Fine-scale environmental, genetic, and temporal factors can drive the coral, sediment, and water column metagenome on reefs

Carly B Scott[1]*, Emily N Nixon[2], Alexa K Huzar[3], Mikhail V Matz[3]

[1]Department of Biology, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina, USA

[2]Marine Biology Research Division, Scripps Institution of Oceanography, University of California San Diego, La Jolla, California, USA

[3]Department of Integrative Biology, University of Texas at Austin, Austin, Texas, USA

*Corresponding Author

CB Scott, cabscott@unc.edu; EN Nixon, enixon@ucsd.edu; A K Huzar, alexahuzar@utexas.edu; MV Matz, matz@utexas.edu

Funding Statement

This research was supported by NSF Grant OCE-2318775 to MVM, Human Frontiers Science Program Grant 201903079001 to MVM, a grant from the Integrative Biology Department at the University of Texas at Austin to CBS, and a grant from the International Women's Fishing Association to CBS. The funders played no role in the design or analysis of this study.

Abstract

Understanding the stability of microbial community assembly on coral reefs is crucial for determining their response to changing environments. Here, we evaluate how the marine sediment, water column, and mountainous star coral (*Orbicella faveolata*) metagenomic communities shift over temporal, spatial, and genetic gradients, centered on the annual coral spawning event on St. Croix, USVI. Using low-coverage shotgun sequencing and a reference-free approach, we characterized metagenomic communities and additionally report 15 metagenome-assembled genomes from environmental samples. We show that diel time can be as important as kilometer-scale distance in structuring the coral-associated and marine sediment community structure, albeit with other highly local factors (genotype, within-reef location) driving most variation. Through multivariate machine learning methods, we find several taxa and metabolic pathways enriched over time, where broad metagenomic community shifts may be linked with concurrent temperature increase or the annual coral spawning event. As understanding the coral microbiome in the context of its environment becomes increasingly important, our research underscores the importance of small-scale variation, both spatially and temporally, in structuring coral reef metagenomic communities.

Keywords

coral reefs; metagenomics; environmental metagenome; coral spawning; *Orbicella faveolata*

Introduction

Coral resilience to environmental change can depend both on the coral host and the algal and microbial symbionts it associates with (together, the "holobiont"; Knowlton & Rohwer, 2003; Rohwer et al., 2001). Local environment plays a key role in determining these relationships, with corals exhibiting flexible, but site-specific associations (Chan et al., 2024; Deignan & McDougald, 2022; Dunphy et al., 2019; Wegley Kelly et al., 2014). Holobiont membership may be selected by the host to provide environment-specific benefits and potentially be altered to provide a rapid mechanism to acclimatize to future conditions (Reshef et al., 2006). Alternatively, holobiont community assembly may instead reflect neutral ecological process, such as the environmental abundance of microorganisms at ecologically relevant times of the year (e.g., coral spawning), or a coevolutionary relationship between coral individuals and their holobionts (Ziegler et al., 2019).

Partitioning the drivers of holobiont composition requires understanding the coral host in the context of its reef environment. Holobiont associations are relatively well characterized for coral colonies, but parallel temporal data on both the surrounding sediment and water column metagenomic community (the full diversity of DNA present in the sample) is rarely measured (but see Glasl et al., 2019). The drivers of seawater microorganisms are better understood than the reef sediment. The reef water column microbiome is environmentally sensitive and has already been proposed as an indicator of reef health (Glasl, Bourne, et al., 2019; Ma et al., 2022; Terzin et al., 2024; Weber & Apprill, 2020). In comparison, little work has been done to determine the drivers of the reef sediment metagenomic community, despite its proximity to coral colonies. It is clear the reef sediment community differs in response to large environmental perturbations and across broad scale geography, but it has not been well quantified at smaller scales (Dong et al., 2023a; Krausfeldt et al., 2023).

In general, the coral reef microbial community has been described well at extremely detailed (e.g., hours-to-days, over a single colony) and coarse (e.g., seasons-to-years, between reefs and coral species) resolutions. However, microbiome stability over more intermediate scales (e.g., days-to-weeks, within reefs and between genotypes at this timescale) is relatively less defined, especially for non-Symbiodiniaceae members. For example, diurnal shifts can affect the coral and seawater microbiome (Caughman et al., 2021; Weber & Apprill, 2020), but how these changes connect across days is not well characterized. Local environment plays a large role in structuring the reef holobiont (Kriefall et al., 2022), but spatial variability within a single reef is not well measured, especially for the marine sediment. Finally, at the genetic level, it is clear that different

coral species and cryptic lineages can host unique holobiomes (Grupstra et al., 2024; Scott, Schott, et al., 2024; Ziegler et al., 2019). Further, individual coral genotypes can host distinct microbial communities (Glasl, Smith, et al., 2019), but it is less clear what the relative role of genotype is in structuring holobionts in the natural environment.

Most studies of the coral holobiont to date have used amplicon sequencing to quantify dinoflagellate and prokaryotic community composition. However, these methods fail to capture broader eukaryotic and viral diversity and do not offer direct estimates of microbiome functional potential. By leveraging low-coverage whole genome shotgun sequencing (lcWGS), both species and gene composition of the metagenomic communities can be assessed in tandem, giving insights into whether community turnover is functionally redundant (Stothart et al., 2023). Already, implementations of genome-scale methods have offered functional insights to the reef microbiome across seasons (Glasl et al., 2020), diel cycles (Wegley Kelly et al., 2019), and bleaching events (Littman et al., 2011).

Here, we integrate lcWGS data from the sediment, water column, and coral hosts both within and between reefs across a four-week timeframe centered on the annual coral spawning event. We use this data to explore the drivers of taxonomic community structure and functional potential in these samples, showing that diel time can be just as important as the broad-scale environment in structuring host and environmental metagenomic datasets.

Materials and Methods

Sample Collections

Samples were collected on the north shore of St. Croix, USVI between August 9, 2022, and August 31, 2022. We collected *Orbicella faveolata*, marine sediment, and water column samples at two reefs (Cane Bay, CB, and Deep End, DE) across 4-5 time points (Figure 1). All samples were collected between 10AM and 2PM, except for sediment samples taken on the night of spawning at DE. Notably, this sampling coincided with an increase in mean sea surface temperature (meanSST; Figure 2). For the coral, we repeatedly sampled three tagged genotypes at each site and verified the identity of these corals through DNA sequencing. Inadvertently, we sampled an additional three genotypes across all time points and sites. For each *O. faveolata* tissue sample, three polyps were scratched (two on opposing sides, one on top of colony) using a 16G needle attached to a 20 mL luer-lock syringe. The three polyps collected from each colony were combined during sampling. Therefore, downstream processing and analysis values reflect the “colony-average” community resulting from the combination of these three polyps.

