BSS (42.6%) and $\geq 2\text{ mm}$ (30.7%) fractions. Despite each fraction being categorized as distinct motile and sessile fractions, $\geq 2\text{ mm}$ and BSS had the greatest MOTU overlap between size fractions (2.7%). In contrast, the $106\text{ }\mu\text{m}$ (9.7%) and $500\text{ }\mu\text{m}$ (8.8%) fractions had the least unique MOTUs with the lowest overlap (0.7%) between each other and other size fractions.

3.2 | Diversity patterns and relationships with abiotic and biotic variables

Community analyses based on Jaccard similarity showed spatial variation among sites for metazoans but not non-metazoans (Figure 3). ARMS units from sites R and K formed distinct clusters, while units from sites T and S were generally intermixed regardless whether metazoan MOTUs from across size fractions and gene markers were combined or separated. This was supported by significant differences in metazoan communities between sites with both 18S (PERMANOVA: $df = 3, F = 1.373, R^2 = 0.346, p = 0.003$) and COI (PERMANOVA: $df = 3, F = 1.623, R^2 = 0.372, p = 0.0001$). Furthermore, within each size fraction, the differences were driven mainly by metazoan communities with 18S BSS (PERMANOVA: $df = 3, F = 1.260, R^2 = 0.107, p = 0.034$), COI BSS (PERMANOVA: $df = 3, F = 1.565, R^2 = 0.370, p = 0.002$) and COI $\geq 2\text{ mm}$ barcoding (PERMANOVA: $df = 3, F = 1.720, R^2 = 0.391, p = 0.0001$). Correspondingly, these size fractions also showed clear spatial patterns between sites (Figure S1a,d,f). On the contrary, there were neither observable spatial patterns of site nor significant differences with non-metazoan communities (Figure 3a,b and Figure S1b,c), for both 16S (PERMANOVA: $df = 3, F = 1.013, R^2 = 0.275, p = 0.354$) and 18S (PERMANOVA: $df = 3, F = 0.967, R^2 = 0.284, p = 0.601$). Cognizant that most non-metazoans are ubiquitous, planktonic taxa, we removed two known planktonic non-metazoans with unambiguous species identities from the 18S eukaryote community data set to examine if spatial patterns would be more distinct post-removal. There remained no significant differences in the communities between sites (PERMANOVA: $df = 3, F = 0.923, R^2 = 0.069, p = 0.604$), though omitting phytoplankton *Skeletonema tropicum* and diatom *Tribonema marinum* MOTUs allowed site K to be distinguished in the BSS section and $500\text{ }\mu\text{m}$ fraction (Figure S1h,i). Planktonic non-metazoan MOTUs were therefore unlikely to contribute significantly to community differences among sites.

Comparative analyses of the ARMS communities with environmental variables revealed several significant positive correlations based on Mantel tests (Table 3, Figure S1). Non-metazoan communities of the BSS section were distinguished between sites by the combination of environmental factors that included SST($^\circ\text{C}$) and TSS (mg/L), with a Mantel r statistic of 0.36 ($p = 0.024$) for 16S microbes and 0.41 ($p = 0.050$) for 18S eukaryotes. Conversely, metazoan

communities were significantly associated with distance from mainland and deployment depth but not with the remaining abiotic environmental variables (SST($^\circ\text{C}$), TSS (mg/L)). In particular, the $\geq 2\text{ mm}$ MOTUs from COI barcoding varied between sites according to differences in their distance from mainland with r statistic of 0.32 ($p = 0.034$), and all COI communities differed based on ARMS deployment depth with r statistic of 0.23 ($p = 0.031$).

The ranking of site richness for COI communities (see section 3.1, Figure 2a) had limited correspondence with the ranking of ARMS plate encrustation cover ($R > K > T > S$) and coral cover ($R > T > K > S$) (Figure 4a). Firstly, differences in ARMS plate encrustation across sites were found to be marginally significant (ANOVA: $F[3, 68] = [2.306], p = 0.08$). Post-hoc analyses showed significant difference only between sites R and S (Tukey's honest significance difference [HSD]: $p = 0.05$), where site R was the most encrusted (mean $74.9\% \pm \text{SD } 27.7\%$) and site S was the least encrusted (mean $57.6\% \pm \text{SD } 30.8\%$), while site K (mean $69.0\% \pm \text{SD } 18.8\%$) and site T (mean $67.9\% \pm \text{SD } 34.1\%$) were moderately encrusted (Figure 4a). Secondly, the trends for coral cover differences among sites were similar to ARMS plate encrustation trends, and were highly significant (ANOVA: $F[3, 16] = [50.283], p < 0.01$). Apart from few insignificant differences between sites R and T (Tukey's HSD: $p = 0.78$), sites S and K (Tukey's HSD: $p = 0.13$), all other pairwise differences in coral cover between sites were significant, with the most significant difference found between sites R and S (Tukey's HSD: $p = 1\text{e}^{-7}$). Likewise, site R had the highest (mean $60.4\% \pm \text{SD } 11.2\%$) and site S the lowest coral cover (mean $9.5\% \pm \text{SD } 3.3\%$), while site K (mean $21.4\% \pm \text{SD } 5.5\%$) and site T (mean $55.6\% \pm \text{SD } 14.6\%$) had intermediate coral cover.

In contrast, there seemed to be a slight inverse trend in reef invertebrate cover ($S > K > T > R$) that was found to be insignificant (ANOVA: $F[3, 16] = [0.36], p = 0.78$), with site S having the highest (mean $4.0\% \pm \text{SD } 3.5\%$) and site R the lowest invertebrate cover (mean $2.3\% \pm \text{SD } 1.1\%$). Furthermore, diversity trends of COI metazoan communities with plate encrustation ($p = 0.85, R^2 = 0.004$) and coral cover ($p = 0.43, R^2 = 0.063$) seemed to be positively correlated (Figure 4b,c), although the relationships were insignificant. In particular, Shannon diversity was the highest for site R and lowest for site S, and the linear regression model demonstrated a positive relationship between COI metazoan diversity with both ARMS plate encrustation and coral cover (Figures 4b,c and 5). However, these diversity trends were not observed with the communities generated from 16S and 18S, and showed negative correlations with both ARMS plate encrustation (16S: $p = 0.33, R^2 = 0.096$; 18S: $p = 0.09, R^2 = 0.250$) and coral cover (16S: $p = 0.37, R^2 = 0.0784$; 18S: $p = 0.67, R^2 = 0.0196$) (Figures 4b,c and 5). Despite boxplots showing differences in Shannon diversity across sites of varying coral and invertebrate cover, the one-way ANOVA tests showed that these differences were insignificant (Figures 4c,d): 16S microbial (ANOVA: $F[3, 8] = [0.34], p = 0.80$); 18S non-metazoans (ANOVA: $F[3, 8] = [0.24], p = 0.87$); 18S metazoans (: $F[3, 8] = [1.11], p = 0.40$); COI metazoans (ANOVA: $F[3, 8] = [0.64], p = 0.61$).

