Vol. 88: 81–94, 2022
https://doi.org/10.3354/ame01985

AQUATIC MICROBIAL ECOLOGY
Aquat Microb Ecol

Published online March 31

# Biogeography of reef water microbes from within-reef to global scales

**Lei Ma[1,2], Cynthia Becker[1,2], Laura Weber[1], Chris Sullivan[3], Brian Zgliczynski[3], Stuart Sandin[3], Marilyn Brandt[4], Tyler B. Smith[4], Amy Apprill[1,*]**

[1]Woods Hole Oceanographic Institution, Woods Hole, MA 02543, USA
[2]MIT-WHOI Joint Program in Oceanography/Applied Ocean Science & Engineering, Woods Hole office, Woods Hole, MA 02543, USA
[3]Scripps Institution of Oceanography, UC San Diego, La Jolla, CA 92037, USA
[4]University of Virgin Islands, St. Thomas 00802, USVI

ABSTRACT: Seawater microorganisms play an important role in coral reef ecosystem functioning and can be influenced by biological, chemical, and physical features of reefs. As coral reefs continue to respond to environmental changes, the reef seawater microbiome has been proposed as a conservation tool for monitoring perturbations. However, the spatial variability of reef seawater microbial communities is not well studied, limiting our ability to make generalizable inferences across reefs. In order to better understand how microorganisms are distributed at multiple spatial scales, we examined seawater microbial communities in Florida Reef Tract and US Virgin Islands reef systems using a nested sampling design. On 3 reefs per reef system, we sampled seawater at regular spatial intervals close to the benthos. We assessed the microbial community composition of these waters using ribosomal RNA gene amplicon sequencing. Our analysis revealed that reef water microbial communities varied as a function of reef system and individual reefs, but communities did not differ within reefs and were not significantly influenced by benthic composition. For the reef system and inter-reef differences, abundant microbial taxa were found to be potentially useful indicators of environmental difference due to their high prevalence and variance. We further examined reef water microbial biogeography on a global scale using a secondary analysis of 5 studies, which revealed that microbial communities were more distinct with increasing geographic distance. These results suggest that biogeography is a distinguishing feature for reef water microbiomes, and that development of monitoring criteria may necessitate regionally specific sampling and analyses.

KEY WORDS: 16S rDNA · Biogeography · Coral reef communities · Free-living bacteria · Marine microbial ecology · Meso-scale spatial variations · Secondary analysis · Microbial diversity



## 1. INTRODUCTION

Coral reefs are currently experiencing significant challenges due to global and local factors (Hughes et al. 2017). Among them, climate change and ocean acidification affect corals worldwide while stressors such as human impacts and disease outbreaks are more localized. These crises are driving the development of new management and conservation strategies to preserve and monitor reef biodiversity. Awareness of the coral as a holobiont — an assemblage of a host and all of its associated symbiotic microorganisms (Knowlton & Rohwer 2003, Rosenberg et al. 2007) — has spurred research into establishing microbial solutions to reef stress, such as coral probiotics and microbiome-based monitoring (Glasl et al. 2017, Peixoto et al. 2017). In particular, a holistic characterization of microbes in coral reefs will

*Corresponding author: aapprill@whoi.edu

aid in predicting reef resilience and environmental threats (Kelly et al. 2018).

Reefs harbor many distinct niches for bacterial and archaeal communities, including corals, sponges, sediments, and the water column itself (Tout et al. 2014, McDevitt-Irwin et al. 2017). Free-living water column microbes, residing above the reef substrate, are influenced by hydrological conditions (Sweet et al. 2010, Becker et al. 2020), general benthic community composition (Haas et al. 2011, Kelly et al. 2014), local nutrient regimes (van Duyl & Gast 2001, Nelson et al. 2011), and temporal dynamics (Weber & Apprill 2020, Becker et al. 2020). Combined, these influences cause reef-associated seawater microbiomes to be readily distinguishable between reefs as well as between zones within a reef (Jeffries et al. 2015, Salerno et al. 2016, Frade et al. 2020). Microbial communities can be powerful indicators of reef health and environmental conditions (Glasl et al. 2017). Indeed, as reefs transition from coral- to algae-dominated, the exudates released from the benthos also likely shift, causing increased heterotrophy and decreased oligotrophy in the seawater microbiome (Haas et al. 2011, Nelson et al. 2013). In a process called microbialization, the heterotroph-dominated microbial community further depresses coral growth and encourages the growth of algae (Haas et al. 2016, Kelly et al. 2018). This microbial phase shift may be an important process to monitor in at-risk reefs. Additionally, reef microorganisms respond rapidly to nutrient and temperature fluctuations, potentially providing a sensitive and non-invasive diagnostic or predictive tool for perturbations that may provide knowledge prior to visible reef changes (Glasl et al. 2019, Becker et al. 2020).

Implementation of large-scale reef water monitoring efforts for reef microorganisms is partially limited by our understanding of reef seawater microbial diversity across spatial scales (Bourne et al. 2016, Glasl et al. 2017). Biogeographic patterns of coral reef microbial assemblages have been found at a variety of spatial scales. Small-scale patterns such as within a coral skeleton (Marcelino et al. 2018), in the boundary layer overlying the coral mucus (Weber et al. 2019), and in micro-habitats generated by coral structures (Schöttner et al. 2012) highlight potential mechanisms affecting reef microbial composition, but may not represent the state of an entire reef. On the other hand, studies and models of marine microbial distribution at the scale of oceans (Amend et al. 2013, Hellweger et al. 2014) provide insight into the global drivers of microbial abundance but are not specific to the unique environments of reefs. Therefore, a better understanding of the biogeography of coral reef seawater microbes across distinct spatial scales is warranted.

The goal of this study is to understand the variability of coral reef seawater microbial communities across different spatial scales. We examined this question in 2 parts. For the first part, we examined reef water microbial communities within and between 2 reef systems to understand the influence of both reef and reef benthic composition on microbial diversity (Fig. 1A). Secondly, in order to quantify the impact of larger geographic distances on reef water microbial communities, we conducted a secondary analysis of aggregated 16S rRNA gene sequences from 5 studies that used similar sampling methodologies (Fig. 1B, see Table S1 in Supplement 2 at www.int-res.com/articles/suppl/a088p081_supp2.xlsx for supplementary tables). We predicted that microbial community structure would differ primarily on the scale of individual reefs and secondarily on underlying benthic structure. Additionally, we expected the secondary analysis to recapitulate the individual reef and reef system-based biogeographic patterns seen in the Florida and Virgin Islands systems on a more global scale.

## 2. MATERIALS AND METHODS

### 2.1. Florida and Virgin Islands transects: sampling

The first part of this study took place in 2 reef systems: the Florida Reef Tract in the Florida Keys (Fl) in June 2019 and off the southern coast of St. Thomas in the US Virgin Islands (VI) in February 2020. A total of 3 reefs were sampled in Fl; the northernmost reef was Biscayne, located within the boundaries of Biscayne National Park, the reef Grecian was located at the Grecian Rocks reef off the coast of Key Largo, and the reef Dry Tortugas was located within Dry Tortugas National Park. All reefs in Fl were forereefs within the barrier reef. Similarly, 3 reefs were sampled in VI; Brewers Bay and Black Point were forereef zones on fringing reefs a few hundred meters from the coast while Flat Cay was a fringing reef located near an uninhabited island named Flat Cay about 2 km off the coast (Table 1). Average sampling depth was between 5.0 and 7.1 m, with the exception of the Dry Tortugas, which was deeper with an average depth of 18.0 m. Due to the difference in season of sampling, the average temperature in the Fl reefs was slightly higher than in the VI reefs (28.7 and 26.9°C, respectively).