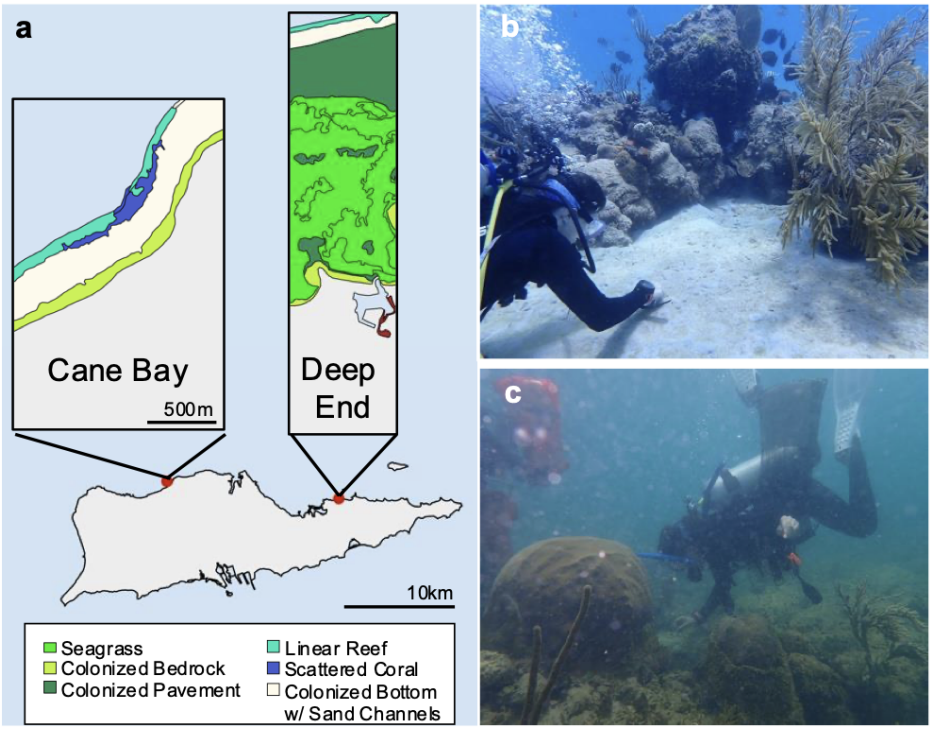


Figure 1. Environmentally distinct sampling sites from St. Croix, USVI. (a) Location of Cane Bay (CB) and Deep End (DE) reefs as well as substrate composition. (b) Photo of sampling at CB. CB is a high-visibility (10m - 100+m) deep site, with patch reef leading up to a continuous wall. (c) Photo of sampling at DE, a shallow (<5 m in photo), turbid reef near a boat ramp and resort. Photo Credits: Daisy Flores and Kristina Black.

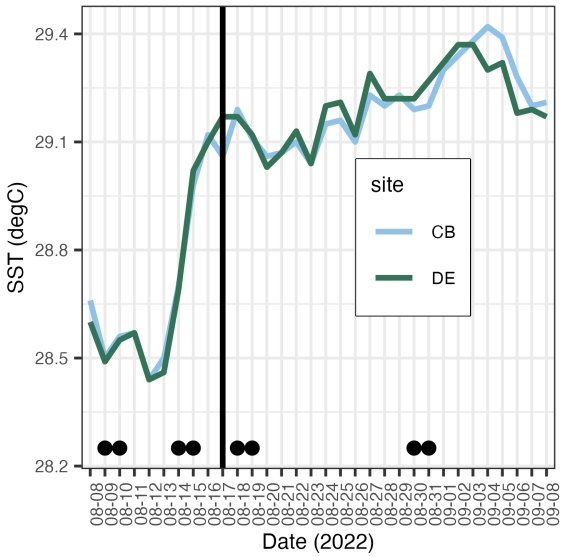


Figure 2. Daily mean sea surface temperature during the sampling window at CB and DE. Mean SST (NOAA Coral Reef Watch, 2023). Black dots along the x-axis give the sampling timepoints included in this study and the vertical black line indicates the night that spawning was observed at DE. CB was always sampled one day prior to DE.

121 Water was collected at depth <1m above sampled *O. faveolata* patches in sterile 500mL
122 containers. Water samples were collected in duplicate, where for a small subset of site
123 and time point combinations only one water sample was successfully obtained.

124 Marine sediment was sampled at three "subsites", moving cross-shelf, at each reef. At
125 each of the subsites, sediment samples were collected in triplicate within 5m of each
126 other using 2mL cryovial tubes. We collected parallel *O. faveolata* tissue, water column,
127 and sediment samples from the first subsite (Subsite 1), which was tagged to ensure we
128 collected from the exact same within-reef location at each timepoint. We did not tag the
129 other two subsites but sampled at the same depths at each timepoint moving cross-shelf.

130 Samples were processed within two hours of collection at The Nature Conservancy's Coral
131 Innovation Hub, in St. Croix, USVI. Coral tissue was first rinsed in 100% EtOH, then fixed in
132 100% EtOH and stored at -20°C until DNA extraction. Sediment samples were centrifuged
133 in a tabletop centrifuge for two minutes. Seawater was then removed from top of the vial
134 and replaced with 100% EtOH and samples were stored at -20°C. For the water samples,
135 300 mL of the collection was passed through a 0.22 µm filter (Millipore Sterivex), which
136 was then stored at -20°C.

137 Mass coral spawning was observed for *O. faveolata* at DE on the night of August 16, 2022.
138 Due to the distance between CB and DE and personnel limitations, we could not verify
139 spawning *in situ* at CB. However, it is highly likely that *O. faveolata* spawning occurred at
140 CB on August 16-17, as spawning was reported or predicted for the species during this
141 window at sites as distant as the Florida Keys and the Bahamas (Willams et al., 2022;
142 Hurtado-López et al., 2022). Thus, all measurements of temporal proximity to spawning
143 assume *O. faveolata* spawning occurred at both sites on this date.

144 *DNA Extraction & Library Prep*

145 DNA was extracted from coral tissue samples using CTAB lysis buffer and an organic
146 extraction (see Supplemental File 1 for full protocol). Coral extracts were then cleaned
147 with the Zymo Clean & Concentrator kit [D4011]. A similar method was used to extract
148 DNA from the Sterivex filters. The protocol was modified such that 2000 µL of CTAB buffer
149 was put into each Sterivex filter using a sterile 5 µL syringe with a luer lock tip. Prior to
150 incubating the Sterivex filter in CTAB buffer, excess water was removed by pushing air
151 through the filter two times with a 1000 µL syringe with a luer lock tip. The open ends of the
152 Sterivex filter were sealed with caps and then wrapped in parafilm to prevent leakage.
153 Filters containing CTAB buffer which were incubated at 42°C in CTAB buffer overnight in a
154 rocking incubator. The following day, we removed the CTAB buffer from the filters using
155 sterile 5 µL syringes and carried it into an organic extraction as detailed in Supplemental

File 1. Sediment samples were processed using the Qiagen PowerSoil Pro Kit. We used 0.125g of sediment as the input for each extraction and added an additional 50 µL of Qiagen Buffer ATL to 750 µL of Qiagen Buffer C1 during the lysis step. We additionally prepared one CTAB and PowerSoil Kit extraction negative, using 2 µL MilliQ water as a blank input.