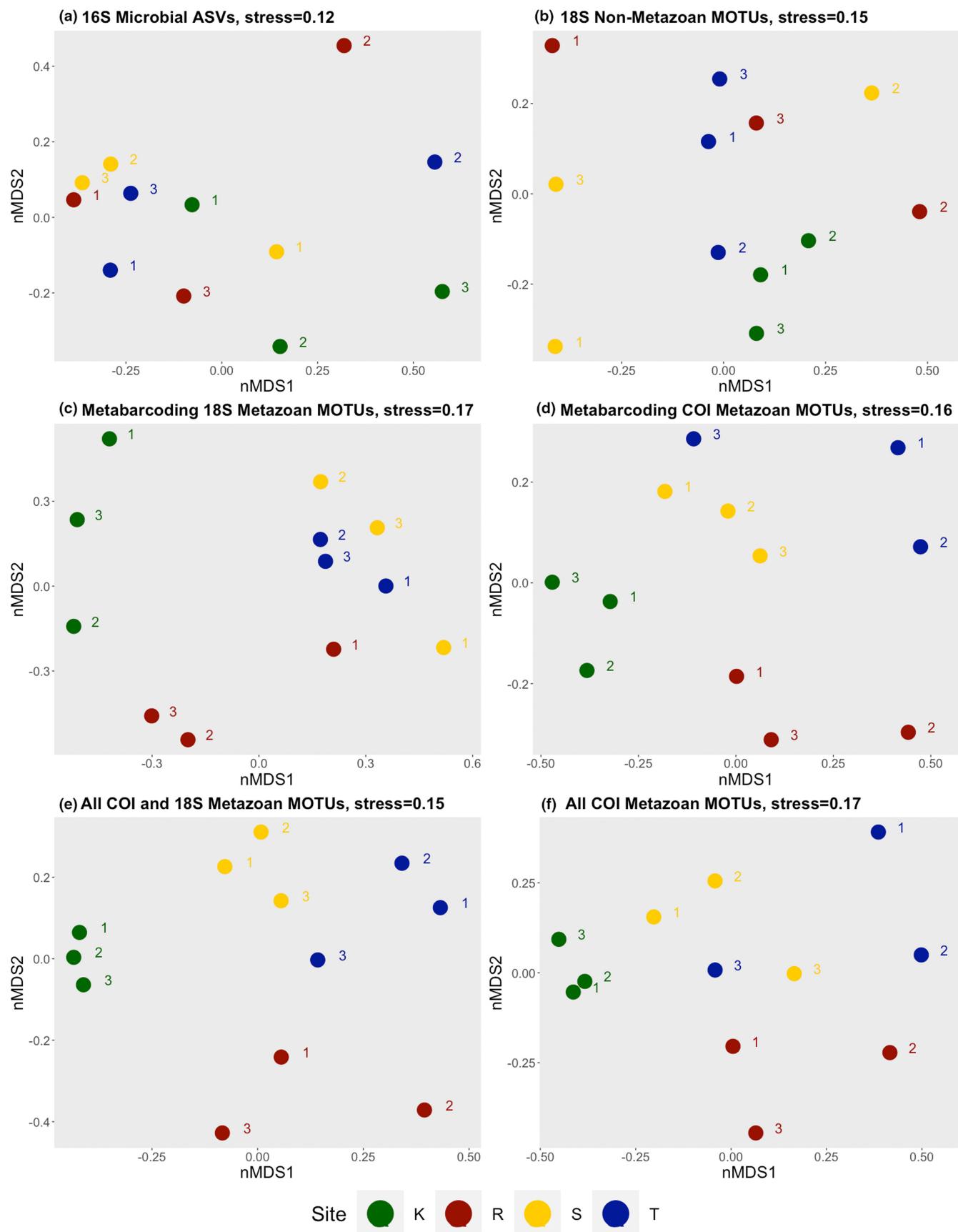


FIGURE 3 Nonmetric multidimensional scaling (nMDS) based on Jaccard similarity among amplicon sequence variants (ASVs) (16S microbial only), non-metazoan and metazoan eukaryote molecular operational taxonomic units (MOTUs) detected across 12 Autonomous Reef Monitoring Structure (ARMS) units using three gene markers, 16S, 18S and cytochrome oxidase I (COI). Colour and number of each dot corresponds to the ARMS locality and unit, respectively. (a) 16S metabarcoding microbial ASVs. (b) 18S metabarcoding non-metazoan eukaryote MOTUs. (c) 18S metabarcoding metazoan MOTUs. (d) COI metabarcoding metazoan MOTUs. (e) all metazoan MOTUs (COI barcoding, COI and 18S metabarcoding). (f) COI barcoding and metabarcoding metazoan MOTUs

TABLE 3 Mantel tests based on Bray-Curtis dissimilarity of environmental variables SST(°C) and TSS (mg/L), mainland and geographical distance (matrix 2) against community distance (matrix 1)

Matrix 1	Matrix 2	Mantel's R	p-value
16S microbial BSS ASVs	All env parameters	0.3561	0.0239
18S eukaryote BSS MOTUs	All env parameters	0.4108	0.0496
COI metazoan barcoding MOTUs	Depth (m)	0.1838	0.0762
COI metazoan all MOTUs	Depth (m)	0.229	0.0314
COI metazoan barcoding MOTUs	Distance mainland (km)	0.3191	0.0341
16S microbial 106 µm MOTUs	Distance mainland (km)	0.2255	0.0584

Note: Statistically significant associations are indicated with bold p-values; remaining comparisons are marginally significant. Refer to Appendix S1 for full list of mantel test results.

Abbreviations: ASVs, amplicon sequence variants; BSS, blended sessile section; MOTUs, molecular operational taxonomic units.

3.3 | Overall richness estimates and identification confidence

Accumulation curves of MOTUs/ASVs showed that 12 ARMS units were insufficient to provide total richness estimates for Singapore's reef cryptobiome. This was consistent across all genes and taxa, with no asymptotes reached on the accumulation curves and the interpolated rarefaction curves (Figure 6). Conversely, sequencing depths were sufficient for recovering cryptobenthic diversity from the ARMS units, where asymptotes were rapidly reached on accumulation curves both within and across sites for all three gene markers (Figure S1). Nevertheless, the ≥ 2 mm barcoding accumulation curve was the closest to attaining a plateau (Figure 6b). The rest of the communities generated with metabarcoding showed rapid rises in MOTU/ASV accumulation, decelerating only slightly from approximately eight ARMS units. Analyses predicted that at least 40–50 ARMS units would be required for more accurate estimates of overall species richness (Figure 6c).

Despite high richness revealed by the accumulation curves, most MOTUs/ASVs remained without names or identity matches to GenBank, SILVA, PR2 or MIDORI databases (Figure S1). MOTUs generated with the COI marker experienced the best coverage in GenBank as expected, followed by 18S non-metazoan eukaryotes, 18S metazoans and lastly, 16S microbes. Species-level matches ($\geq 97\%$) were possible across all metazoan phyla with barcoding, albeit disproportionately, with proportions of identifiable MOTUs ranging from 2.5% in Annelida to 42.9% and 100.0% in Mollusca and Cnidaria respectively. Only 54.6% of all phyla from COI metabarcoding had species-level identifications; Echinodermata (30.0%), Porifera (23.6%) and Cnidaria (17.7%) had the most species-level identifications, while Mollusca (5.7%) and Arthropoda (0.0%) had the least. Notably, COI metabarcoding uncovered phyla that were not present with barcoding, such as Nemertea, Nematoda and

Sipuncula. Although 18S metabarcoding of metazoans had lower overall proportions of species-level identification (38.5%) than COI (54.6%), it fared better at identifying MOTUs from Annelida (16.1%), Arthropoda (2.4%) and Mollusca (27.8%) to species. Additionally, rare phyla missed by COI metabarcoding were detected by 18S metabarcoding with 33.3% identifiable at least to genus. As for the non-metazoan phyla, metabarcoding with 18S gene marker allowed them to be better identified (79.0%) at least to genus level ($\geq 99\%$ without GenBank names) than metazoan phyla (53.9%). Further, more than half of the microbial ASVs (62.9%) from 16S metabarcoding were identifiable only to phylum level.

4 | DISCUSSION

Testing marine biodiversity patterns across varying spatial and temporal scales is necessary amid rapid environmental change caused by anthropogenic and natural impacts. The widespread extinction of species that could occur before their ecological significance is fully recognized (Costello & Wilson, 2011) underscores the urgency in understanding these patterns to better enhance monitoring and management approaches. Conservation strategies tracking changes in biodiversity trends would benefit from the use of standardized methodologies like ARMS, which allows comparisons of communities across space and time (Carvalho et al., 2019; Pearman et al., 2018, 2019).

4.1 | Characterizing cryptobiome communities in tropical reefs

Marine invertebrates are involved in key biological processes occurring within reef ecosystems (Glynn & Enochs, 2011; Plaisance