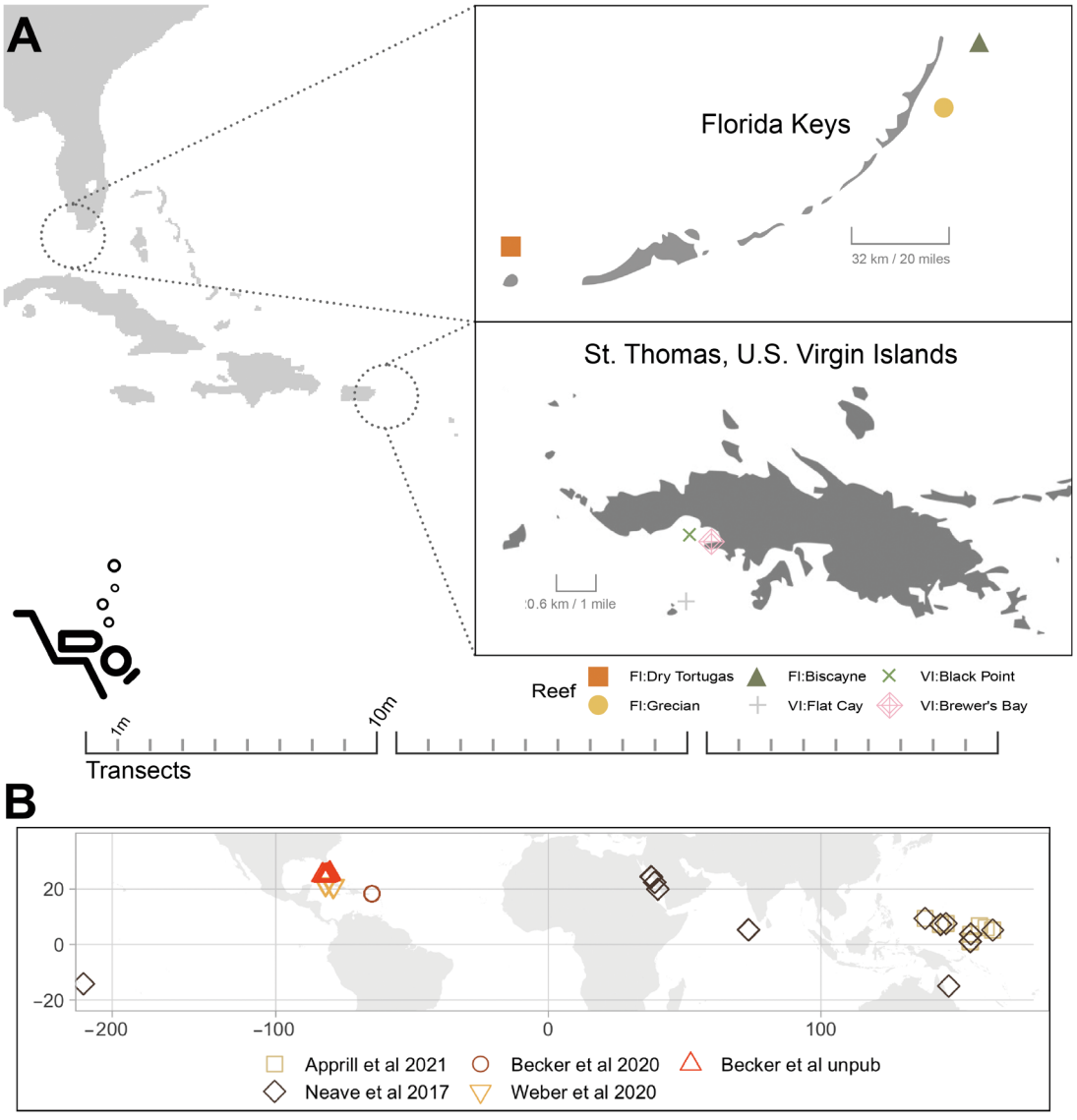


Fig. 1. (A) Sampling locations and method for the transect-based study. Six reefs were surveyed across 2 reef systems. At each reef, three 10 m transects were laid, and divers used a syringe to sample seawater just above the benthos at 1 m intervals. (B) Locations of the 5 studies in the secondary analysis, spanning many major reef systems across the globe. Collections for all studies were performed by the same lab group using nearly identical techniques in the field and lab

Table 1. Florida reef tract and St. Thomas sampling locations. ND: no data

| Reef System | Reef | Transect | Sample | Avg. depth (m) | Latitude, longitude | Date | Temp (°C) | Reef type |
|---|---|---|---|---|---|---|---|---|
| Florida Keys | Dry Tortugas | 3 | 30 | 18.0 | 24.722° N, 82.828° W | 7 Jun 2019 | 28.7 | Forereef |
| Florida Keys | Grecian | 3 | 30 | 6.7 | 25.110° N, 80.303° W | 14 Jun 2019 | 28.7 | Forereef |
| Florida Keys | Biscayne | 1 | 9 | 5.2 | 25.386° N, 80.162° W | 17 Jun 2019 | 28.7 | Forereef |
| St. Thomas, USVI | Flat Cay | 3 | 30 | 7.1 | 18.316° N, 64.987° W | 12 Feb 2020 | 26.8 | Fringing |
| St. Thomas, USVI | Black Point | 3 | 30 | 7.0 | 18.344° N, 64.986° W | 12 Feb 2020 | 27.0 | Fringing |
| St. Thomas, USVI | Brewer's Bay | 3 | 27 | 5.9 | 18.343° N, 64.980° W | 12 Feb 2020 | ND | Fringing |

At each reef, three 10 m transects were taken by laying down a 10 m weighted line that was marked every 1 m. Water samples were taken by a diver using a 60 or 100 ml syringe positioned approximately 5 cm above the benthos at each 1 m line. The transects were laid haphazardly but did not intersect with each other. Because of inclement conditions, only 1 transect was collected at the Biscayne reef. At the Fl reefs, benthic composition — represented by percent cover of coral skeleton, crustose coralline algae, cyanobacteria, hard coral, macroalgae, non-biological, other invertebrates, soft coral, sponge, and turf algae — was determined using large-area imagery collected from 10 × 10 m area plots. All transects were placed within these 100 $m^2$ plots. Stratified random points (2500) were dropped across the reef area and classified to generate reef-wide cover estimates; see full methods in Fox et al. (2019). At the VI reefs, benthic composition was recorded at the precise location of each syringe sample using a video survey of the transect line as well as being noted in writing by a diver during sampling. Video and written records were cross-referenced, and each sample was then classified into a single category as algae, dead coral, live coral, rock, sand, sponge, or undetermined.

To capture the seawater microbial community, 60 ml of the seawater was filtered through a 0.22 µm Supor filter (25 mm; Pall Corporation). The water volume of 60 ml has previously been found to be comparable to larger volumes (1–2 l) for characterizing seawater microbial communities using amplicon sequencing (Weber et al. 2019). Filters were placed in 2 ml cryovials, flash-frozen in a liquid nitrogen dry shipper, and processed upon returning to Woods Hole, MA.

### 2.2. Fl and VI transects: DNA extraction, PCR amplification, and sequencing

DNA was extracted from the filters using the DNeasy PowerBiofilm Kit (Qiagen) according to the manufacturer's protocols. Seven DNA extraction controls, consisting of unused 0.22 µm filters, were processed alongside samples. Extracted DNA was quantified using the Qubit 2.0 fluorometer HS dsDNA assay (ThermoFisher Scientific). Primers 515FY (Parada et al. 2016) and 806RB (Apprill et al. 2015) containing Illumina overhang adapter sequences were used to amplify the V4 region of the small subunit (SSU) rRNA gene in bacteria and archaea. PCR reactions contained 14.75 µl molecular-grade water, 5 µl GoTaq Flexi 5× buffer (Promega Corporation), 2.5 µl of 25 mM $MgCl_2$, 1 µl of 10 mM dNTPs, 1 µl of 10 mM forward and reverse primers, 0.5 µl GoTaq DNA polymerase (Promega), and 1 µl of DNA template. Three PCR controls consisting of 1 µl of PCR-grade water as template were also included as well as microbial genomic DNA from a Human Microbiome Project mock community (BEI Resources, National Institute of Allergy and Infectious Diseases [NIAID], National Institutes of Health [NIH] as part of the Human Microbiome Project: 'Genomic DNA from Microbial Mock Community B [Even, Low Concentration], v5.1L, for 16S rRNA Gene Sequencing, HM-782D'). The first stage PCR conditions were 28 cycles (95°C for 20 s, 55°C for 20 s, 72°C for 5 min) with a 2 min 95°C hot start and 10 min 72°C final elongation. PCR products were screened for quality using gel electrophoresis and purified using the MinElute PCR purification kit (Qiagen). PCR products were then barcoded using the Nextera XT Index Kit v2 set A primers (Illumina) using the following conditions: 8 cycles (95°C for 30 s, 55°C for 30 s, 72°C for 30 s) with a 3 min 95° hot start and 5 min 72°C final elongation. Barcoded products were purified as above, and concentrations of the purified products were assessed using the HS dsDNA assay on the Qubit 2.0 fluorometer (ThermoFisher Scientific). Products were diluted with Tris HCl to 5 nM before being pooled randomly into 2 libraries. The libraries were diluted to a final loading concentration of 50 pM with a 5% spike-in of 50 pM PhiX. The libraries were then sequenced on the iSeq 100 System (Illumina) using paired-end 150 bp reads. Data are accessible in the NCBI Sequence Read Archive under Bioproject PRJNA733652.