Low-coverage whole-genome shotgun sequencing libraries were prepared utilizing unloaded Tn5 transposase according to a protocol developed by our lab (full protocol in Supplemental File 2). While a Tn5 transposase has been used in other studies to prepare libraries, these previous studies all relied on either pre-loaded Tn5 transposase and/or Nextera kits which increase costs and reduce flexibility (e.g., Therkildsen & Palumbi, 2017). Our protocol uses an unloaded Tn5 transposase allowing us to design our own adaptors that work with any previously designed barcode system (in our case, Wang et al., 2012). Briefly, we create the tagmentase complex by loading the Tn5 transposase with our custom designed adaptors that are complementary to the 19 bp consensus region on the Tn5 transposase as well as to the barcodes. After that incubation, the tagmentase complex is diluted to 50 ng/µl before being added to DNA (2.5 total ng). In a 10-minute 55°C incubation, the enzyme shears the DNA and adds the adaptor sequences in the same step. DNA is then amplified and barcoded in one PCR reaction. Libraries can then be pooled and quantified for sequencing submission. An entire plate of samples can be prepped in less than 4 hours, of which approximately 1.5 hours is hands-on time. All libraries were amplified along with a PCR negative control, which consistently showed no product. We sequenced libraries at the University of Texas Genomic Sequencing and Analysis Facility on the Illumina NovaSeq SP with a target of five million PE 100bp reads per sample.

Read Processing

Illumina adapters were trimmed from paired reads with Trim Galore using default settings and allowing the autodetection of adapters (https://github.com/FelixKrueger/TrimGalore). The remaining “mosaic” portions of adaptors, designed to achieve base call diversity at the beginning of reads, were trimmed using Cutadapt (Martin, 2011).

Sediment and water column community read processing and standardization

SingleM was used to profile our shotgun metagenomes from the processed reads (Woodcroft et al., 2024). SingleM uses hidden markov models to identify single copy archaeal, bacterial, and cross-domain OTUs in metagenomic sequencing reads. This improves abundance estimation and enables the removal of eukaryotic host contamination from microbial community data. In brief, singleM was run on paired-end inputs, using the “smafa_naive_then_diamond” assignment method, and requiring a

minimum taxon coverage of 0.1 in the consensus profile table. Both the consensus taxonomic profile table and the raw OTU table were output from singleM. The raw OTU table was used for subsequent community-level analyses and the consensus profile table for genus-level analyses. SingleM did not identify any OTUs in negative controls for CTAB and PowerSoil extractions.

For the community-level analysis, any OTUs that were classified as Metazoan were removed, as these likely represent coral or human contamination. Sequencing effort was corrected for each OTU, such that:

$$[Adjusted\ OTU\ Abundance]_{SampleN} = \frac{[OTU\ Abundance]_{SampleN}}{[Reads\ Sequenced]_{SampleN}} * 1e6$$

Each substrate type (sediment and water, we did not recover coral-associated OTUs) was analyzed separately and we removed OTUs that were not present in at least three samples. Many OTUs were highly covarying across samples, likely because they represented different marker genes from the same microbial species identified by singleM. To reduce the computational burden of our models, OTUs were clustered based on their sample similarity to identify OTUs which frequently co-occurred at similar abundances in samples. We created a hierarchical clustering tree of the Euclidean distances between center log ratio standardized OTU abundances using the method “complete” and set a clustering threshold at tree height (h) = 0.05. The OTUs in each group were aggregated based on their unstandardized median abundance, reducing the 6,576 OTUs returned by singleM to 4,391 OTU groups. Each group consisted of a median of 3 OTUs. When assigning taxonomy to these OTU groups later, we took the lowest level of consensus taxonomic assignment across ⅔ of the OTUs in the group. Seventy-three percent of OTU groups had a consensus call at the genus level.

Using R package ‘vegan’, the OTU group table was standardized using a centered log ratio adding a pseudocount of one and then calculated the Euclidean distances (Oksanen et al., 2024).

Sediment and water column community analysis

In an initial analysis of the drivers of community structure, we conducted PERMANOVA tests using vegan’s ‘adonis2’ function. For the sediment community, we used the model, ‘Community Distance~Site + Depth + Time Point + Site:Time Point’. Since all water samples are from the same depth at each site, we used the model, ‘community distance ~ Site + Time Point + Site:Time Point’. Terms were added from left to right.

Since depth gives partially redundant information to both site and time point (depth ranges at CB and DE are mutually exclusive, some depths only sampled at some time points), a

225 second analysis was conducted at just Subsite 1 (CB1-1 and DE1-1), where sampling was
226 guaranteed to be spatially consistent through time at each reef. For this analysis, we
227 conducted a PERMANOVA test on the subset sediment data with the model, 'community
228 distance ~ Site + Time Point + Site:Time Point'.

229 To elucidate drivers of coral reef metagenomic community structure, we built gradient
230 forest models to estimate the overall importance of time since spawning, depth, sampling
231 site, temperature, and host genotype in determining which OTUs were abundant. In brief,
232 gradient forest analysis is a multivariate extension of random forests and enables the
233 investigation of complex drivers of community structure. Gradient forest analysis identifies
234 predictor variables and microbial taxa that are the most associated with community
235 change simultaneously (Ellis et al., 2012). Importantly, gradient forest does not assume
236 linearity, and allows the effect of coral spawning on the community to be non-monotonic
237 through time (for example, rise followed by fall). For each substrate type, we built a
238 separate gradient forest model to predict the drivers of community composition based on
239 the variables we measured for the given substrate. Additionally, we visually inspected the
240 cumulative importance of 'days post spawn' to assess how community structure changed
241 at different points through time.

242 After identifying drivers of community composition, the mrIML package was used to
243 determine which microbial OTUs were most associated with the spawning event and
244 changes in sea surface temperature (Fountain-Jones et al., 2021). Accumulated local
245 effects (ALE) were calculated for OTUs in response to the mean sea surface temperature
246 (meanSST) and days post spawning (dps). ALE plots facilitate understanding of how
247 individual predictor variables affect the response of a machine learning model, while
248 controlling for cross-correlation and interactions among predictors. We calculated the
249 standard deviation of ALE scores across dps or meanSST for each OTU to determine which
250 OTUs were most variable with respect to our predictor of interest. We visualized how the
251 top 1% of OTUs with the highest standard deviation changed over the range of the predictor
252 variable. We only visualized OTUs for which singleM determined a taxonomic classification
253 at any level.

254 *Coral-associated community analysis using Bracken*

255 Unfortunately, singleM did not identify sufficient marker genes in our coral reads to take
256 into downstream analysis. Therefore, we additionally classified all our samples using
257 Kraken2 against the prebuilt "PlusPF – updated 6/5/2024" database, which contains
258 Refseq archaea, bacteria, viral, plasmid, fungi, and protozoa sequences (Wood et al.,
259 2019; database from: https://benlangmead.github.io/aws-indexes/k2). We then used
260 Bracken to estimate genus-level abundances of our classified reads (Lu et al., 2017).

Coral community analysis

A similar analysis as detailed above was conducted for coral samples, using the Bracken abundances at the genus level as an input rather than the singleM OTU table. We removed any genera that occurred in our extraction negative controls. To predict the drivers of taxon abundances, a gradient forest model was built with genotype, site, meanSST, and dps as predictors. The same workflow outlined above was then followed to investigate the cumulative importance of dps to our model and determine which genera were most associated with dps and meanSST. We ran a PERMANOVA test with the model, 'Community Distance ~ Genotype + Site + Time Point + Site:Time Point'.