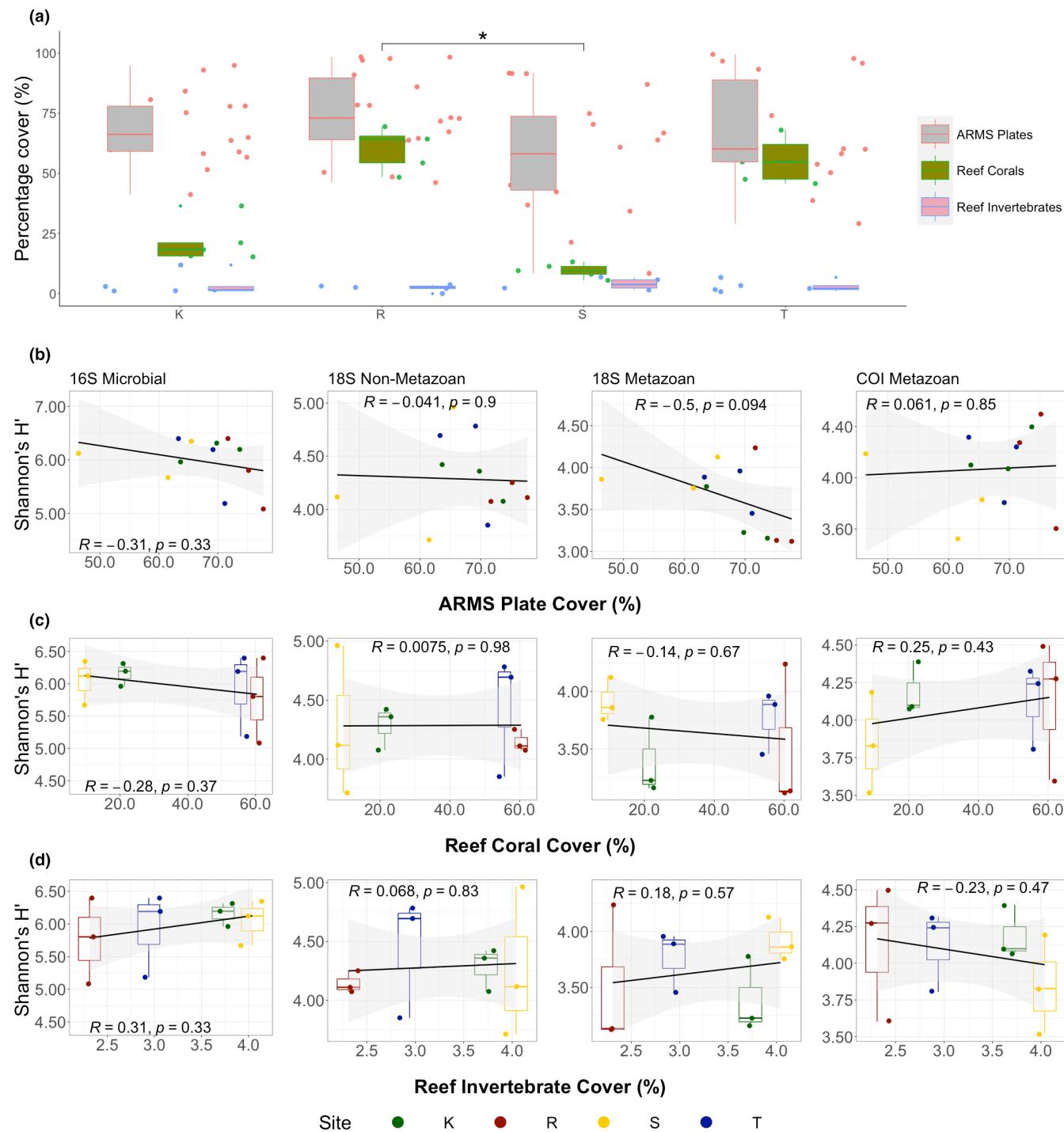


FIGURE 4 (a) Boxplot showing percentage of Autonomous Reef Monitoring Structure (ARMS) plate cover (grey fill/pink data points), coral cover (brown fill/green data points) and invertebrate cover (pink fill/blue data points) at each of the four sampling sites. The percentage cover differences were significant between R and S, indicated by **. (b) Scatter plots illustrating the relationships between alpha diversity for all three genes/taxa based on Shannon diversity index (H') for each gene/taxon, against percentage of ARMS plate cover from each of the three units at each site. A linear regression model is fitted for each gene/taxon, with R and p -values indicated in each plot. (c,d) boxplots with data points showing Shannon diversity of communities for each gene/taxon, against percentages of coral (c) and invertebrate (d) cover at each site. Linear regression models are fitted for all genes/taxa, with R and p -values indicated in each plot

et al., 2011). Here, we found Annelida, Arthropoda, and Mollusca were the most dominant motile metazoan phyla, while Porifera and Chordata (Asciidiacea) were the most abundant sessile metazoan phyla, corroborating results of other ARMS studies (Carvalho

et al., 2019; Pearman et al., 2020). Notably, sponges and ascidians play functional roles in cycling of nitrogen and particulate organic carbon (De Goeij et al., 2013; Ribes et al., 2005), and many epifauna influence food availability as bridges between primary producers and

FIGURE 5 Macrofaunal diversity and encrustation on Autonomous Reef Monitoring Structures (ARMS) have positive relationships with coral cover. High coral cover sites have high habitat complexity, hence more crevices and cavities in the reef matrix that provide microhabitats and refugia for reef associates, especially larger-sized metazoan species (≥ 2 mm). ARMS mimic the three-dimensional structure of the reef, so these larger-sized reef associates may utilize ARMS in a similar manner. Microbial communities are relatively homogenous among sites of varying coral cover. Therefore, ARMS deployed at high coral cover sites have increased encrustation with higher diversity of metazoan macrofauna but lower diversity of microbes. This figure was partially created with BioRender.com

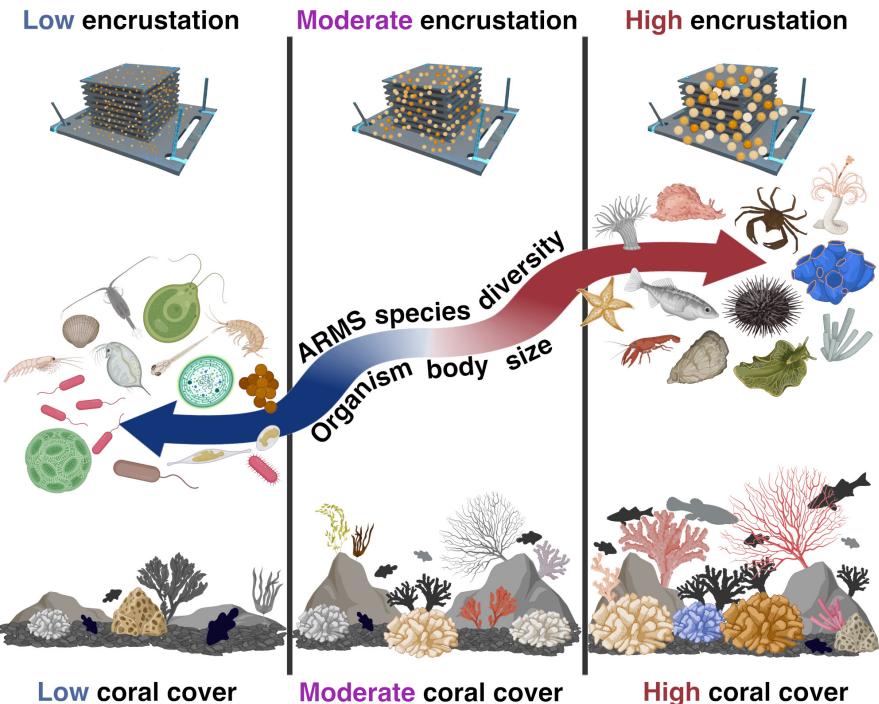
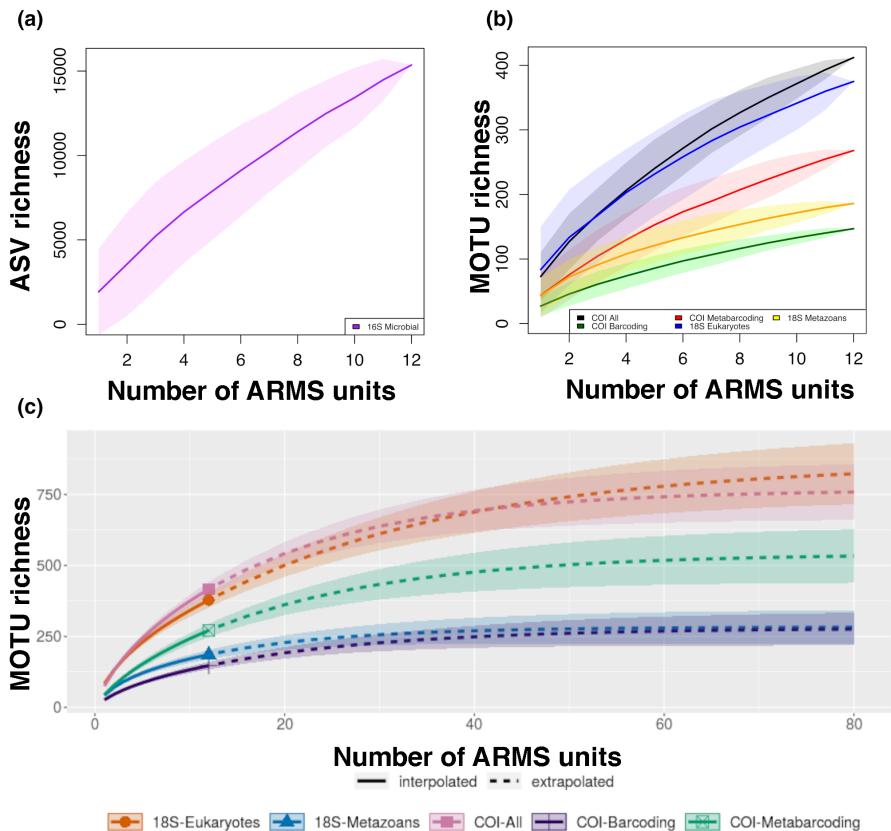


FIGURE 6 Putative species accumulation curves for each of the three gene markers targeting microbial amplicon sequence variants (ASVs) (a), non-metazoan or metazoan molecular operational taxonomic units (MOTUs) (b) from all 12 Autonomous Reef Monitoring Structure (ARMS) units. Shaded areas denote 95% confidence intervals. In (c), the MOTU rarefaction curves show projections of additional sampling with ARMS units needed for overall richness estimates of non-metazoan and metazoan MOTUs with 18S and cytochrome oxidase I (COI) markers. Dotted lines show projected values and shaded areas denote 95% confidence intervals



secondary consumers (Kramer et al., 2017). Further characterization of metazoan invertebrate community structure, especially with ARMS that can better survey invertebrates from hidden reef cavities (Vicente et al., 2021), will improve our understanding of shallow reef food webs (Fraser et al., 2021). On the one hand, dominant taxa form the foundational communities of the cryptobiome and can indicate

the degree of connectivity between sites (Volkov et al., 2007). On the other hand, rare metazoans (e.g., Platyhelminthes and Nematoda in this study) can be highly specialized with limited niche preferences and are expected to drive changes in spatial patterns between sites (Logares et al., 2014; Magurran & Henderson, 2003). Several rare taxa, such as *Pseudobiceros hancockanus* (Platyhelminthes:

Pseudocerotidae) and *Liomera venosa* (Arthropoda: Xanthidae) found at specific study sites highlight the presence of unique microhabitats on the reefs (Jousset et al., 2017).