### 2.3. Fl and VI transects: data analysis

The code used to generate the figures and analyses in this paper is publicly available on GitHub (https://github.com/microlei/AME_biogeography_2021). Sequences were processed using the DADA2 package (v.1.12.1) in R (v.3.6.2) (Callahan et al. 2016). Due to the short length of iSeq reads (150 bp) it was not possible to merge the reads, and therefore only forward reads were used in the analysis. Forward reads were filtered using the default parameters of the function 'filterAndTrim' in DADA2 except 'trimLeft=20' (to remove the primer), 'truncLen=125', 'truncQ=2', and 'maxEE=1'. The parameter of 'truncLen' was determined after observing quality dropping during the

last 5 bp of the fastq reads. Chimera removal and amplicon sequence variant (ASV) generation was also done by DADA2. Taxonomy was assigned, without percent identity clustering, using the naive Bayesian classifier method of Wang et al. (2007) trained on the Silva SSU rRNA database (v.138) (Pruesse et al. 2007). Putative contaminant reads were identified using the prevalence method in the R package 'decontam' (v.1.4.0) (Davis et al. 2018) by using the negative controls to identify contaminants; contaminant reads were subsequently removed. Reads matching Kingdom Eukaryota or Order Chloroplast were also removed.

Our data analysis was completed in RStudio (v.1.2.5.001) (RStudio Team 2019) using, primarily, the packages 'phyloseq' (v.1.28.0) and 'vegan' (v.2.5-7) (McMurdie & Holmes 2013, Oksanen et al. 2020). Graphics were generated using 'ggplot2' (v.3.3.3) (Wickham 2016). Alpha diversity metrics were estimated using the 'estimate_richness' function in vegan, with unrarefied read counts. Differences in alpha diversity metrics were assessed using pairwise *t*-tests corrected for multiple comparisons using the Holm method (Holm 1979). To understand the variability of microbial community composition across all samples, ASV counts were transformed to relative abundances, and Bray-Curtis dissimilarity was calculated between each sample pair. Dissimilarity values were plotted using non-metric multidimensional scaling ordination (NMDS). Taxa that were differentially abundant were identified using the package 'corncob' (v.0.2.0) (Martin et al. 2021), with a false discovery rate cutoff of 0.05 using the Benjamini-Hochberg correction.

In order to account for differences in read counts arising from variances in sequencing depth and read quality as well as to improve the quality of the distance-decay analyses, the Aitchison distance was used instead of Bray-Curtis when comparing community similarity across studies in the secondary analysis and, for consistency, between samples in the transect-based study (Gloor et al. 2017, Clark et al. 2021). The package 'zCompositions' (v.1.3.4) (Palarea-Albaladejo & Martín-Fernández 2015) was used to impute zeroes before performing a centered-log-ratio (CLR) transformation on the count data. Taking the Euclidean distances of the CLR-transformed data generated the Aitchison distances. Geographic distances were calculated using the 'gdist' function in the package 'Imap' (v.1.32) (Wallace 2012), which uses the Vincenty inverse formula for ellipsoids. The 'adonis2' function in the vegan package was used to perform permutational multivariate analysis of variance (PERMANOVA) analysis (999 iterations) on the dissimilarity indices at the scales of transect (within-reef), reef, and reef system. Differences in dispersion at various spatial scales were calculated using the vegan function 'betadisper' and tested using the vegan function 'permutest', which performs a permutation test (999 permutations) of multivariate homogeneity of groups dispersions. The 'mantel' function in the vegan package was used (999 permutations) to test for correlation between the geographic distance matrix and the community similarity for both the secondary analysis and transect-based study.

### 2.4. Secondary analysis: sample information

The methods for sample collection, DNA extraction, PCR amplification, and sequencing used by the 5 studies in the secondary analysis are highly similar with small variations. The studies collected seawater from reefs at a variety of depths, ranging from surface (0.3 m) to benthic (13 m). Seawater sampling was done by filtering replicate 2 l volumes of seawater through 0.22 µm pore size, 25 mm Supor® filters using a peristaltic pump. For DNA extraction, all studies used bead beating followed by spin column purification, although the DNA extraction reagents differed. Neave et al. (2017) used the PowerPlant Pro DNA Isolation Kit (Qiagen) while Weber et al. (2020) used a sucrose-lysis with bead beating method followed by column purification with the Qiagen DNeasy Blood and Tissue Kit (Santoro et al. 2010) as well as a phenol chloroform protocol (Urakawa et al. 2010) and pooled the extracts. C. Becker et al. (unpubl. data) used the DNeasy PowerBiofilm kit (Qiagen). Full methods for these unpublished data are included in Text S1 in Supplement 1 at www.int-res.com/articles/suppl/a088p081_supp1.pdf. The remaining 2 studies used the extraction method described in Santoro et al. (2010).

All studies amplified the V4 hypervariable region of the 16S rRNA gene using the reverse primer described in Apprill et al. (2015), but 2 studies (Neave et al. 2017 and Apprill et al. 2021) used the forward primer not optimized for *Thaumarchaeota* (Caporaso et al. 2011) while the others used the forward primer described in Parada et al. (2016). All studies used the 250 bp paired-end Illumina MiSeq platform, although Weber et al. (2020) used the Fluidgm® platform (Fluidgm Corporation) for library preparation while others followed the methods described in Kozich et al. (2013). Primer choice, sequencing tech-

nology, and DNA extraction method are known to influence downstream 16S rRNA gene sequence analysis, such as in marine biofilms and seawater (Urakawa et al. 2010, Corcoll et al. 2017). A comparison of different DNA extraction techniques on aquatic samples concluded that rare taxa are more affected by differing extraction techniques, driving small but significant differences in Bray-Curtis distances (Liu et al. 2019). However, the secondary analysis is based on pairwise Aitchison distances, which are less influenced by presence/absence of individual taxa (Gloor et al. 2017), and does not seek to compare groups of samples based on distances.

### 2.5. Secondary analysis: data acquisition and processing

Raw sequence data from the 5 studies in the secondary analysis were collected from the NCBI Sequence Read Archive (SRA) and for the unpublished study, with consent from the authors. Because the transect comparisons collected in Fl and VI were sequenced with shorter reads, these samples were excluded from the secondary analysis. Using metadata from the studies, sequence files were filtered such that only samples taken from reef-associated seawater (and not controls) were included. Samples were classified based on the reef that was sampled as well as the overall reef system (Table S1). Primer sequences were removed using 'cutadapt' (Martin 2011). Sequences were processed in DADA2 as above with the parameters 'trimLeft=(20,20)', 'truncLen=(205,205)', 'truncQ=(2,2)', and 'maxEE=(1,1)', and error estimation was performed by pooling all sequences into one error model. Paired forward and reverse reads were assembled into one contig and trimmed to 230 bp. Chimera removal, ASV generation, and taxonomy assignment were performed as in Section 2.3. Four samples with fewer than 10 000 reads were removed. Because negative controls are specific to each study, contaminant reads were not identified or removed, but reads matching Kingdom Eukaryota and Order Chloroplast were removed. Data analysis was performed as described in Section 2.3. Briefly, the package 'zCompositions' (Palarea-Albaladejo & Martín-Fernández 2015) was used to impute zeroes before using the CLR transform to normalize the read counts. The Aitchison distance was then plotted against the geographic distance between samples to examine the distance-decay relationship between samples.