Determining covarying taxa across substrate types

To connect across substrate types, we used Bracken-calculated abundances for coral, sediment, and water samples. Using R package 'vegan', the OTU table was standardized using a centered log ratio adding a pseudocount of one. We sought to determine which genera covaried in the sediment and coral communities across time. Our overall analysis suggested that highly local factors (e.g., within-reef location) affected community dynamics, so we restricted our analysis to just subsites CB1-1 and DE1-1, which were immediately next to our tagged coral colonies and sampled in triplicate at each timepoint (compared to the other subsites we sampled, which remained constant in depth, but did not have a tagged location on the reef). One sediment sample was randomly selected per timepoint and one coral genotype was selected randomly across timepoints and used to calculate the covariance between the coral and sediment microbial genus abundance values over time. We iterated this process 100 times at each site for each genus, sampling sediment samples and coral genotypes with replacement, and calculated the mean and standard error in the covariance values across these replicates. We ran this analysis for CB and DE separately. Then, we visualized how the mean abundance of four most covarying genera changed in each substrate type over our study window.

Pathway Analysis

Pathway abundances were determined in the shotgun sequencing reads using the HUMAnN 3.0 pipeline (Beghini et al., 2021). We concatenated our paired end reads for each sample into one file which was classified. For the coral samples, only reads which did not map to the *Orbicella faveolata* genome were included in the analysis (Young et al., 2024). The unstratified output from HUMAnN was used for downstream analyses, as it could not assign taxonomy to many of the samples. Finally, pathway counts were normalized based on their relative abundance as an input into MaAsLin2

295 (human_renorm_table, with option "—units relab"). We removed any pathways that were
296 unmapped, unintegrated, or engineered in the lab from the analysis.

297 Using the same method as described above, normalized pathway counts were used to
298 build a gradient forest model (pathway abundance ~ Substrate Type + Depth + Site +
299 meanSST + dps). Similarly, we determined which pathways were most associated with dps
300 and meanSST by calculating each pathway's ALE values across the predictor's ranges.

301 *Metagenomic Assembly*

302 We co-assembled contigs across samples from each substrate type separately using
303 MEGAHIT (Li et al., 2015). For coral and water samples, MEGAHIT was run under 'meta'
304 mode (--min-count 2 --k-list 21,41,61,81,99). For sediment samples, which we expected to
305 have higher complexity, MEGAHIT was run under 'meta-large' mode (--min-count 2 --k-list
306 27,37,47,57,67,77,87). Contigs >2500bp were kept for subsequent steps. We then mapped
307 our raw reads back onto the contigs assembled from their substrate type using bowtie2
308 ('local' mapping mode) and calculated contig coverage depth. We binned contigs into
309 MAGs in parallel using MetaBat2 (Kang et al., 2019) and Concoct (Alneberg et al., 2014).
310 We used DAS_Tool (Sieber et al., 2018) to produce our final MAGs from their combined
311 outputs and ran CheckM (Parks et al., 2015) on these MAGs to evaluate completeness and
312 contamination. We report MAGs with >75% completeness and <5% contamination. Finally,
313 we assigned taxonomy to the high-quality MAGs using GTDB-tk (Chaumeil et al., 2022).

314 **Results**

315 *Fine-scale spatiotemporal and genetic factors drive large differences in reef metagenomic*
316 *communities*

317 We first looked at how day-scale time, meter-scale depth, and host genotype affected the
318 metagenomic community we recovered in the coral host, reef sediment, and water
319 column. For each substrate type, we constructed a separate gradient forest model to
320 evaluate the importance of mean sea surface temperature (meanSST), days post-
321 spawning (dps), sampling site, sampling depth (for sediment), and genotype (for coral
322 colonies). For the coral colonies, we determined that genotype was the most important
323 predictor of the metagenome community, followed by dps and mean SST. Site was the
324 least important predictor (Supplemental Table 1, Figure 3a). For the sediment community,
325 the depth the sample was collected from was the most important predictor, though dps,
326 and meanSST were similarly important to the model, followed by sampling site
327 (Supplemental Table 2, Figure 3a). For the water column, dps, meanSST, and Site had
328 similar importances to the model (Supplemental Table 3, Figure 3a). Based on the

cumulative importance gained by the model over the spawning window, we determined that the sediment community was most affected by the spawning event (Figure 3b).

In all cases, our PERMANOVA analysis yielded a qualitatively similar result: that fine-scale genetic, spatial, or temporal factors were as important as between-site differences. Notably, in our PERMANOVA analyses, we recovered the significant effect of the interaction of site and time point in structuring the sediment community, suggesting CB and DE have different environmental dynamics over time (Supplemental Figure 1, Supplemental Tables 4 - 6).

For each substrate type, we also identified the taxa which were most associated with dps and meanSST based on their ALE values (Supplemental Figures 4 - 6). These taxa were predominantly bacterial in origin, though we did infer an increase in eukaryotic *Naegleria* spp. in coral tissue at temperature >29.2°C.