The majority of non-metazoan eukaryote taxa are small and difficult to sample. ARMS becomes advantageous as it provides an efficient method to collect, study and characterize such communities (Casey et al., 2021; Palomino-Alvarez et al., 2021; Pearman et al., 2019, 2020). Here, we identified marine fungi Ascomycota, Basidiomycota, and red algae Rhodophyta as the main non-metazoan eukaryote contributors to a tropical urban reef cryptobiome (Figure 2b; Appendix S1). These taxa are essential components of the marine environment. In particular, primary producers such as macroalgae, marine plants and phytoplankton are major food sources for primary and secondary consumers (Jones et al., 2015). Additionally, marine fungi play fundamental roles in the ecosystem as saprotrophs and are known for their significant contributions to decomposition cycles (Ainsworth et al., 2017), which in turn provide food for detritivores (Jones et al., 2015). More importantly, the establishment of marine microeukaryote community structure can also be potentially used as bioindicators of environmental change or pollution (Xu et al., 2015). Characterization of eukaryote communities from ARMS will thus have important implications for understanding reef ecosystem health and species trophic functions, since these non-metazoans are involved in food production and detrital nutrient cycling that support significant reef organismal biomass (Rogers et al., 2018).

Although ARMS was initially designed for analysing communities of marine invertebrates, we also examined marine bacteria from the two smaller-sized fractions and sessile section that were consolidated as bulk samples (i.e., 106, 500 µm and BSS). Metabarcoding analyses revealed that the bacterial communities in the cryptobiome were dominated by Proteobacteria, Planctomycetes, and Actinobacteria (Figure 2b). Within Proteobacteria, classes Alphaproteobacteria and Gammaproteobacteria were the most abundant bacteria (Figure 2b; Appendix S1), corroborating findings by Pearman et al. (2019). Both classes are known to be associated with reef sediment (Rusch et al., 2009) and host organisms such as sponges, corals and macroalgae (Olson & Kellogg, 2010; Wainwright, Afiq-Rosli, et al., 2019; Wainwright, Bauman, et al., 2019). Additionally, Planctomycetes, typically associated with sponges and macroalgae, is involved in ammonium oxidation (Kaboré et al., 2020), while Actinobacteria is commonly found in various marine organisms like ascidians, corals, sponges, and molluscs, taking on roles in carbon and nitrogen cycling (Valliappan et al., 2014). These bacteria-host relationships correspond with the dominant metazoan communities found colonizing ARMS based on COI and 18S, especially the sessile section with high relative abundances of sponges and ascidians that are known to have significant impacts on bacterial community composition (Kelly et al., 2014). More broadly, bacteria play fundamental roles in biogeochemical cycles and nutrient cycling and have major responsibilities in ensuring the functioning of coral reefs (Garren & Azam, 2012). As such, our characterization of bacterial communities within cryptobiomes provides an important baseline for future studies focusing

on assessments of environmental change impacts on tropical urban reef ecosystems.

Most ARMS studies have used a single gene marker to analyse either barcoding or metabarcoding sections, rendering it challenging to characterize cryptozoa at higher taxonomic classifications and across all size fractions (Appendix S1). Our approach with three gene markers across various ARMS fractions has revealed a broader range of taxa across the domains of life, enabling standardized comparisons of diversity patterns among more than 50 microbial, 19 non-metazoan eukaryote and 15 metazoan phyla. This has also allowed documentation of diverse organism groups without sole reliance on taxonomic expertise. Despite reference databases being integral for taxonomic assignments, we continue to observe incompleteness of database coverage even for frequently-used gene regions like COI (Carugati et al., 2015; Ip, Tay, et al., 2021). Large proportions of sequences did not attain species-level matches, especially for the metabarcoding of smaller-sized fractions but unexpectedly even for the more prominent ≥2 mm samples (Figure S1). Unsurprisingly, microbial 16S sequences have the lowest identity coverage from SILVA since microorganisms are less commonly named at the species level (Trüper, 1999). Nevertheless, the use of different primers to collectively target metazoans proved to be relatively effective in circumventing incomplete database coverage. We found 18S to complement COI for better identification within Annelida and Mollusca, while also detecting several rare metazoan phyla missed by the COI gene marker.

Despite uncovering high cryptobiome diversity across all sites, many MOTUs remained as putative species units and were unidentified without names (Figure S1), highlighting that more taxonomic work is needed to complement the molecular results to provide meaningful sequence matching and identification. However, sorting and identification of such a wide range of taxa using traditional methods would require high levels of skilled taxonomic expertise and manpower that are considerably costly. We estimate that the work here in Singapore that was completed within 6 months, costing approximately USD\$50,000, would have taken at least 2 years requiring >USD\$120,000 to manually sort and identify specimens to the lowest taxonomic level (Table S1). More importantly, the mass collection of genetic material from ARMS can be subjected to wide-ranging molecular and sequencing methods, which enhance the workflow efficiency for species discovery, documentation of biodiversity and assessing changes in community patterns.

4.2 | Distinct spatioenvironmental drivers of cryptobenthic and macrobenthic communities

Reef coral cover is often associated with habitat structural complexity, which increases the availability of refugia for various reef-associated fauna (Fraser et al., 2021; Idjadi & Edmunds, 2006; Komyakova et al., 2013; Rogers et al., 2018). While there is a well-established positive relationship between coral and invertebrate diversity using invasive sampling methods (Nelson et al., 2016), we

observed from LIT surveys that site R had the highest coral cover, followed by sites T, K and then S, but this trend was reversed for invertebrate cover, with only the coral cover differences being significant (Figure 4a). This does not necessarily reflect low invertebrate diversity at high coral cover sites, because reef fauna can be hidden within the reef structures and utilize cavities for protection against predation (Fraser et al., 2021; Rogers et al., 2018).

Since ARMS mimics the physical complexity of coral reefs, organisms that are dependent on reef cavities would likely also find ARMS suitable for inhabitation (Palomino-Alvarez et al., 2021; Pearman et al., 2020; Plaisance et al., 2011; Zimmerman & Martin, 2004). Interestingly, we found that higher coral cover was associated with higher COI metazoan diversity though not with 18S metazoan diversity (Figure 4c), so it is possible that non-invasive survey methods focused on conspicuous reef organisms have overlooked the “hidden majority” concealed within crevices that could drive these invertebrate richness patterns (Pearman et al., 2018). Specifically, ARMS units from site R (high ARMS plate and coral cover) were encrusted to a significantly higher degree than site S (low ARMS plate and coral cover) (Figure 4a). The plates were furthermore encrusted by large sessile metazoans like sponges (Figure 2b) that depend on the three-dimensional reef structure. These organisms reside within habitats in hidden spaces of the reef (Richter et al., 2001) and are likely missed by traditional visual surveys. However, we note that ARMS encrustation was much less variable than coral cover, and while the latter at site T (mean 55.6%) was much higher than at site K (mean 21.4%), ARMS plate cover (Figure 4a,b) and metazoan diversity (Figure 4c) were largely similar between the two sites. Moreover, none of the other taxa examined displayed consistent, significant relationships with reef coral cover. Therefore, there was overall limited support for the hypothesis that cryptobenthic cover and diversity are positively associated with reef coral cover. The fact that high coral cover sites were associated with low invertebrate cover (Figure 4a) suggests that there may be complex trade-offs in reef diversity between corals competing for space and facilitating biodiversity (Brandt et al., 2019; Idjadi & Edmunds, 2006; McWilliam et al., 2018). More broadly, reef conservation strategies relying on diversity measures need to look beyond coral species richness (e.g., Veron et al., 2009, 2015) because biodiversity patterns are likely to be distinct across the tree of life (see also Pearman et al., 2018, 2019, 2020; Plaisance et al., 2021).

There were clear spatial differences in ARMS-recruited metazoan communities among reefs (Figure 3), with significant positive relationships between community and depth, and distance from mainland (Table 3, Figure S1). In particular, larger metazoans (i.e., $\geq 2\text{ mm}$ and BSS sections) were observed to drive spatial patterns here (Figure S1), and this was supported by known contributions of larger-sized reef-associates to assemblage dissimilarity among sites (Fraser et al., 2021; Pearman et al., 2018). Macrofauna that are known to depend on reef cavities (Fraser et al., 2021; Nelson et al., 2016) would also find the ARMS suitable for inhabitation because of the structurally complex niches it mimics (Palomino-Alvarez et al., 2021; Plaisance et al., 2011; Zimmerman & Martin, 2004) and the refugia it

provides (Rogers et al., 2018). This suggests that higher coral cover sites may host more unique metazoan communities than lower coral cover sites (Figures 4 and 5). Nevertheless, the close proximity between sites T and S (Figure 1; see full list of Mantel test results in Appendix S1) could potentially mean higher degree of connectivity between these sites and more similar communities, while higher nutrient levels at nearshore site K compared to offshore site R (Gin et al., 2003) could influence ARMS community similarity (Pearman et al., 2018). Indeed, the abundance of larger metazoan species can collectively influence the spatial patterning of ARMS communities (Cottenie, 2005; Fraser et al., 2021).