## 3. RESULTS

### 3.1. Fl and VI transects: site characteristics

In the Fl reef system, Dry Tortugas and Grecian were dominated by macroalgae (56–67% of cover), while Biscayne was dominated by turf algae (45%). Hard coral was more abundant at Dry Tortugas (21%) and Biscayne (20%), but only comprised 3% of Grecian (Fig. S1A in Supplement 1). In the VI system, live coral predominated at Brewer's Bay and Flat Cay (40–43%), but algae were slightly more prevalent at Black Point (40%) (Fig. S1B).

At the time of sampling, all reefs with the exception of Dry Tortugas had been experiencing outbreaks of stony coral tissue loss disease (SCTLD) to varying degrees of severity and duration (Precht et al. 2016, Brandt et al. 2021).

### 3.2. Fl and VI transects: sequence output

After quality control of the Fl and VI 16S rRNA gene amplicons from reef water transects, a total of 13 382 051 reads were retained, and the number of reads per sample ranged from 19 686 to 186 683 with a median of 78 976. A total of 20 488 ASVs were identified over 156 samples. Per-sample unique ASVs averaged 461. The abundance matrices of the ASV counts per sample were very sparse, comprising 97.7% zeros, indicating that a small number of taxa comprised the majority of the data set. Specifically, only 1228 ASVs made up the top 90% of observations across all samples.

### 3.3. Fl and VI transects: alpha and beta diversity metrics

Alpha diversity metrics of the reef water microbiomes, including observed ASV richness, Shannon index (a measure of evenness), and Simpson's index (a measure of dominance) measured at the transect (within-reef), individual reef, and reef system level in Fl and VI showed comparable values with some notable differences. At the transect level (within reefs), observed ASVs were most variable (highest and lowest values) at the Dry Tortugas reefs, with some outliers at both Fl and VI reefs. Simpson's index was most variable at the Dry Tortugas and Brewer's Bay reef (Fig. S2; pairwise *t*-test: $p < 0.0001$ for all significant comparisons involving Dry Tortugas and Brewer's Bay). At the individual reef level, signifi-

cantly lower Simpson's index values were detected at Dry Tortugas (mean: 0.93) and Brewer's Bay (mean: 0.94) compared to the other reefs (mean: 0.97). Both reefs also displayed significantly lower Shannon index values than 3 other reefs, except Biscayne (Fig. S2; pairwise *t*-test: $p < 0.002$ for significant differences, $p > 0.1$ for non-significant differences). At the reef system level (Fl compared to VI), each of the diversity metrics were significantly different (Fig. S2).

An NMDS ordination of Bray-Curtis dissimilarities between reef water microbiomes showed that the VI reefs clustered together along with the Dry Tortugas reef water microbial communities, and Fl reefs Grecian and Biscayne were more separated (Fig. 2A). Bray-Curtis dissimilarities for reef system (Fl and VI) and individual reefs differed significantly, but transects within a reef were not significantly different from each other (PERMANOVA) (Table 2).

Table 2. PERMANOVA results based on Bray-Curtis dissimilarities of microbial community abundances grouped at the spatial scales of within reefs, between reefs, and between reef regions. Pseudo-*F* values derived from 999 permutations. Values in **bold** are significant at $p < 0.05$

| Scope | df | SS | $R^2$ | Pseudo-*F* | p |
|---|---|---|---|---|---|
| Transect (within reefs) | 15 | 8.4 | 0.65 | 17.5 | 0.412 |
| Residual | 140 | 4.5 | 0.35 | | |
| Total | 155 | 12.8 | 1.00 | | |
| Site (between reefs) | 5 | 8.1 | 0.63 | 50.6 | **0.001** |
| Residual | 150 | 4.8 | 0.37 | | |
| Total | 155 | 12.8 | 1.00 | | |
| Reef system (between reef regions) | 1 | 3.3 | 0.26 | 53.0 | **0.001** |
| Residual | 154 | 9.5 | 0.74 | | |
| Total | 155 | 12.8 | 1.00 | | |

A comparison of microbial community beta dispersion, calculated as the distance from the centroid of each reef's reef microbial communities in principal coordinate space, showed significant differences at the level of reef system (permutest; $F = 30.12$, $p = 0.001$), individual reef ($F = 4.87$, $p = 0.001$), and within-reef transects (permutest; $F = 1.90$, $p = 0.026$). However, a post hoc test (Tukey's HSD) on the reef- and transect-based beta dispersions found that the difference was driven solely by the comparisons between the Dry Tortugas reef versus Grecian and Black Point (Fig. 2B). The Dry Tortugas reef had low variance among its samples while Grecian and Black Point had a larger variance among samples. When Dry Tortugas was removed from the analysis, there was no longer a significant effect of individual reef or transect on beta dispersion.

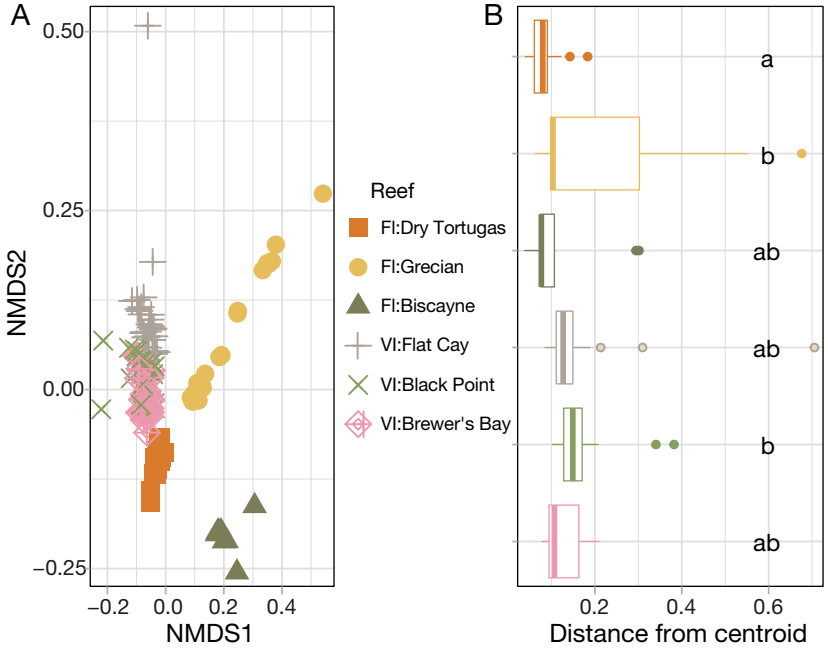


Fig. 2. Reefs have distinct microbial communities and variable group dispersions. (A) A non-metric multidimensional scaling ordination (NMDS) ordination using Bray-Curtis dissimilarities. 2D stress: 0.1006136. (B) Group dispersions calculated as distance from centroid using the Bray-Curtis dissimilarity metric. Vertical line in each boxplot: median; box hinges: first and third quartiles; whiskers: 1.5 × interquartile range; dots beyond whiskers: outliers. Colors correspond to the individual reef; different letters indicate significance differences (pairwise *t*-test; p-adjusted < 0.05)

While only reef-wide benthic composition was recorded for the Fl reefs, we recorded the underlying benthic composition for each sample in the VI reefs to enable comparison between substrate type and the overlying seawater microbiome. When all samples from VI were considered together, Bray-Curtis dissimilarities weakly correlated with benthic substrate type (PERMANOVA; $R^2 = 0.093$, Pseudo-*F* = 1.64, $p = 0.042$). However, a comparison of live coral compared to other categories (grouped together) did not show a correlation (PERMANOVA; $R^2 = 0.0091$, Pseudo-$F = 0.77$, $p = 0.62$), and when samples were nested by their respective reef, the correlation with benthic substrate was no longer

significant (PERMANOVA; Pseudo-$F$ = 1.64, p = 0.23). Group dispersions were not different between benthic substrate classes or between live coral and other substrates, and no taxa were found to be differentially abundant in either of these contrasts (permutest; $F$ = 2.32, p = 0.063).