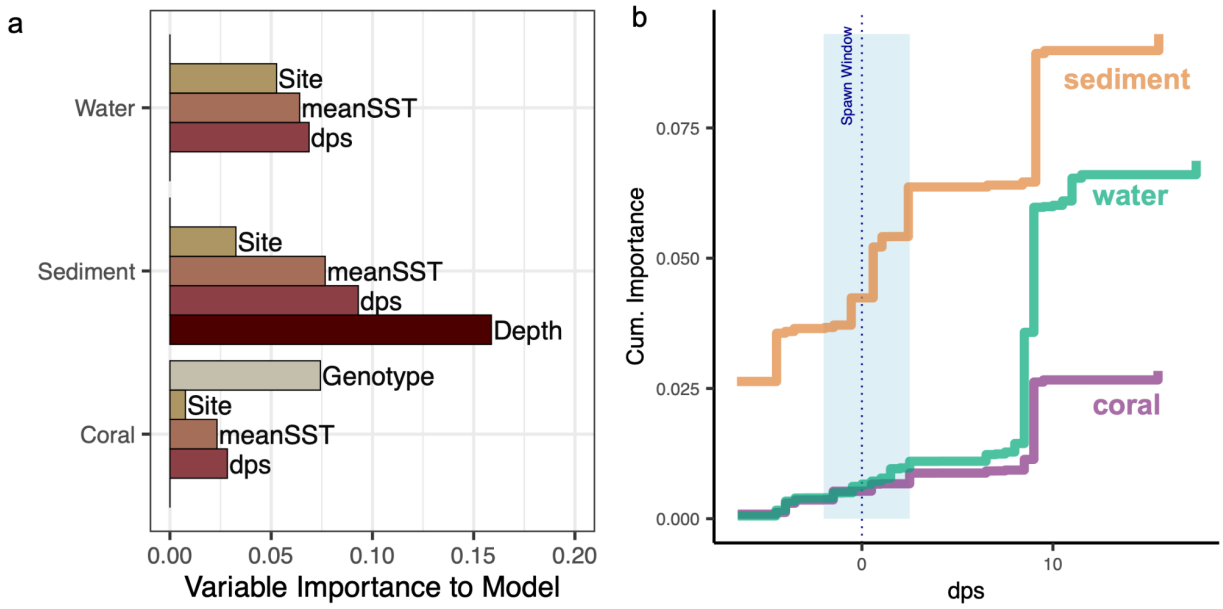


Figure 3. Fine-scale spatial, temporal, or genetic information is most important in predicting metagenomic communities for all substrate types, but the impact of spawning varies. (a) R^2 weighted variable importances to gradient forest model for site, days post spawning (dps), within-site depth and genotype for the different substrate types. Fine-scale differences play a large role in each case, with days post spawning (dps) playing a large role for both the sediment and water column communities. For the coral host, genotype appears to play the largest role in predicting metagenome community. (b) The cumulative importance of dps to each substrate type. The sediment community shows the largest importance gained over the spawning window, indicating that including these days in the model was important to predicting the overall metagenomic community. The same trend is not reflected for the water and coral communities, which are most influenced by timepoints outside of the spawning window.

Temporal factors are more important than between-site differences in determining the sediment metagenomic community

From evaluating the cumulative importance plot for depth, we found that depth accumulated most of its importance to the model at ~20-35 m (Supplemental Figure 2), a range of depths which only occurs at CB and not at DE. Therefore, to better compare how temporal factors and sampling sites influenced the reef sediment metagenomic community outside of cross-reef variation, we restricted our analysis to just subsites CB1-1 and DE1-1 (which were tagged for repeated sampling at each timepoint).

Restricting the analysis in this way, our gradient forest model found meanSST and dps most heavily influence the sediment community, with sampling site (CB/DE) much less important to the model (Supplemental Figure 3, Supplemental Table 7). This was supported by a parallel PERMANOVA analysis, which found the site by timepoint interaction to have the largest significant effect on the sediment community at CB1-1 and DE1-1 (Supplemental Table 8).

Evidence for parallel temporal change in metagenomic community across coral and sediment

To evaluate concurrent changes across the substrate types, we calculated the covariance between microbial genera in the sediment and coral across timepoints at the same location on the reef (sites CB1-1 and DE1-1). At both sites, there were over 2,000 genera present, and we found many (>100 at each site) which likely exhibited parallel dynamics through time in the coral-associated and sediment communities (Supplemental File 3). Here, we visualize just the top 30 covarying genera at CB and DE and explore abundance changes in the top four over time. The four most covarying genera across the sediment and coral communities at CB were *Alteromonas, Vibrio, Methanocaldoccus,* and *Pseudoalteromonas* (Figure 4a). All of these taxa exhibited a change in abundance in one or both substrate types across the spawning event (Figure 4b). At DE, the four most covarying genera across the sediment and coral communities were *Lactiplantibacillus, Moorena, Granulicatella,* and *Glutamicibacter* (Figure 5a).

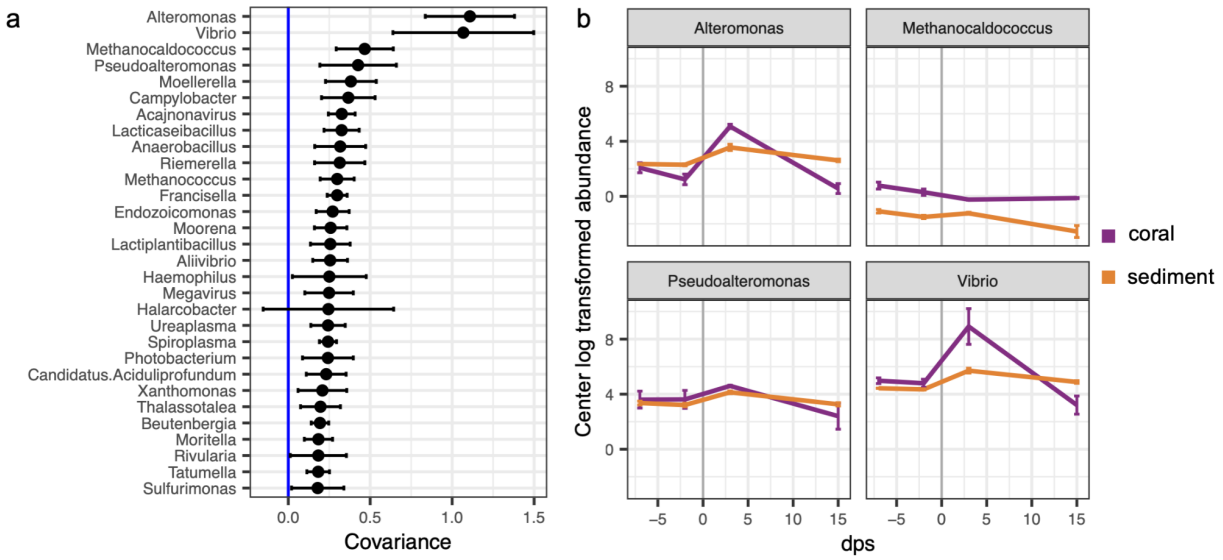


Figure 4. Top 30 highly covarying genera between the sediment and coral communities at CB. (a) Mean and standard error of the through-time covariance determined via iterating over a randomly selected coral genotype and sediment sample 100 times. (b) Visualization of the mean abundance across time (center log transformed) for the top four most covarying genera across the coral and sediment communities. Error bars give the 95% confidence interval around the mean abundance. The vertical line indicates the spawning event.

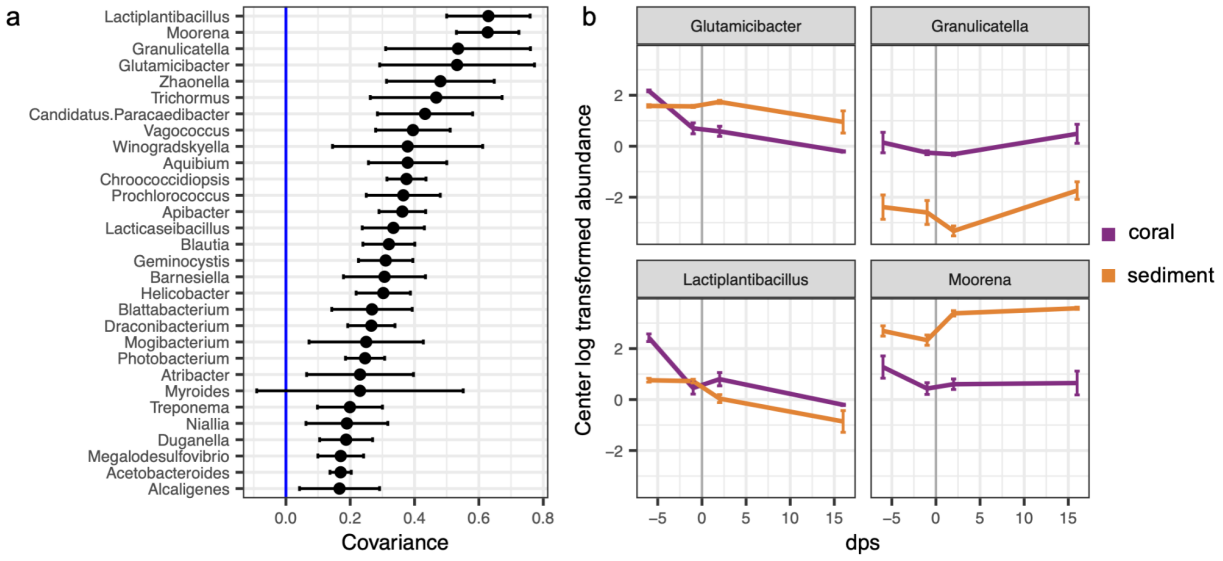


Figure 5. Top 30 highly covarying genera between the sediment and coral communities at DE. (a) Mean and standard error of the through-time covariance determined via iterating over a randomly selected coral genotype and sediment sample 100 times. (b) Visualization of the mean abundance across time (center log transformed) for the top four most covarying genera across the coral and sediment communities. Error bars give the 95% confidence interval around the mean abundance. The vertical line indicates the spawning event.