No observable spatial differences in diversity among sites were found for the smaller-sized metazoan (106 , $500\text{ }\mu\text{m}$) and non-metazoan communities (Figure 3 and Figure S1b,c,g). These mainly comprised planktonic taxa (Appendix S1) that have limited requirements for the three-dimensional reef structure (Ainsworth et al., 2017; Jones et al., 2015), hence, their recruitment may be less dependent on coral cover or habitat complexity. Instead, the prevalence of marine fungi, bacteria, phytoplankton and other planktonic organisms are contingent on biogeochemical cycles or other intricate biological pathways in the ecosystem (Jones et al., 2015; Kling et al., 2020; Rusch et al., 2009). Considering the well-established transport models in Singapore's coastal waters (Tay et al., 2012), the higher degree of taxon overlap between sites (Figure 2 and Figure S1) implies that there is high site connectivity with the prevailing current patterns here (Afif-Rosli et al., 2021) resulting in homogenization of non-metazoan communities (Limberger & Wickham, 2012; Rabbani et al., 2021) and metazoan pelagic larvae (Molen et al., 2018; Siegel et al., 2008). Despite the lack of spatial patterns, there was a significant positive relationship between the differences in non-metazoan communities with the combination of all environmental variables tested (Table 3, Figure S1). These parameters, especially SST and TSS, can influence water quality and nutrient levels, and are known to be important drivers of non-metazoan community structure (Humphreys et al., 2018). Since different communities were influenced by various biotic and abiotic factors, it is clear that ARMS diversity patterns need to be analysed in conjunction with comprehensive environmental and benthic survey data to better understand the drivers of diversity among reef-associated taxa, from microbes to macrofauna.

5 | CONCLUSIONS

Utilizing ARMS units and multiple gene markers to study the overlooked cryptobiome diversity across the tree of life on coral reefs in Singapore resulted in the detection of 971 eukaryote and 15,356 bacterial putative species. However, 12 ARMS units were inadequate, and up to 50 units may be needed in conjunction with increased sequencing depth to sufficiently estimate the overall cryptobenthic richness of this small yet diverse marine area. Nevertheless, we suggest that ARMS can functionally mimic reef structural complexity for settlement of organisms with specialized habitat requirements within reef cavities, revealing distinct diversity patterns among taxa

across the tree of life. Consequently, high coral cover sites contained greater diversity of larger-sized metazoans as recovered from ARMS. In contrast, undifferentiated planktonic non-metazoan communities among sites reflected community connectivity facilitated by current transport. The ARMS taxonomic components did not covary with one another, where only diversity patterns of the COI communities containing more larger-sized metazoans can be explained by biotic factors like coral and invertebrate cover. The remaining taxa consisting of planktonic microeukaryotes and non-metazoans are known to be closely linked to biogeochemical pathways (Peixoto et al., 2017). Therefore, their community patterns were better explained by abiotic factors such as STT, TSS, deployment depth and distance from mainland.

This study furthermore highlighted the inadequacy of sequence databases for meaningful taxonomic assignments, underscoring the importance of barcoding vouchered collections to fill these gaps (Ip et al., 2019) so that specimen sorting and identification of under-studied fauna can be achieved much more efficiently using reverse workflow methods (Wang et al., 2018). Nonetheless, the comprehensive approach applied here has helped further our understanding of the reef cryptobiome by uncovering diversity that would otherwise be challenging to sample from the mature and complex reef matrix.

AUTHOR CONTRIBUTIONS

Danwei Huang and Andrew G. Bauman conceived the study. ARMS units were purchased by Andrew G. Bauman and Danwei Huang secured funding for the project. Yin Cheong Aden Ip, Yong Kit Samuel Chan, Andrew G. Bauman, and Danwei Huang deployed the units. Yin Cheong Aden Ip and Jia Jin Marc Chang led the retrieval and sorting workflows. The molecular work and data analysis for COI and 18S were performed by Yin Cheong Aden Ip, while 16S was performed by Ren Min Oh and Zheng Bin Randolph Quek. MiNiON sequencing and analyses were performed by Jia Jin Marc Chang and Yin Cheong Aden Ip. Yin Cheong Aden Ip drafted the manuscript and prepared the figures with input from Jia Jin Marc Chang and Danwei Huang.

ACKNOWLEDGEMENTS

We are extremely grateful to the following for their help in fieldwork and processing: Sudhanshi Sanjeev Jain, Jovena Chun Ling Seah, Zack Chen, Chin Soon Lionel Ng, Joy Shu Yee Wong, Sherlyn Sher Qing Lim, Zhi Ting Yip, and Jun Wei Phua. We also acknowledge the National Supercomputing Centre, Singapore (<https://www.nscc.sg>) for providing the computational resources for analysis. This research was supported by the Singapore Ministry of Education Academic Research Fund Tier 1 (R-154-000-A63-114), the National Research Foundation, Prime Minister's Office, Singapore under its Marine Science R&D Programme (MSRDP-P03), and an AXA Postdoctoral Fellowship (R-154-000-649-507) to A.G.B.

CONFLICT OF INTERESTS

All authors approved the final version for publication and declare that the research was conducted in the absence of any commercial

or financial relationships that could be construed as a potential conflict of interest.

DATA AVAILABILITY STATEMENT

Ip, Y. C. A., Chang, J. J. M., Oh, R. M., Quek, Z. B. R., Chan, Y. K. S., Bauman, A., Huang, D. Singapore ARMS raw sequence reads have been made available on NCBI Sequence Read Archive under BioProject PRJNA801066.

ORCID

Yin Cheong Aden Ip  <https://orcid.org/0000-0002-9472-1161>
 Jia Jin Marc Chang  <https://orcid.org/0000-0001-5208-4929>
 Ren Min Oh  <https://orcid.org/0000-0002-6256-6291>
 Zheng Bin Randolph Quek  <https://orcid.org/0000-0001-7998-7052>
 Yong Kit Samuel Chan  <https://orcid.org/0000-0003-0922-9357>
 Andrew G. Bauman  <https://orcid.org/0000-0001-9260-2153>
 Danwei Huang  <https://orcid.org/0000-0003-3365-5583>

REFERENCES

- Afiq-Rosli, L., Wainwright, B. J., Gajanur, A. R., Lee, A. C., Ooi, S. K., Chou, L. M., & Huang, D. (2021). Barriers and corridors of gene flow in an urbanised tropical reef system. *Evolutionary Applications*, 14(10), 2502–2515.
- Ainsworth, T. D., Fordyce, A. J., & Camp, E. F. (2017). The other micro-eukaryotes of the coral reef microbiome. *Trends in Microbiology*, 25(12), 980–991.
- Albano, P. G., Sabelli, B., & Bouchet, P. (2011). The challenge of small and rare species in marine biodiversity surveys: Micromollusk diversity in a complex tropical coastal environment. *Biodiversity and Conservation*, 20(13), 3223–3237.
- Anderson, M. J. (2001). A new method for non-parametric multivariate analysis of variance. *Austral Ecology*, 26(1), 32–46.
- Bellwood, D. R., & Meyer, C. P. (2009). Searching for heat in a marine biodiversity hotspot. *Journal of Biogeography*, 36(4), 569–576.
- Biondi, P., Masucci, G. D., & Reimer, J. D. (2020). Coral cover and rubble cryptofauna abundance and diversity at outplanted reefs in Okinawa, Japan. *PeerJ*, 8, e9185.
- Bouchet, P., Ng, P. K., Largo, D., & Tan, S. H. (2009). PANGLAO 2004: Investigations of the marine species richness in The Philippines. *Raffles Bulletin of Zoology, Supplement*, 20, 1–19.
- Boyer, F., Mercier, C., Bonin, A., Le Bras, Y., Taberlet, P., & Coissac, E. (2016). Obitools: A unix-inspired software package for DNA metabarcoding. *Molecular Ecology Resources*, 16(1), 176–182.
- Brandl, S. J., Goatley, C. H., Bellwood, D. R., & Tornabene, L. (2018). The hidden half: Ecology and evolution of cryptobenthic fishes on coral reefs. *Biological Reviews*, 93(4), 1846–1873.
- Brandl, S. J., Tornabene, L., Goatley, C. H., Casey, J. M., Morais, R. A., Côté, I. M., Baldwin, C. C., Parravicini, V., Schietekat, N. M., & Bellwood, D. R. (2019). Demographic dynamics of the smallest marine vertebrates fuel coral reef ecosystem functioning. *Science*, 364(6446), 1189–1192.
- Brandt, M. E., Olinger, L. K., Chaves-Fonnegra, A., Olson, J. B., & Gochfeld, D. J. (2019). Coral recruitment is impacted by the presence of a sponge community. *Marine Biology*, 166(4), 1–13.
- Callahan, B. J., McMurdie, P. J., Rosen, M. J., Han, A. W., Johnson, A. J. A., & Holmes, S. P. (2016). DADA2: High-resolution sample inference from Illumina amplicon data. *Nature Methods*, 13(7), 581–583.
- Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K., & Madden, T. L. (2009). BLAST+: Architecture and applications. *BMC Bioinformatics*, 10(1), 1–9.