### 3.4. Fl and VI transects: differentially abundant taxa

To investigate which taxa may be driving differences in community composition with respect to individual reefs, a differential abundance (DA) analysis was performed on the data set. DA analysis revealed 138 taxa that were significantly differentially abundant (p < 0.05, false discovery rate corrected using the Benjamini-Hochberg method) across the 6 reefs sampled. These significant taxa included common oligotrophic marine groups, such as the SAR11 and SAR86 clade, *Cyanobiaceae*, SAR116, and the Archaean Marine Group II. Opportunistic copiotrophs such as *Flavobacteriaceae*, *Rhodobacteraceae*, and *Vibrionaceae* were also well represented. All significant taxa were among the most abundant and most variable (displayed the highest variance in their relative abundances across samples) in the data set (Fig. S3). A list of the significant taxa along with their sequences is provided in Table S2.

### 3.5. Secondary analysis: sequence output and methods analysis

After assembly and quality control of the raw sequence reads from the 5 studies comprising the secondary analysis, a total of 8 761 462 reads were retained. The number of reads per sample ranged from 10441 to 159388 with a median of 35 967. A total of 15 005 ASVs were identified across the samples of all studies. Per-sample unique ASVs averaged 272. Although the Fl and VI transect samples had on average greater sequencing depth and ASV counts per sample than the studies in the secondary analysis, the relationship between sequencing depth and observed ASVs does not appear to have been saturated in either case (Fig. S4).

Because 2 studies in the analysis used the 806R primer while the other 3 used the 806RB primer, we evaluated the impact of this difference on the study. The group dispersions between the 2 primer sets are not significantly different (permutest; $F$ = 2.11, p = 0.14), indicating that primer choice did not significantly contribute to community variability. While DNA extraction methods were generally similar across studies, this was not similarly tested as a factor because only 2 studies shared the same method.

### 3.6. Distance-decay relationship

In order to investigate the impact of geographic distance on the microbial community, a geographic distance matrix was generated using the samples' physical location and compared to the Aitchison distances calculated between microbial communities (a measure of dissimilarity). The comparisons of samples within transects at an individual reef spanned <10 m, while individual reefs were 1–3 km apart for VI reefs and 34–279 km for Fl reefs. Mantel tests of these 2 matrices revealed that physical distance was significantly correlated with the Aitchison distance between samples within each (Fl and VI) reef system, with a stronger relationship in Fl (r = 0.45, p = 0.001 for Fl and r = 0.10, p = 0.004 for VI) (Table 3). There was also a relationship between geographic distances and microbial communities for both Fl and VI reef systems combined, with comparisons spanning 0–1978 km (r = 0.34, p = 0.001).

Incorporating the additional reef water microbiomes from the secondary analysis provided us with the opportunity to extend this study to 1000s of km. Distances between reefs within a reef system ranged from 1 to 2775 km while reef systems were separated by 421 to 16 874 km. The Aitchison distance between these secondary analysis samples showed a significant relationship to geographic distance (r = 0.28, p = 0.001) (Table 3).

Table 3. Distance-decay relationships at multiple spatial scales. Fl: Florida reef tract; VI: St. Thomas. Linear fit values significant at p < 0.01 are in **bold**

| Scope | Spatial scale | Correlation (Mantel r) | Mantel p-value | Slope of linear fit |
|---|---|---|---|---|
| Within transects | 0–9 m | NA[a] | NA[a] | −0.002 |
| (Reef system) USVI | 0–3 km | 0.101 | 0.004 | **−0.008** |
| (Reef system) Florida Keys | 0–279 km | 0.449 | 0.001 | **−3.208** |
| All samples in Fl/VI-based study | 0–1978 km | 0.336 | 0.001 | **−16.246** |
| All samples in secondary analysis | 0–16 874 km | 0.284 | 0.001 | **−113.128** |

[a]Mantel test not performed because distances exist for samples within transects but not between, therefore matrix had >50% missing values

In the Fl and VI study, distance-decay plots showed no relationship between geographic distance and microbial community similarity within reefs (at the transect level) at the scale of meters ($p > 0.05$) (Fig. 3A). However, at the scale of kilometers, there was a negative relationship between geographic distance and microbial community similarity for the Fl and VI study (Fig. 3B), as well as for the more expansive secondary analysis (Fig. 3C), with microbial communities becoming less similar with increasing distance ($R^2 = 0.11$, $p < 0.001$ and $R^2 = 0.08$, $p < 0.001$, respectively). The slope of the negative distance-decay relationship increased in magnitude as the geographic extent of the samples increased (Table 3, Fig. 3C).

### 3.7. Drivers of distance-decay relationship

In the secondary analysis, the effects of collection depth, temperature, and reef type were examined as potential drivers of community similarity. A PERMANOVA assessing the marginal impacts of these abiotic factors as well the effect of study found that study accounted for the most variation ($R^2 = 0.14$, Pseudo-$F = 9.32$, $p = 0.001$), distantly followed by reef type ($R^2 = 0.03$, Pseudo-$F = 2.06$, $p = 0.001$), collection depth ($R^2 = 0.017$, Pseudo-$F = 4.64$, $p = 0.001$), and finally temperature ($R^2 = 0.017$, Pseudo-$F = 4.49$, $p = 0.001$). Despite explaining the least amount of variation, difference in temperature was significantly correlated with Aitchison distance (Mantel: $r = 0.18$, $p = 0.001$), meaning communities that were more different (distant) in temperature were also more dissimilar. A similar correlation for depth was not found (Mantel: $r = -0.00035$, $p = 0.49$).

## 4. DISCUSSION

In this study, we used a nested distance sampling design to examine how reef seawater microbiomes vary at multiple spatial scales, including within reefs, between individual reefs in a reef system, and across northern Caribbean (Fl and VI) forereef systems. Overall, we found that individual reef and reef system-related features had the largest influence on microbial community diversity and composition. No differences in microbial community diversity or composition were detected within different locations on individual reefs, and there was a weak correlation with the benthic substrate underlying the sample. Despite the large number of observed microbial taxa in the transect-based study, just over 100 of the most abundant were identified as differentially abundant between reefs, suggesting that these abundant taxa may be useful indicators of reef change. We also used data from 5 previous studies in a secondary analysis to understand the biogeography of more distant reef seawater microbiomes; these data revealed that microbial communities are more distinct with increasing geographic distance.

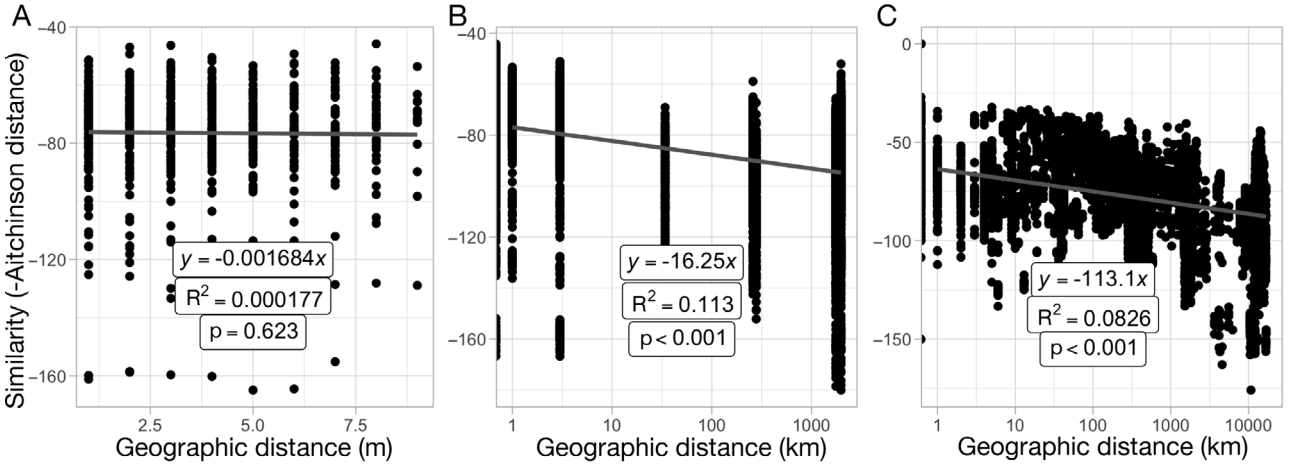


Fig. 3. Reef seawater communities exhibit a distance–decay relationship at the 100 km but not <10 m scale. (A) Pairwise geographic distances between samples within each transect in the Florida/Virgin Islands (Fl/VI)-based study versus the corresponding Aitchison distances. Between-transect distances are not known and therefore not included. (B) Distance-decay plot of all pairwise distances (i.e. not just within a transect) between samples collected in the Fl/VI-based study. (C) Distance-decay plot using the samples from the 5 studies included in the secondary analysis