Metabolic changes in the reef environment due to the spawning event and temperature change

Using the metabolic pathways detected in our shotgun sequencing reads, we investigated functional shifts in the coral reef metagenome over our sampling window. We built a gradient forest model to predict pathway abundance and found that substrate type was the most important variable in the model, followed by depth, sampling site, meanSST and dps (Figure 6a, Supplemental Table 9).

To infer which pathways were most related to the spawning event and temperature shifts, we analyzed how the accumulated local effects (ALE) plots changed over the sampling range in the top 5% of the most variable pathways for both dps and meanSST. ALE plots facilitate understanding of how individual predictor variables affect the response of a machine learning model, while controlling for cross-correlation and interactions among predictors. In our case, an increase in ALE corresponds to predicted increased abundance of a given metabolic pathway or taxon at a parameter value or vice versa. We found the ALE of superpathway of 5'-phosphate biosynthesis and salvage to be elevated across the spawning window, whereas the glycolysis IV and rubisco shunt pathways qualitatively exhibited a decline in ALE across spawning (Figure 6b, top). In response to elevated temperatures (>29.2°C), it appeared that several pathways exhibited a decline in ALE, including the glycolysis IV, 2-amino-3-carboxymuconate semialdehyde degradation to 2-hydroxypentadienotate, 4-aminobutonate degradation V, L-glutamate degradation XI, and heterolactic fermentation pathways (Figure 6b, bottom).

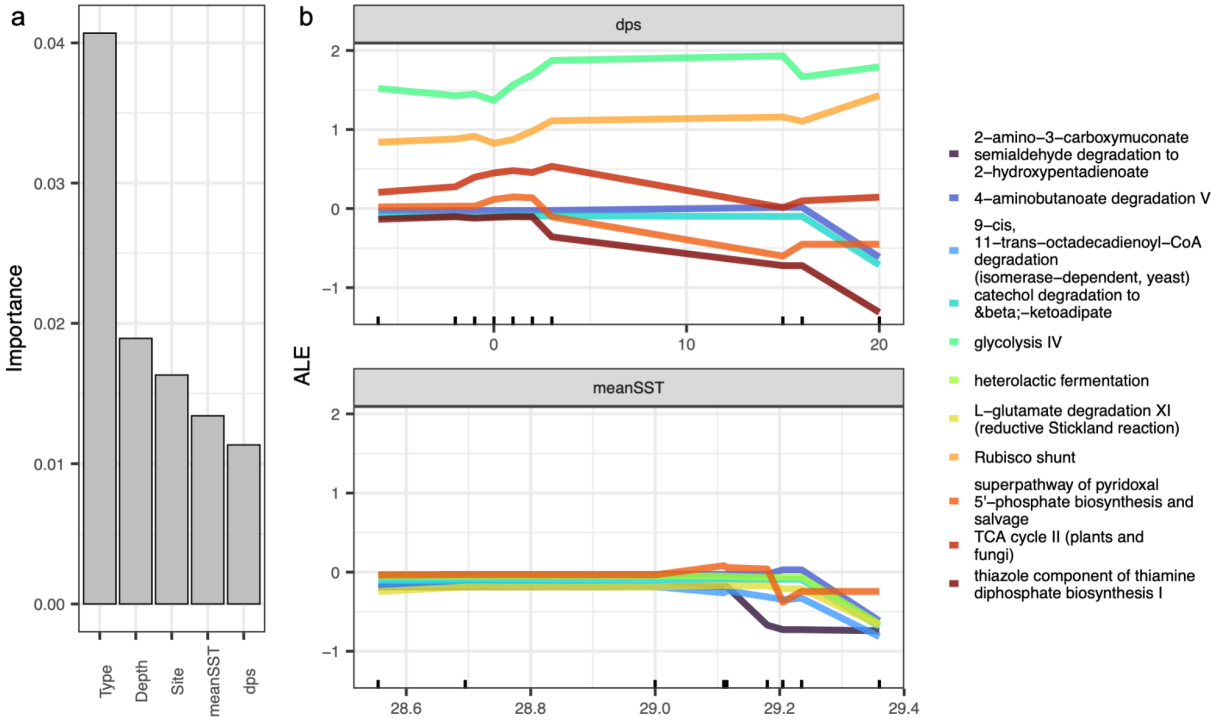


Figure 6. Metabolic pathways detected in the coral reef metagenome respond to the spawning

event and elevated temperatures. (a) Gradient forest variable R^2 weighted importance in predicting the abundance of metabolic pathways. Substrate type (Type) is the most important variable, indicating that the water column, sediment, and coral hosts harbor distinct metabolic profiles. (b) ALE plots of dps (top) and meanSST (bottom) showing the most variable 5% of metabolic pathways over the variable ranges.

15 Metagenome assembled genomes recovered from water column samples

From our whole-metagenome shotgun sequencing data, we were able to assemble 15 high completeness (>75%), low contamination (<5%) metagenome assembled genomes (MAGs; Supplemental Table 10). All 15 of these MAGs were from our water column samples as MAGS from sediment samples did not pass our completeness or contamination threshold and we were not able to bin any contigs from our coral samples. Of the 15 high quality MAGs we assembled, 13 were bacterial in origin and 2 were archaeal in origin. We were able to classify all MAGs at the genus level, but only six down to the species level, suggesting we recovered diversity not present in existing databases (Supplemental Table 10).

Discussion

By measuring the reef metagenomic community in parallel we: (1) estimated the relative role of fine-scale spatial, temporal, and genetic processes compared to large, reef-scale differences in determining community and gene family structure, (2) present evidence for parallel community changes over time across reef substrates, (3) estimate functional shifts in gene and metabolic pathway abundance in response to environmental changes, and (4) assemble 15 high-quality coral reef MAGs from our environmental sequencing data.

The relative roles of spatial and temporal processes to the reef metagenome

We find the host and environmental metagenomic communities are highly variable across time and space, with processes spanning meters or occurring over days to be as important as between-reef variation spanning kilometers (Figure 3).