- Caporaso, J. G., Lauber, C. L., Walters, W. A., Berg-Lyons, D., Huntley, J., Fierer, N., Owens, S. M., Betley, J., Fraser, L., Bauer, M., Gormley, N., Gilbert, J. A., Smith, G., & Knight, R. (2012). Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms. *The ISME Journal*, 6(8), 1621–1624.
- Carugati, L., Corinaldesi, C., Dell'Anno, A., & Danovaro, R. (2015). Metagenetic tools for the census of marine meiofaunal biodiversity: An overview. *Marine Genomics*, 24, 11–20.
- Carvalho, S., Aylagas, E., Villalobos, R., Kattan, Y., Berumen, M., & Pearman, J. K. (2019). Beyond the visual: Using metabarcoding to characterize the hidden reef cryptobiome. *Proceedings of the Royal Society of London. Series B, Biological sciences*, 286(1896), 20182697.
- Casey, J. M., Ransome, E., Collins, A. G., Mahardini, A., Kurniasih, E. M., Sembiring, A., Schiettekatte, N. M., 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(6), 1229–1246.
- Chang, J. J. M., Ip, Y. C. A., Bauman, A. G., & Huang, D. (2020). MinION-in-ARMS: Nanopore sequencing to expedite barcoding of specimen-rich macrofaunal samples from Autonomous Reef Monitoring Structures. *Frontiers in Marine Science*, 7, 448.
- Chang, J. J. M., Ip, Y. C. A., Ng, C. S. L., & Huang, D. (2020). Takeaways from mobile DNA barcoding with BentoLab and MinION. *Genes*, 11(10), 1121.
- Charpy, L., Casareto, B. E., Langlade, M. J., & Suzuki, Y. (2012). Cyanobacteria in coral reef ecosystems: A review. *Journal of Marine Biology*, 2012, 1–9. <https://doi.org/10.1155/2012/259571>
- Chong-Seng, K. M., Mannering, T. D., Pratchett, M. S., Bellwood, D. R., & Graham, N. A. (2012). The influence of coral reef benthic condition on associated fish assemblages. *PLoS One*, 7(8), e42167.
- Costello, M. J., & Wilson, S. P. (2011). Predicting the number of known and unknown species in European seas using rates of description. *Global Ecology and Biogeography*, 20(2), 319–330.
- Cotténe, K. (2005). Integrating environmental and spatial processes in ecological community dynamics. *Ecology Letters*, 8(11), 1175–1182.
- Danovaro, R., Carugati, L., Berzano, M., Cahill, A. E., Carvalho, S., Chenuil, A., Corinaldesi, C., Cristina, S., David, R., Dell'Anno, A., Dzhembekova, N., Garces, E., Gasol, J. M., Priscila, G., Féral, J.-P., Ferrera, I., Forster, R. M., Kurekin, A. A., Rastelli, E., ... Borja, A. (2016). Implementing and innovating marine monitoring approaches for assessing marine environmental status. *Frontiers in Marine Science*, 3, 213.
- David, R., Uyarra, M. C., Carvalho, S., Anlauf, H., Borja, A., Cahill, A. E., Carugati, L., Danovaro, R., De Jode, A., Féral, J. P., Guillemain, D., Martired, M. L., Thierry De Ville D'Avray, L., 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.
- Davis, N. M., Proctor, D. M., Holmes, S. P., Relman, D. A., & Callahan, B. J. (2018). Simple statistical identification and removal of contaminant sequences in marker-gene and metagenomics data. *Microbiome*, 6(1), 1–14.
- De Goeij, J. M., Van Oevelen, D., Vermeij, M. J., Osinga, R., Middelburg, J. J., De Goeij, A. F., & Admiraal, W. (2013). Surviving in a marine desert: The sponge loop retains resources within coral reefs. *Science*, 342(6154), 108–110.
- Duhamel, S., & Jacquet, S. (2006). Flow cytometric analysis of bacteria- and virus-like particles in lake sediments. *Journal of Microbiological Methods*, 64(3), 316–332.
- Enochs, I. C., & Manzello, D. P. (2012). Responses of cryptozoan species richness and trophic potential to coral reef habitat degradation. *Diversity*, 4(1), 94–104.
- Enochs, I. C., Toth, L. T., Brandtneris, V. W., Afflerbach, J. C., & Manzello, D. P. (2011). Environmental determinants of motile cryptozoa on an eastern Pacific coral reef. *Marine Ecology Progress Series*, 438, 105–118.
- Fabricius, K. E., De'ath, G., Noonan, S., & Uthicke, S. (2014). Ecological effects of ocean acidification and habitat complexity on reef-associated macroinvertebrate communities. *Proceedings of the Royal Society of London. Series B, Biological sciences*, 281(1775), 20132479.
- Fisher, R., Knowlton, N., Brainard, R. E., & Caley, M. J. (2011). Differences among major taxa in the extent of ecological knowledge across four major ecosystems. *PLoS One*, 6(11), e26556.
- Fraser, K. M., Stuart-Smith, R. D., Ling, S. D., & Edgar, G. J. (2021). High biomass and productivity of epifaunal invertebrates living amongst dead coral. *Marine Biology*, 168(7), 1–12.
- Garren, M., & Azam, F. (2012). New directions in coral reef microbial ecology. *Environmental Microbiology*, 14(4), 833–844.
- Gibson, R., Atkinson, R., Gordon, J., Smith, I., & Hughes, D. (2011). Coral-associated invertebrates: Diversity, ecological importance and vulnerability to disturbance. *Oceanography and Marine Biology*, 49, 43–104.
- Gin, K. Y. H., Zhang, S., & Lee, Y. K. (2003). Phytoplankton community structure in Singapore's coastal waters using HPLC pigment analysis and flow cytometry. *Journal of Plankton Research*, 25(12), 1507–1519.
- Glynn, P. W., & Enochs, I. C. (2011). Invertebrates and their roles in coral reef ecosystems. In *Coral reefs: An ecosystem in transition* (pp. 273–325). Springer.
- Graham, N. A. J., & Nash, K. L. (2013). The importance of structural complexity in coral reef ecosystems. *Coral Reefs*, 32(2), 315–326.
- 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. (2012). The protist ribosomal reference database (PR2): A catalogue of unicellular eukaryote small sub-unit rRNA sequences with curated taxonomy. *Nucleic Acids Research*, 41(D1), D597–D604.
- Hoeksema, B. W. (2007). Delineation of the Indo-Malayan Centre of maximum marine biodiversity: The coral triangle. In *Biogeography, time, and place: Distributions, barriers, and islands* (pp. 117–178). Springer.
- Hsieh, T. C., Ma, K. H., & Chao, A. (2016). iNEXT: An R package for rarefaction and extrapolation of species diversity (hill numbers). *Methods in Ecology and Evolution*, 7(12), 1451–1456.
- Huang, D., Goldberg, E. E., Chou, L. M., & Roy, K. (2018). The origin and evolution of coral species richness in a marine biodiversity hotspot. *Evolution*, 72(2), 288–302.
- Huang, D., Tun, K. P., Chou, L. M., & Todd, P. A. (2009). An inventory of zoanthellate scleractinian corals in Singapore, including 33 new records. *Raffles Bulletin of Zoology*, 22(22), 69–80.
- Humphreys, A. F., Halfar, J., Ingle, J. C., Manzello, D., Reymond, C. E., Westphal, H., & Riegl, B. (2018). Effect of seawater temperature, pH, and nutrients on the distribution and character of low abundance shallow water benthic foraminifera in the Galápagos. *PLoS One*, 13(9), e0202746.
- Idjadi, J. A., & Edmunds, P. J. (2006). Scleractinian corals as facilitators for other invertebrates on a Caribbean reef. *Marine Ecology Progress Series*, 319, 117–127.
- Ip, Y. C. A., Chang, J. J. M., Lim, K. K., Jaafar, Z., Wainwright, B. J., & Huang, D. (2021). Seeing through sedimented waters: Environmental DNA reduces the phantom diversity of sharks and rays in turbid marine habitats. *BMC Ecology and Evolution*, 21(166), 166.
- Ip, Y. C. A., Tay, Y. C., Chang, J. J. M., Ang, H. P., Tun, K. P. P., Chou, L. M., Huang, D., & Meier, R. (2021). Seeking life in sedimented waters: Environmental DNA from diverse habitat types reveals ecologically significant species in a tropical marine environment. *Environmental DNA*, 3(3), 654–668.
- Ip, Y. C. A., Tay, Y. C., Gan, S. X., Ang, H. P., Tun, K., Chou, L. M., Huang, D., & Meier, R. (2019). From marine park to future genomic