### 4.1. Microbial communities differentiate by individual reef (>1 km) and reef system (>100 km), but not within each reef (<10 m)

Counter to expectations, the benthic substrate did not have a strong influence on the composition of the seawater microbiome. Although there is evidence that substrate type influences the surrounding seawater microbial community (Schöttner et al. 2012, Tout et al. 2014), we did not find a correlation in these data, nor did we find differentially abundant microbial taxa between substrate types. It is likely that differential hydrodynamic conditions, which have not yet been measured in the context of coral reef benthic–pelagic microbial interactions, may play a role in these differential results. As such, additional research concerning benthic–pelagic exchange on coral reefs is needed to understand the impact of substrate seawater microorganisms close to the reef surface.

All 6 reefs were distinguishable in terms of the composition of their reef water microbial communities, despite the 3 VI reefs being separated by only 1–3 km. These results align with our expectation that microbes in the water column above reefs would display reef-specific signatures because marine microbial communities are reflective of their physical and chemical environment (Azam & Malfatti 2007, Kelly et al. 2018). While benthic substrate had a weak influence on the seawater microbes, there were other influences that varied between the reefs, including season, time of sampling, depth, and reef type, but these generally were consistent within VI or Fl and therefore difficult to statistically examine. Indeed, in our secondary analysis, which included a larger number of reef sites and more geographic locations, temperature, depth, and reef type were small but significant contributors to the community variation. It must be noted, however, that all variables examined were highly confounded by the specific sampling scheme of each study. For example, Becker et al. (2020) only sampled reef seawater from 4 forereefs in the VI at 0.3 m depth, and Weber et al. (2020) contained a disproportionate number of samples at cooler temperatures (67 out of 82 samples below average temperature of the secondary analysis), all from Cuban reefs. In addition to our study, other studies of the Indian Ocean and Northwestern Hawaiian islands have shown strong microbial biogeographic signatures (Jeffries et al. 2015, Salerno et al. 2016). Hydrodynamics likely plays a major role in explaining some of the biogeographical portioning between reef water microbial communities because it impacts distribution and transport of nutrients and facilitates dispersal of pelagic microorganisms. Previous studies have suggested links between water masses and microbial community composition (Varela et al. 2008, Galand et al. 2010, Jeffries et al. 2015). Comparison of reef water microbial communities within and between hydrographic regimes and current systems could help us better understand this influence.

Surprisingly, we did not identify consistent differences in the microbial communities within each reef (at the transect level). Samples collected within meters of each other were indistinguishable in VI and Fl, but trends in beta dispersion did suggest some within-reef variability, indicating differences in beta diversity among one or more groups (Anderson et al. 2006). Dispersion was greatest at Grecian and lowest at Dry Tortugas, both Fl reefs. Grecian reef is located in the Upper Keys, which on average has elevated nutrients, organic carbon, and turbidity compared to the Lower Keys (Lirman & Fong 2007). In contrast, the Dry Tortugas reef is located within a marine protected zone (US National Park), is more distant from the shore, was the deepest reef sampled (60 feet; 18 m), and was the only reef not experiencing active outbreaks of SCTLD at the time of sampling. These factors could have contributed to the relative homogeneity of the samples collected at Dry Tortugas. Reef depth and coastal influence may be among the regional geographic conditions that influence the variability of microbial communities in reef-associated seawater (Frade et al. 2020, Weber et al. 2020).

### 4.2. Community similarity decays with distance beginning at the km scale

The distance-decay of community similarity — a widely studied relationship in ecology (Soininen et al. 2007) — quantifies the decrease in community similarity with increased geographic distance. Typically, communities that are closer geographically are also more similar to each other compositionally (Soininen et al. 2007). One mechanism that drives this relationship is spatial structuring, where locations closer together have more similar environments, thus leading to selection of more similar communities. In the absence of selection (e.g. in a homogeneous environment), neutral drift interacts with dispersal limitation to differentiate communities over space (Soininen et al. 2007, Hanson et al. 2012). These mechanisms represent 2 hypotheses for what drives species distributions: environmental selection and historical contin-

gency (Martiny et al. 2006). A number of studies have examined the distance-decay relationship in both soil and marine environments and found that microorganisms tend to display a weaker (i.e. less negative) relationship compared to macroorganisms on the same scale, a phenomenon attributed to the small size and large populations of microorganisms leading to greater dispersal (Green & Bohannan 2006, Martiny et al. 2006, Meyer et al. 2018).

The sampling pattern in this study allowed us to assess this biogeographic pattern in the context of coral-reef-associated seawater. Within transects in a reef (<10 m scale), no distance-decay relationship was found, likely due to high mixing rates on the reef. However, there was a significant correlation between community similarity and geographic distance beginning at the reef level (1 km scale), and the steepness of the relationship increased with an increase in geographic extent (10 000 km scale) (Table 3). Differences in the steepness and strength of correlation of the distance-decay relationship may reflect different mechanisms driving the decay at multiple spatial scales (Martiny et al. 2011). The larger correlation and slope observed in the Fl reefs compared to VI reefs may reflect the orientation of Fl reefs in a north–south line along the Florida current, with the most distant reef upstream of the 2 closer reefs, while the VI reefs were closer together and not oriented in relation to the surrounding Caribbean current. The autocorrelation of distance and environmental similarity (Lirman & Fong 2007) along the Fl likely drives the stronger correlation compared to the VI reefs. The steep slope and weaker correlation found in the secondary analysis likely reflects historical factors such as dispersal limitation and drift as distant reefs recruit from different metapopulations (Hellweger et al. 2014, Clark et al. 2021).

### 4.3. Abundant taxa are most variable and more likely to differentiate individual reefs

Although we recovered a total of over 20 000 microbial ASVs from the Fl and VI transect sampling, the vast majority were rare, and samples were dominated by just over 1000 highly abundant taxa. This is a common occurrence in microbiome sequencing, especially with the advent of high throughput deep sequencing, and there is debate over the importance of these rare taxa (McMurdie & Holmes 2014, Cao et al. 2021). Abundant taxa tend to be the most prevalent and, in this study, also displayed the highest variance in their relative abundance values between samples. The ASVs that were identified as differentially abundant between reefs were also among the most abundant taxa (Fig. S3). Glasl et al. (2019) found that the relative abundances of indicator taxa in coral reef seawater that best predict environmental conditions range from 0.5 to 20 %, and those taxa were also prevalent throughout that study's sampling period. Overlap in taxonomic assignment between these indicator taxa and differentially abundant taxa found within this study include *Synechococcus*, *Prochlorococcus*, *Rhodobacteraceae*, unclassified *Alphaproteobacteria*, and others. Because extremely rare taxa can be difficult to reliably detect, more frequent, shallower sequencing may be more important for capturing the salient variability of a reef.