For the reef sediment, we found that meter-scale variation within reef sites was the strongest predictor of the microbial community. This effect was most pronounced at CB (Supplemental Figure 2), where depth rapidly increases at the Puerto Rico Trench reef wall. Thus, proximal sampling sites (< 5m apart) across the reef may fall at distinct depths. Up to this point, there exists little quantification of reef sediment microbial community at the meter scale (though see Dong et al., 2023). Broader work on marine sediments has found that sediment microbial communities can vary with dissolved oxygen, nitrate, total organic carbon, and other environmental variables that often covary with depth (Dong et al., 2023b; Hoshino et al., 2020). With our data we cannot disentangle the degree to which

meter-scale variation in microbial communities are structured by neutral spatial processes or the local microenvironment.

With respect to time, prior work has shown that diel-level variation can exceed between-site variation in the water column metagenomic community (Wegley Kelly et al., 2019). We build on this to show that marine sediment and coral-associated microbial communities are affected by temporal processes on the same scale (Figure 3a). Our data was collected at the same time each day (between 10 am and 2pm), so these shifts are believed to represent between-day changes rather than shifts related to diel cycling. Importantly, we again found time to be the most important variable affecting sediment communities at our tagged reef patches (Supplemental Figure 3). This suggests that time point differences do not simply recapitulate variation captured by sampling different locations on the reef at each timepoint.

The effects of time were also not uniform between environmentally distinct reefs. The marine sediment had strong site by time interactions structuring its metagenomic communities (Supplemental Table 5). In some ways, this is not surprising, as CB and DE are highly distinct ecologically (Cabacungan et al., 2025), despite being geographically proximal (<20 km apart on north coastline of St. Croix). It is already clear that environment affects baseline differences between environmental and coral microbiomes (Botté et al., 2022; Deignan et al., 2023; Parker et al., 2020). However, the large interaction we find between time and location implies that the reef's microbial community response to future environmental change may not be generalizable across even geographically close reefs.

The role of genotype in coral holobiont trajectories

For coral-associated communities, we found a strong role of genotype in determining the holobiont community. While temporal factors (dps and meanSST) affected the holobiont, their role was smaller compared to host identity (Figure 3a, Supplemental Table 4). This aligns with previous work, suggesting that coral-associated microbiomes may show genotype-specific and unique responses to temporal and environmental changes, especially when compared to the coral reef environmental microbiome (Dunphy et al., 2019; Glasl, Bourne, et al., 2019). However, the strong role we find for genotype in structuring the holobiont community may partially reflect the presence of cryptic genetic lineages of *O. faveolata* on St. Croix. Since *O. faveolata* cryptic lineages on St. Croix were found to segregate entirely by location (Cabacungan et al., 2025), the strong effect of genotype we find in determining the coral holobiont may partially reflect site-specific cryptic lineages acquiring the holobiont from locally available microbes.

Opportunistic microbial taxa may covary across coral and reef sediment in response to spawning

We detected many taxa which varied across the spawning window or in response to sea surface temperature changes. We note that coral spawning often co-occurs with the hottest heat-ramp month of the year (Keith et al., 2016). We detected an approximate increase of 1ºC in meanSST which was highly correlated with proximity to the August spawning event (Figure 2), making it difficult to attribute community shifts to spawning or changing temperature, even with methods such as gradient forest which attempt to correct for this. However, changes in community structure or taxon abundance at just 1-2 days post spawn, which return to baseline conditions at two weeks post spawn, we believe to represent real, transient effects of the spawning event. Coral spawning has been previously found to affect the microbial taxa associated with hosts (Ceh et al., 2012) and these changes are likely due to the increase of organic matter in the water column (Pattern et al., 2008; Wild et al., 2007).

Within this subset of taxa, we identified many such genera which likely covaried through time across the reef sediment and coral colonies in parallel at CB and DE (Figures 4 - 5). This included changes in putative coral pathogen *Vibrio* spp. at CB. Prior studies have suggested that *Vibrio* spp. is a pathogen which increases in abundance in response to thermal stress and is positively correlated with some coral diseases (Arotsker et al., 2009; Tout et al., 2015; Wright et al., 2017). We found that *Vibrio* spp. increases in abundance immediately post-spawning in both coral and marine sediment communities, supporting prior observation that *Vibrio* abundance can change in response to mass coral spawning (Ceh et al., 2012). While we did not measure host physiology, we did not observe severe bleaching or tissue necrosis in any colonies sampled for this study. Given its increase across both substrate types and its quick return to baseline levels, we believe this increase in *Vibrio* spp. is an opportunistic response to the mass release of gametes and thus dissolved organic matter during the spawning event, rather than a hallmark of disease (Sparagon et al., 2024). Notably, this response was observed at CB and not DE (Figures 4 - 5), providing further evidence that *Vibrio* spp. dynamics may be highly dependent on the local environment (Takemura et al., 2014; Zhao et al., 2023).

Metabolic pathway abundance and taxonomic shifts in response to spawning and temperature increase

By leveraging lcWGS data, we found that temporal changes in the coral reef microbial community corresponded with putatively functional shifts in the metabolomic pathways detected in our reads (Figure 6). We found that the 5’-phosphate biosynthesis and salvage was elevated around the spawning event, suggesting a high degree of biomass degradation

being carried out by the microbial community. This parallels prior work finding functional changes in the nutrient cycling on the reef, stemming from the degradation of organic matter in the environment (Eyre et al., 2008; Simpson et al., 1993; Wild et al., 2008). In response to the concurrent increase in meanSST, we inferred a decline in several metabolic pathways in the metagenomic community. Except for the glycolysis IV pathway, the most variable pathways found in relation to shifts in meanSST are commonly attributed to bacterial taxa (Karp et al., 2019). Thus, the decreased likelihood of detecting these pathways at temperatures > 29.2°C likely reflects a change in bacterial energy metabolism during heat stress (Sun et al., 2025).

Recommendations for future study design

The large role we find for within-reef temporal, spatial, and genotype-level processes has implications for future study design. For example, given we find the effect of day-scale time can be as large as source reef (Figure 3a), corals sampled on different days may exhibit different communities not due to between-reef environmental differences, but instead within-reef temporal change. These effects are likely most pronounced at rapidly changing times of the year, such as the spawning event captured here. Future studies would benefit from: collecting multiple samples over time at geographically distinct reefs when assessing the role of the local environment in structuring the microbial community, collecting environmental samples from the same depth at each timepoint, and minimizing the time between collecting at environmentally distinct locations.

The strong influence of host genotype on microbial communities underscores the importance of tracking individual colonies over time. By following the same hosts longitudinally, researchers can better control for genetic differences, such as cryptic lineages, that might otherwise confound interpretations of holobiont community composition (e.g., Grupstra et al., 2024; Scott, Schott, et al., 2024; Ziegler et al., 2019). This approach is especially valuable in studies where it is not feasible to collect larger sample sizes from a reef site at a single time point, thereby increasing the robustness of the results.

Finally, increased implementation of lcWGS for metagenomics or metatranscriptomics on coral reefs will also advance our understanding of the role of eukaryotic and viral microorganisms in reef health. This may be particularly important in the Caribbean, which has high disease mortality and low coral recruitment (Roff, 2021; van Woesik et al., 2014). Generating a complete picture of reef microorganism biodiversity - including taxa not detectable with 16S amplicon sequencing - is crucial in defining ecosystem baselines and understanding these ecological outcomes (Pinzón et al., 2015; Rosales et al., 2022). This is

still challenging for coral colonies, as high host "contamination" limits metagenomic assembly of other taxa, as described here.