- observatory? Enhancing marine biodiversity assessments using a biocode approach. *Biodiversity Data Journal*, 7, e46833.
- Jaccard, P. (1901). Étude comparative de la distribution florale dans une portion des Alpes et des Jura. *Bulletin De La Société Vaudoise Des Sciences Naturelles*, 37, 547–579.
- Jones, E. G., Suetrong, S., Sakayaroj, J., Bahkali, A. H., Abdel-Wahab, M. A., Boekhout, T., & Pang, K. L. (2015). Classification of marine ascomycota, basidiomycota, blastocladiomycota and chytridiomycota. *Fungal Diversity*, 73(1), 1–72.
- Jousset, A., Bienhold, C., Chatzinotas, A., Gallien, L., Gobet, A., Kurm, V., Küsel, K., Rillig, M. C., Rivett, D. W., Salles, J. F., Van Der Heijden, M. G., Youssef, N. H., Zhang, X., Wei, Z., & Hol, W. G. (2017). Where less may be more: How the rare biosphere pulls ecosystems strings. *The ISME Journal*, 11(4), 853–862.
- Kaboré, O. D., Godreuil, S., & Drancourt, M. (2020). Planctomycetes as host-associated bacteria: A perspective that holds promise for their future isolations, by mimicking their native environmental niches in clinical microbiology laboratories. *Frontiers in Cellular and Infection Microbiology*, 10, 729.
- Kelly, L. W., Williams, G. J., Barott, K. L., Carlson, C. A., Dinsdale, E. A., Edwards, R. A., Haas, A. F., Haynes, M., Lim, Y. W., McDole, T., Nelson, C. E., Sala, E., Sandin, S. A., Smith, J. E., Vermeij, M. J. A., Youle, M., & Rohwer, F. (2014). Local genomic adaptation of coral reef-associated microbiomes to gradients of natural variability and anthropogenic stressors. *Proceedings of the National Academy of Sciences of the United States of America*, 111(28), 10227–10232.
- Kling, J. D., Lee, M. D., Fu, F., Phan, M. D., Wang, X., Qu, P., & Hutchins, D. A. (2020). Transient exposure to novel high temperatures reshapes coastal phytoplankton communities. *The ISME Journal*, 14(2), 413–424.
- Knowlton, N. (2001a). Coral reef biodiversity--habitat size matters. *Science*, 292(5521), 1493–1495.
- Knowlton, N. (2001b). The future of coral reefs. *Proceedings of the National Academy of Sciences of the United States of America*, 98(10), 5419–5425.
- Komyakova, V., Munday, P. L., & Jones, G. P. (2013). Relative importance of coral cover, habitat complexity and diversity in determining the structure of reef fish communities. *PLoS One*, 8(12), e83178.
- Kramer, M. J., Bellwood, D. R., Taylor, R. B., & Bellwood, O. (2017). Benthic crustacea from tropical and temperate reef locations: Differences in assemblages and their relationship with habitat structure. *Coral Reefs*, 36(3), 971–980.
- Lane, D. J. (1991). 16S/23S rRNA sequencing. In E. Stackebrandt & M. Goodfellow (Eds.), *Nucleic acid techniques in bacterial systematic* (pp. 115–175). John Wiley and Sons.
- Laporte, M., Reny-Nolin, E., Chouinard, V., Hernandez, C., Normandea, E., Bougas, B., Côté, C., Behmel, S., & Bernatchez, L. (2021). Proper environmental DNA metabarcoding data transformation reveals temporal stability of fish communities in a dendritic river system. *Environmental DNA*, 3(5), 1007–1022.
- 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.
- Leray, M., Ho, S. L., Lin, I. J., & Machida, R. J. (2018). MIDORI server: A webserver for taxonomic assignment of unknown metazoan mitochondrial-encoded sequences using a curated database. *Bioinformatics*, 34(21), 3753–3754.
- Leray, M., & Knowlton, N. (2015). DNA barcoding and metabarcoding of standardized samples reveal patterns of marine benthic diversity. *Proceedings of the National Academy of Sciences of the United States of America*, 112(7), 2076–2081.
- 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), 1–14.
- Limberger, R., & Wickham, S. A. (2012). Transitory versus persistent effects of connectivity in environmentally homogeneous metacommunities. *PLoS One*, 7, e44555.
- Lobo, J., Costa, P. M., Teixeira, M. A., Ferreira, M. S., Costa, M. H., & Costa, F. O. (2013). Enhanced primers for amplification of DNA barcodes from a broad range of marine metazoans. *BMC Ecology*, 13(1), 1–8.
- Logares, R., Audic, S., Bass, D., Bittner, L., Boutte, C., Christen, R., Claverie, J. M., Decelle, J., Dolan, J. R., Dunthorn, M., Edvardsen, B., Gobet, A., Kooistra, W. H. C. F., Mahé, F., Not, F., Ogata, H., Pawłowski, J., Pernice, M. C., Romac, S., ... Massana, R. (2014). Patterns of rare and abundant marine microbial eukaryotes. *Current Biology*, 24(8), 813–821.
- Machida, R. J., Leray, M., Ho, S. L., & Knowlton, N. (2017). Metazoan mitochondrial gene sequence reference datasets for taxonomic assignment of environmental samples. *Scientific Data*, 4(1), 1–7.
- Magurran, A. E., & Henderson, P. A. (2003). Explaining the excess of rare species in natural species abundance distributions. *Nature*, 422(6933), 714–716.
- Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal*, 17(1), 10–12.
- McMurdie, P. J., & Holmes, S. (2013). Phyloseq: An R package for reproducible interactive analysis and graphics of microbiome census data. *PLoS One*, 8(4), e61217.
- McWilliam, M., Chase, T. J., & Hoogenboom, M. O. (2018). Neighbor diversity regulates the productivity of coral assemblages. *Current Biology*, 28(22), 3634–3639.
- Medlin, L., Elwood, H. J., Stickel, S., & Sogin, M. L. (1988). The characterization of enzymatically amplified eukaryote 16S-like rRNA-coding regions. *Gene*, 71(2), 491–499.
- Meier, R., Wong, W., Srivathsan, A., & Foo, M. (2016). \$1 DNA barcodes for reconstructing complex phenomes and finding rare species in specimen-rich samples. *Cladistics*, 32(1), 100–110.
- Molen, J. V. D., García-García, L. M., Whomersley, P., Callaway, A., Posen, P. E., & Hyder, K. (2018). Connectivity of larval stages of sedentary marine communities between hard substrates and offshore structures in the North Sea. *Scientific Reports*, 8(1), 1–14.
- Murali, A., Bhargava, A., & Wright, E. S. (2018). IDTAXA: A novel approach for accurate taxonomic classification of microbiome sequences. *Microbiome*, 6(1), 1–14.
- Nelson, H. R., Kuempel, C. D., & Altieri, A. H. (2016). The resilience of reef invertebrate biodiversity to coral mortality. *Ecosphere*, 7(7), e01399.
- Ng, C. S. L., Chan, Y. K. S., Nguyen, N. T. H., Kikuzawa, Y. P., Sam, S. Q., Toh, T. C., Mock, A. Y. J., Chou, L. M., & Huang, D. (2021). Coral community composition and carbonate production in an urbanized seascape. *Marine Environmental Research*, 168, 105322.
- Oh, R. M., Bollati, E., Maithani, P., Huang, D., & Wainwright, B. J. (2021). The microbiome of the reef macroalga *Sargassum ilicifolium* in Singapore. *Microorganisms*, 9(5), 898.
- Oksanen, J., Blanchet, F. G., Kindt, R., Legendre, P., Minchin, P. R., O'hara, R. B., Simpson, G. L., Solymos, P., Stevens, M. H. H., Wagner, H., & Oksanen, M. J. (2013). Package 'vegan'. *Community Ecology Package*, Version, 2(9), 1–295.
- Olson, J. B., & Kellogg, C. A. (2010). Microbial ecology of corals, sponges, and algae in mesophotic coral environments. *FEMS Microbiology Ecology*, 73(1), 17–30.
- Palomino-Alvarez, L. A., Vital, X. G., Castillo-Cupul, R. E., Suárez-Mozo, N. Y., Ugalde, D., Cervantes-Campero, G., Muciño-Reyes, M. R., Homá-Canché, P., Hernández-Díaz, Y. Q., Sotelo-Casas, R., García-González, M., Avedaño-Peláez, Y. A., Hernández-González, A., Paz-Ríos, C. E., Lizaola-Guillermo, J. M., García-Venegas, M., Dávila-Jiménez, Y., Ortigosa, D., Hidalgo, G., ... Guerra-Castro, E. J. (2021). Evaluation of the use of Autonomous