### 4.4. Caveats

Although the nested design surveyed seawater microbial communities across multiple spatial scales, the temporal scale of seawater variability was not considered in this study. The Fl samples were taken in June, while the VI samples were taken in February. Each reef was only sampled once and not throughout the day. Seasonal as well as diurnal/tidal cycles in reef seawater microbial communities are well documented (Becker et al. 2020, Frade et al. 2020, Glasl et al. 2020, Weber & Apprill 2020). The differences between the Fl and VI microbial communities may be due in part to the different seasons in which the samples were taken. Temporal differences in sampling can make direct comparisons between distant reefs challenging, even within the same study (Weber & Apprill 2020). While microbial communities are sensitive to environmental conditions, coral reef seawater remains distinct from other seawater habitats (Becker et al. 2020), and variation of the microbial community is better explained by reef-level environmental parameters rather than seasonal differences (Glasl et al. 2019). Within a reef, repeated sampling throughout the calendar year may be needed to establish the baseline variability.

In contrast to many other seawater microbiome studies, including those in the secondary analysis, this study filtered a small volume of seawater (60 ml) for each sample rather than the more typical 1–2 l. Weber et al. (2019) directly examined the effect of sampling volume and found that while species richness modestly increased with larger volumes due to sampling rare taxa, beta diversity and overall community composition were not influenced by sampling volume.

### 4.5. Conclusions

We suggest that due to reef and reef system-level influences, the development of reef water microbiome monitoring criteria may need to be regionally tailored. We found that the community composition of reef seawater microbiomes is distinguishable even when reefs are a few km apart and that there can be large differences in beta dispersion within a reef. Detecting a shift in the community composition as a whole will necessitate an understanding of each reef region's variability. Individual reefs within a reef system may also be experiencing different regional stressors, such as varying degrees of anthropogenic influence. Such differences may be reflected in both the baseline microbial community composition and variability as reef conditions change. Additionally, microbial taxa common between reef regions are vastly outnumbered by taxa that are unique, making it difficult to develop a generalized database of indicator microbial taxa for reef environmental conditions. Overall, we found that the seawater microbial communities of reefs closer together are more similar, and the local oceanographic conditions which differentiate these communities are important to investigate.

*Acknowledgements*. We thank Carolyn Miller, Joseph Townsend, Kathryn Cobleigh, Sonora Meiling, Kelsey Beavers, Erinn Muller, Abigail Clark, and the captain and crew of the M/V 'Alucia' for field assistance and Illumina Inc. for technical support. Samples were collected under Florida Keys National Marine Sanctuary permit no. FKNMS-2019-046, US National Park Service permit DRTO-2019-SCI-0010 and USVI territory permit DFW19057U. This project was supported by an OceanX-sponsored cruise on the M/V 'Alucia', with additional funding provided to A.A. by the Rockefeller Philanthropy Advisors, Dalio Foundation, and other generous donors of the Oceans 5 project and NSF OCE-1938147 and NSF OCE-1928761 awards.

LITERATURE CITED

Amend AS, Oliver TA, Amaral-Zettler LA, Boetius A and others (2013) Macroecological patterns of marine bacteria on a global scale. J Biogeogr 40:800–811

Anderson MJ, Ellingsen KE, McArdle BH (2006) Multivariate dispersion as a measure of beta diversity. Ecol Lett 9: 683–693

Apprill A, McNally S, Parsons R, Weber L (2015) Minor revision to V4 region SSU rRNA 806R gene primer greatly increases detection of SAR11 bacterioplankton. Aquat Microb Ecol 75:129–137

Apprill A, Holm H, Santoro AE, Becker C and others (2021) Microbial ecology of coral-dominated reefs in the Federated States of Micronesia. Aquat Microb Ecol 86: 115–136

Azam F, Malfatti F (2007) Microbial structuring of marine ecosystems. Nat Rev Microbiol 5:782–791

Becker CC, Weber L, Suca JJ, Llopiz JK, Mooney TA, Apprill A (2020) Microbial and nutrient dynamics in mangrove, reef, and seagrass waters over tidal and diurnal time scales. Aquat Microb Ecol 85:101–119

Bourne DG, Morrow KM, Webster NS (2016) Insights into the coral microbiome: underpinning the health and resilience of reef ecosystems. Annu Rev Microbiol 70: 317–340

Brandt ME, Ennis RS, Meiling SS, Townsend J and others (2021) The emergence and initial impact of stony coral tissue loss disease (SCTLD) in the United States Virgin Islands. Front Mar Sci 8:1105

Callahan BJ, McMurdie PJ, Rosen MJ, Han AW, Johnson AJA, Holmes SP (2016) DADA2: high-resolution sample inference from illumina amplicon data. Nat Methods 13: 581–583

Cao Q, Sun X, Rajesh K, Chalasani N and others (2021) Effects of rare microbiome taxa filtering on statistical analysis. Front Microbiol 11:607325

Caporaso JG, Lauber CL, Walters WA, Berg-Lyons D and others (2011) Global patterns of 16S rRNA diversity at a depth of millions of sequences per sample. Proc Natl Acad Sci USA 108:4516–4522

Clark DR, Underwood GJC, McGenity TJ, Dumbrell AJ (2021) What drives study-dependent differences in distance–decay relationships of microbial communities? Glob Ecol Biogeogr 30:811–825

Corcoll N, Österlund T, Sinclair L, Eiler A, Kristiansson E, Backhaus T, Eriksson KM (2017) Comparison of four DNA extraction methods for comprehensive assessment of 16S rRNA bacterial diversity in marine biofilms using high-throughput sequencing. FEMS Microbiol Lett 364: fnx139

Davis NM, Proctor DM, Holmes SP, Relman DA, Callahan BJ (2018) Simple statistical identification and removal of contaminant sequences in marker-gene and metagenomics data. Microbiome 6:226

Fox MD, Carter AL, Edwards CB, Takeshita Y and others (2019) Limited coral mortality following acute thermal stress and widespread bleaching on Palmyra Atoll, Central Pacific. Coral Reefs 38:701–712

Frade PR, Glasl B, Matthews SA, Mellin C and others (2020) Spatial patterns of microbial communities across surface waters of the Great Barrier Reef. Commun Biol 3:442

Galand PE, Potvin M, Casamayor EO, Lovejoy C (2010) Hydrography shapes bacterial biogeography of the deep Arctic Ocean. ISME J 4:564–576

Glasl B, Webster NS, Bourne DG (2017) Microbial indicators as a diagnostic tool for assessing water quality and climate stress in coral reef ecosystems. Mar Biol 164:91

Glasl B, Bourne DG, Frade PR, Thomas T, Schaffelke B, Webster NS (2019) Microbial indicators of environmental perturbations in coral reef ecosystems. Microbiome 7:94

Glasl B, Robbins S, Frade PR, Marangon E, Laffy PW, Bourne DG, Webster NS (2020) Comparative genome-centric analysis reveals seasonal variation in the function of coral reef microbiomes. ISME J 14:1435–1450

Gloor GB, Macklaim JM, Pawlowsky-Glahn V, Egozcue JJ (2017) Microbiome datasets are compositional: and this is not optional. Front Microbiol 8:2224

Green J, Bohannan BJM (2006) Spatial scaling of microbial biodiversity. Trends Ecol Evol 21:501–507

Haas AF, Nelson CE, Kelly LW, Carlson CA and others (2011) Effects of coral reef benthic primary producers on dissolved organic carbon and microbial activity. PLOS ONE 6:e27973

Haas AF, Fairoz MFM, Kelly LW, Nelson CE and others (2016) Global microbialization of coral reefs. Nat Microbiol 1:16042

Hanson CA, Fuhrman JA, Horner-Devine MC, Martiny JBH (2012) Beyond biogeographic patterns: processes shaping the microbial landscape. Nat Rev Microbiol 10: 497–506

Hellweger FL, van Sebille E, Fredrick ND (2014) Biogeographic patterns in ocean microbes emerge in a neutral agent-based model. Science 345:1346–1349

Holm S (1979) A simple sequentially rejective multiple test procedure. Scand J Stat 6:65–70

Hughes TP, Barnes ML, Bellwood DR, Cinner JE and others (2017) Coral reefs in the Anthropocene. Nature 546: 82–90

Jeffries TC, Ostrowski M, Williams RB, Xie C and others (2015) Spatially extensive microbial biogeography of the Indian Ocean provides insights into the unique community structure of a pristine coral atoll. Sci Rep 5:15383

Kelly LW, Williams GJ, Barott KL, Carlson CA and others (2014) Local genomic adaptation of coral reef-associated microbiomes to gradients of natural variability and anthropogenic stressors. Proc Natl Acad Sci USA 111: 10227–10232

Kelly LW, Haas AF, Nelson CE (2018) Ecosystem microbiology of coral reefs: linking genomic, metabolomic, and biogeochemical dynamics from animal symbioses to reefscape processes. mSystems 3:e00162-17

Knowlton N, Rohwer F (2003) Multispecies microbial mutualisms on coral reefs: the host as a habitat. Am Nat 162: S51–S62

Kozich JJ, Westcott SL, Baxter NT, Highlander SK, Schloss PD (2013) Development of a dual-index sequencing strategy and curation pipeline for analyzing amplicon sequence data on the MiSeq Illumina sequencing platform. Appl Environ Microbiol 79:5112–5120

Lirman D, Fong P (2007) Is proximity to land-based sources of coral stressors an appropriate measure of risk to coral reefs? An example from the Florida reef tract. Mar Pollut Bull 54:779–791

Liu M, Xue Y, Yang J (2019) Rare plankton subcommunities are far more affected by DNA extraction kits than abundant plankton. Front Microbiol 10:454

Marcelino VR, van Oppen MJ, Verbruggen H (2018) Highly structured prokaryote communities exist within the skeleton of coral colonies. ISME J 12:300–303

Martin M (2011) Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet J 17:10–12

Martin BD, Witten D, Willis AD (2021) corncob: count regression for correlated observations with the beta-binomial. R package version 0.2.0. https://CRAN.R-project.org/package=corncob

Martiny JBH, Bohannan BJM, Brown JH, Colwell RK and others (2006) Microbial biogeography: putting microorganisms on the map. Nat Rev Microbiol 4:102–112

Martiny JBH, Eisen JA, Penn K, Allison SD, Horner-Devine MC (2011) Drivers of bacterial β-diversity depend on spatial scale. Proc Natl Acad Sci USA 108:7850–7854

McDevitt-Irwin JM, Baum JK, Garren M, Vega Thurber RL (2017) Responses of coral-associated bacterial communities to local and global stressors. Front Mar Sci 4:262

McMurdie PJ, Holmes S (2013) phyloseq: an R package for reproducible interactive analysis and graphics of microbiome census data. PLOS ONE 8:e61217

McMurdie PJ, Holmes S (2014) Waste not, want not: why rarefying microbiome data is inadmissible. PLOS Comput Biol 10:e1003531

Meyer KM, Memiaghe H, Korte L, Kenfack D, Alonso A, Bohannan BJM (2018) Why do microbes exhibit weak biogeographic patterns? ISME J 12:1404–1413

Neave MJ, Rachmawati R, Xun L, Michell CT, Bourne DG, Apprill A, Voolstra CR (2017) Differential specificity between closely related corals and abundant endozoicomonas endosymbionts across global scales. ISME J 11:186–200

Nelson CE, Alldredge AL, McCliment EA, Amaral-Zettler LA, Carlson CA (2011) Depleted dissolved organic carbon and distinct bacterial communities in the water column of a rapid-flushing coral reef ecosystem. ISME J 5: 1374–1387

Nelson CE, Goldberg SJ, Wegley Kelly L, Haas AF, Smith JE, Rohwer F, Carlson CA (2013) Coral and macroalgal exudates vary in neutral sugar composition and differentially enrich reef bacterioplankton lineages. ISME J 7: 962–979

Oksanen J, Blanchet FG, Friendly M, Kindt R and others (2020) vegan: community ecology package. R package version 2.5-7. https://CRAN.R-project.org/package=vegan

Palarea-Albaladejo J, Martín-Fernández JA (2015) zCompositions — R package for multivariate imputation of left-censored data under a compositional approach. Chemom Intell Lab Syst 143:85–96

Parada AE, Needham DM, Fuhrman JA (2016) Every base matters: assessing small subunit rRNA primers for marine microbiomes with mock communities, time series and global field samples. Environ Microbiol 18: 1403–1414

Peixoto RS, Rosado PM, Leite DCA, Rosado AS, Bourne DG (2017) Beneficial microorganisms for corals (BMC): proposed mechanisms for coral health and resilience. Front Microbiol 8:341

Precht WF, Gintert BE, Robbart ML, Fura R, van Woesik R (2016) Unprecedented disease-related coral mortality in southeastern Florida. Sci Rep 6:31374

Pruesse E, Quast C, Knittel K, Fuchs BM, Ludwig W, Peplies J, Glöckner FO (2007) SILVA: a comprehensive online resource for quality checked and aligned ribosomal RNA sequence data compatible with ARB. Nucleic Acids Res 35:7188–7196

Rosenberg E, Koren O, Reshef L, Efrony R, Zilber-Rosenberg I (2007) The role of microorganisms in coral health, disease and evolution. Nat Rev Microbiol 5:355–362

RStudio Team (2019) RStudio: integrated development environment for R. RStudio, Boston, MA

Salerno JL, Bowen BW, Rappé MS (2016) Biogeography of planktonic and coral-associated microorganisms across the Hawaiian Archipelago. FEMS Microbiol Ecol 92: fiw109

Santoro AE, Casciotti KL, Francis CA (2010) Activity, abundance and diversity of nitrifying archaea and bacteria in the Central California Current. Environ Microbiol 12: 1989–2006

Schöttner S, Wild C, Hoffmann F, Boetius A, Ramette A (2012) Spatial scales of bacterial diversity in cold-water coral reef ecosystems. PLOS ONE 7:e32093

Soininen J, McDonald R, Hillebrand H (2007) The distance

decay of similarity in ecological communities. Ecography 30:3–12

Sweet MJ, Croquer A, Bythell JC (2010) Temporal and spatial patterns in waterborne bacterial communities of an island reef system. Aquat Microb Ecol 61:1–11

Tout J, Jeffries TC, Webster NS, Stocker R, Ralph PJ, Seymour JR (2014) Variability in microbial community composition and function between different niches within a coral reef. Microb Ecol 67:540–552

Urakawa H, Martens-Habbena W, Stahl DA (2010) High abundance of ammonia-oxidizing *Archaea* in coastal waters, determined using a modified DNA extraction method. Appl Environ Microbiol 76:2129–2135

van Duyl FC, Gast GJ (2001) Linkage of small-scale spatial variations in DOC, inorganic nutrients and bacterioplankton growth with different coral reef water types. Aquat Microb Ecol 24:17–26

Varela MM, Van Aken HM, Sintes E, Herndl GJ (2008) Latitudinal trends of *Crenarchaeota* and *Bacteria* in the meso- and bathypelagic water masses of the eastern North Atlantic. Environ Microbiol 10:110–124

Wallace JR (2012) Imap: interactive mapping. R package version 1.32. https://CRAN.R-project.org/package=Imap

Wang Q, Garrity GM, Tiedje JM, Cole JR (2007) Naïve Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. Appl Environ Microbiol 73:5261–5267

Weber L, Apprill A (2020) Diel, daily, and spatial variation of coral reef seawater microbial communities. PLOS ONE 15:e0229442

Weber L, Gonzalez-Díaz P, Armenteros M, Apprill A (2019) The coral ecosphere: a unique coral reef habitat that fosters coral–microbial interactions. Limnol Oceanogr 64: 2373–2388

Weber L, Gonzalez-Díaz P, Armenteros M, Ferrer VM and others (2020) Microbial signatures of protected and impacted northern Caribbean reefs: changes from Cuba to the Florida Keys. Environ Microbiol 22:499–519

Wickham H (2016) ggplot2: elegant graphics for data analysis. Springer-Verlag, New York, NY

*Editorial responsibility: Eva Lindström, Uppsala, Sweden*
*Reviewed by: 2 anonymous referees*

*Submitted: July 27, 2021*
*Accepted: January 31, 2022*
*Proofs received from author(s): March 18, 2022*