- Reef Monitoring Structures (ARMS) for describing the species diversity of two coral reefs in the Yucatan Peninsula, Mexico. *Diversity*, 13(11), 579.
- Pearman, J. K., Anlauf, H., Irigoién, X., & Carvalho, S. (2016). Please mind the gap—visual census and cryptic biodiversity assessment at Central Red Sea coral reefs. *Marine Environmental Research*, 118, 20–30.
- 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.
- Pearman, J. K., Chust, G., Aylagas, E., Villarino, E., Watson, J. R., Chenail, A., Borja, A., Cahill, A. E., Carugati, L., Danovaro, R., David, R., Irigoién, X., Mendibil, I., Moncheva, S., Rodríguez-Ezpeleta, N., Uyarra, M. C., & Carvalho, S. (2020). Pan-regional marine benthic cryptobiome biodiversity patterns revealed by metabarcoding Autonomous Reef Monitoring Structures. *Molecular Ecology*, 29(24), 4882–4897.
- 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), 1–17.
- Peixoto, R. S., Rosado, P. M., Leite, D. C. D. A., Rosado, A. S., & Bourne, D. G. (2017). Beneficial microorganisms for corals (BMC): Proposed mechanisms for coral health and resilience. *Frontiers in Microbiology*, 8, 341.
- 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.
- Plaisance, L., Matteson, K., Fabricius, K., Drovetski, S., Meyer, C., & Knowlton, N. (2021). Effects of low pH on the coral reef cryptic invertebrate communities near CO₂ vents in Papua New Guinea. *PLoS One*, 16(12), e0258725.
- Rabbani, G., Huang, D., & Wainwright, B. J. (2021). The mycobiome of *Pocillopora acuta* in Singapore. *Coral Reefs*, 40, 1419–1427.
- 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.
- R Core Team (2021). *RStudio: integrated development for R*. RStudio, Inc. <https://www.rstudio.com>
- Ribes, M., Coma, R., Atkinson, M. J., & Kinzie, R. A., III. (2005). Sponges and ascidians control removal of particulate organic nitrogen from coral reef water. *Limnology and Oceanography*, 50(5), 1480–1489.
- Richter, C., Wunsch, M., Rasheed, M., Koëtter, I., & Badran, M. I. (2001). Endoscopic exploration of Red Sea coral reefs reveals dense populations of cavity-dwelling sponges. *Nature*, 413(6857), 726–730.
- Roberts, C. M., McClean, C. J., Veron, J. E., Hawkins, J. P., Allen, G. R., McAllister, D. E., Mittermeier, C. G., Schueler, F. W., Spalding, M., Wells, F., Vynne, C., & Werner, T. B. (2002). Marine biodiversity hotspots and conservation priorities for tropical reefs. *Science*, 295(5558), 1280–1284.
- Rogers, A., Blanchard, J. L., Newman, S. P., Dryden, C. S., & Mumby, P. J. (2018). High refuge availability on coral reefs increases the vulnerability of reef-associated predators to overexploitation. *Ecology*, 99(2), 450–463.
- Rusch, A., Hannides, A. K., & Gaidos, E. (2009). Diverse communities of active bacteria and archaea along oxygen gradients in coral reef sediments. *Coral Reefs*, 28(1), 15–26.
- Shen, W., Le, S., Li, Y., & Hu, F. (2016). SeqKit: A cross-platform and ultrafast toolkit for FASTA/Q file manipulation. *PLoS One*, 11(10), e0163962.
- Siegel, D. A., Mitarai, S., Costello, C. J., Gaines, S. D., Kendall, B. E., Warner, R. R., & Winters, K. B. (2008). The stochastic nature of larval connectivity among nearshore marine populations. *Proceedings of the National Academy of Sciences of the United States of America*, 105(26), 8974–8979.
- Srivathsan, A., Hartop, E., Puniamoorthy, J., Lee, W. T., Kutty, S. N., Kurina, O., & Meier, R. (2019). Rapid, large-scale species discovery in hyperdiverse taxa using 1D MinION sequencing. *BMC Biology*, 17(1), 1–20.
- Srivathsan, A., Sha, J. C. M., Vogler, A. P., & Meier, R. (2015). Comparing the effectiveness of metagenomics and metabarcoding for diet analysis of a leaf-feeding monkey (*Pygathrix nemaeus*). *Molecular Ecology Resources*, 15(2), 250–261.
- Stoeck, T., Bass, D., Nebel, M., Christen, R., Jones, M. D., Breiner, H. W., & Richards, T. A. (2010). Multiple marker parallel tag environmental DNA sequencing reveals a highly complex eukaryote community in marine anoxic water. *Molecular Ecology*, 19, 21–31.
- Tay, Y. C., Todd, P. A., Rosshaug, P. S., & Chou, L. M. (2012). Simulating the transport of broadcast coral larvae among the Southern Islands of Singapore. *Aquatic Biology*, 15(3), 283–297.
- Trüper, H. G. (1999). How to name a prokaryote?: Etymological considerations, proposals and practical advice in prokaryote nomenclature. *FEMS Microbiology Reviews*, 23(2), 231–249.
- Valencia, B., & Giraldo, A. (2021). Coral reef cryptic invertebrates across a gradient of coral cover in Isla Gorgona, Eastern Tropical Pacific Off Colombia. 2021, 68, MDPI Proceedings.
- Valliappan, K., Sun, W., & Li, Z. (2014). Marine actinobacteria associated with marine organisms and their potentials in producing pharmaceutical natural products. *Applied Microbiology and Biotechnology*, 98(17), 7365–7377.
- Veron, J., Stafford-Smith, M., DeVantier, L., & Turak, E. (2015). Overview of distribution patterns of zooxanthellate Scleractinia. *Frontiers in Marine Science*, 1, 81.
- Veron, J. E., Devantier, L. M., Turak, E., Green, A. L., Kininmonth, S., Stafford-Smith, M., & Peterson, N. (2009). Delineating the coral triangle, Galaxea. *Journal of Coral Reef Studies*, 11(2), 91–100.
- Vicente, J., Webb, M. K., Paulay, G., Rakchai, W., Timmers, M. A., Jury, C. P., Bahr, K., & Toonen, R. J. (2021). Unveiling hidden sponge biodiversity within the Hawaiian reef cryptozoa. *Coral Reefs*, 41(3), 727–742.
- Volkov, I., Banavar, J. R., Hubbell, S. P., & Maritan, A. (2007). Patterns of relative species abundance in rainforests and coral reefs. *Nature*, 450(7166), 45–49.
- Wainwright, B. J., Afiq-Rosli, L., Zahn, G. L., & Huang, D. (2019). Characterisation of coral-associated bacterial communities in an urbanised marine environment shows strong divergence over small geographic scales. *Coral Reefs*, 38(6), 1097–1106.
- Wainwright, B. J., Bauman, A. G., Zahn, G. L., Todd, P. A., & Huang, D. (2019). Characterization of fungal biodiversity and communities associated with the reef macroalgae *Sargassum ilicifolium* reveals fungal community differentiation according to geographic locality and algal structure. *Marine Biodiversity*, 49(6), 2601–2608.
- Wang, G., Wang, X., Liu, X., Li, Q. (2012). Diversity and biogeochemical function of planktonic fungi in the ocean. In C. Raghukumar (Ed.), *Biology of Marine Fungi. Progress in Molecular and Subcellular Biology*, (vol. 53, pp. 71–88). Springer.
- Wang, Q., Garrity, G. M., Tiedje, J. M., & Cole, J. R. (2007). Naive Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. *Applied and Environmental Microbiology*, 73(16), 5261–5267.
- Wang, W. Y., Srivathsan, A., Foo, M., Yamane, S. K., & Meier, R. (2018). Sorting specimen-rich invertebrate samples with cost-effective NGS barcodes: Validating a reverse workflow for specimen processing. *Molecular Ecology Resources*, 18(3), 490–501.

- Wells, F. E., Tan, K. S., Todd, P. A., Jaafar, Z., & Yeo, D. C. (2019). A low number of introduced marine species in the tropics: A case study from Singapore. *Management of Biological Invasions*, 10(1), 23–45.
- Wong, J. S., Chan, Y. S., Ng, C. L., Tun, K. P., Darling, E. S., & Huang, D. (2018). Comparing patterns of taxonomic, functional and phylogenetic diversity in reef coral communities. *Coral Reefs*, 37(3), 737–750.
- Wright, E. S., Yilmaz, L. S., & Noguera, D. R. (2012). DECIPHER, a search-based approach to chimera identification for 16S rRNA sequences. *Applied and Environmental Microbiology*, 78(3), 717–725.
- Xu, G., Zhang, W., & Xu, H. (2015). Can dispersions be used for discriminating water quality status in coastal ecosystems? A case study on biofilm-dwelling microbial eukaryotes. *Ecological Indicators*, 57, 208–214.
- Zhang, J., Kobert, K., Flouri, T., & Stamatakis, A. (2014). PEAR: A fast and accurate Illumina paired-end reAd mergeR. *Bioinformatics*, 30(5), 614–620.
- 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(1), 59–64.

SUPPORTING INFORMATION

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

How to cite this article: Ip, Y. C. A., Chang, J. J. M., Oh, R. M., Quek, Z. B. R., Chan, Y. K. S., Bauman, A. G., & Huang, D. (2022). Seq' and ARMS shall find: DNA (meta)barcoding of Autonomous Reef Monitoring Structures across the tree of life uncovers hidden cryptobiome of tropical urban coral reefs. *Molecular Ecology*, 00, 1–20. <https://doi.org/10.1111/mec.16568>