Conclusion

Corals live not in isolation, but within the dynamic framework of their surrounding marine environments. As we work towards a holistic understanding of the reef metagenomic community, it is important to understand potential feedback between hosts and their environments. Here, we showed that the metagenomic communities of different substrates on the reef are affected by different drivers and at different scales. We stress that measuring diel-scale time is imperative in accurately quantifying the reef metagenomic community, and leveraging lcWGS data provides holistic insights into both taxonomic and functional changes on the reef.

Data Availability

Raw sequence reads are deposited in the SRA (BioProject PRJNA1281116). Assembled MAGs and taxonomic annotations are available on Zenodo (doi: 10.5281/zenodo.15721723). The current version of the lab protocols, code, intermediate data products, and metadata used for analyses can be found at https://github.com/cb-scott/StCroix_ReefMetagenomics and are archived on Zenodo (doi:10.5281/zenodo.15721723).

Acknowledgements

We would like to thank Daisy Flores and Kristina Black for their assistance with sample collection in the field. Logistical support was provided by The Nature Conservancy's USVI Coral Innovation Hub. Sampling was conducted in accordance with the U.S. Virgin Islands Department of Planning and Natural Resources (The Nature Conservancy Coral Restoration Permit DFW20052X). All bioinformatic processing was run on the Texas Advanced Computing Cluster. This research was supported by NSF Grant OCE-2318775 to MVM, Human Frontiers Science Program Grant 201903079001 to MVM, a grant from the Integrative Biology Department at the University of Texas at Austin to CBS, and a grant from the International Women's Fishing Association to CBS. The funders played no role in the design or analysis of this study.

Author Contributions

CRediT: Conceptualization: CBS, ENN; Data curation: CBS; Formal Analysis: CBS; Funding acquisition: CBS, MVM; Investigation: CBS, ENN, AKH; Methodology: AKH, MVM; Resources: MVM; Software: CBS; Supervision: MVM; Validation: CBS; Visualization: CBS; Writing – original draft: CBS; Writing – review & editing: CBS, ENN, AKH, MVM

References

Alneberg, J., Bjarnason, B. S., de Bruijn, I., Schirmer, M., Quick, J., Ijaz, U. Z., Lahti, L., Loman, N. J., Andersson, A. F., & Quince, C. (2014). Binning metagenomic contigs by coverage and composition. *Nature Methods*, *11*(11), 1144–1146. https://doi.org/10.1038/nmeth.3103

Apprill, A., Marlow, H. Q., Martindale, M. Q., & Rappé, M. S. (2009). The onset of microbial associations in the coral Pocillopora meandrina. *The ISME Journal*, *3*(6), 685–699. https://doi.org/10.1038/ismej.2009.3

Apprill, A., Marlow, H. Q., Martindale, M. Q., & Rappé, M. S. (2012). Specificity of Associations between Bacteria and the Coral Pocillopora meandrina during Early Development. *Applied and Environmental Microbiology*, *78*(20), 7467–7475. https://doi.org/10.1128/AEM.01232-12

Arotsker, L., Siboni, N., Ben-Dov, E., Kramarsky-Winter, E., Loya, Y., & Kushmaro, A. (2009). Vibrio sp. As a potentially important member of the Black Band Disease (BBD) consortium in Favia sp. Corals. *FEMS Microbiology Ecology*, *70*(3), 515–524. https://doi.org/10.1111/j.1574-6941.2009.00770.x

Beghini, F., McIver, L. J., Blanco-Míguez, A., Dubois, L., Asnicar, F., Maharjan, S., Mailyan, A., Manghi, P., Scholz, M., Thomas, A. M., Valles-Colomer, M., Weingart, G., Zhang, Y., Zolfo, M., Huttenhower, C., Franzosa, E. A., & Segata, N. (2021). Integrating taxonomic, functional, and strain-level profiling of diverse microbial communities with bioBakery 3. *eLife*, *10*, e65088. https://doi.org/10.7554/eLife.65088

Botté, E. S., Cantin, N. E., Mocellin, V. J. L., O'Brien, P. A., Rocker, M. M., Frade, P. R., & Webster, N. S. (2022). Reef location has a greater impact than coral bleaching severity on the microbiome of Pocillopora acuta. *Coral Reefs*, *41*(1), 63–79. https://doi.org/10.1007/s00338-021-02201-y

Cabacungan, G. N., Kankanamalage, T. N. W., Azam, A. F., Collins, M. R., Arratia, A. R., Gutting, A. N., Matz, M. V., & Black, K. L. (2025). Cryptic coral community composition across environmental gradients. *PLOS ONE*, *20*(2), e0318653. https://doi.org/10.1371/journal.pone.0318653

Caughman, A. M., Pratte, Z. A., Patin, N. V., & Stewart, F. J. (2021). Coral microbiome changes over the day–night cycle. *Coral Reefs*, *40*(3), 921–935. https://doi.org/10.1007/s00338-021-02097-8

Ceh, J., Raina, J.-B., Soo, R. M., Keulen, M. van, & Bourne, D. G. (2012). Coral-Bacterial Communities before and after a Coral Mass Spawning Event on Ningaloo Reef. *PLOS ONE*, *7*(5), e36920. https://doi.org/10.1371/journal.pone.0036920

Chan, Y.-F., Chen, Y.-H., Yu, S.-P., Chen, H.-J., Nozawa, Y., & Tang, S.-L. (2024). Reciprocal transplant experiment reveals multiple factors influencing changes in coral microbial communities across climate zones. *Science of The Total Environment*, *907*, 167929. https://doi.org/10.1016/j.scitotenv.2023.167929

Chaumeil, P.-A., Mussig, A. J., Hugenholtz, P., & Parks, D. H. (2022). GTDB-Tk v2: Memory friendly classification with the genome taxonomy database. *Bioinformatics*, *38*(23), 5315–5316. https://doi.org/10.1093/bioinformatics/btac672

Deignan, L. K., & McDougald, D. (2022). Differential Response of the Microbiome of

Pocillopora acuta to Reciprocal Transplantation Within Singapore. *Microbial Ecology*, *83*(3), 608–618. https://doi.org/10.1007/s00248-021-01793-w

Deignan, L. K., Pwa, K. H., Loh, A. A. R., Rice, S. A., & McDougald, D. (2023). The microbiomes of two Singaporean corals show site-specific differentiation and variability that correlates with the seasonal monsoons. *Coral Reefs*, *42*(3), 677–691. https://doi.org/10.1007/s00338-023-02376-6

Dong, X., Lan, H., Huang, L., Zhang, H., Lin, X., Weng, S., Peng, Y., Lin, J., Wang, J., Peng, J., & Yang, Y. (2023a). Metagenomic Views of Microbial Communities in Sand Sediments Associated with Coral Reefs. *Microbial Ecology*, *85*(2), 465–477. https://doi.org/10.1007/s00248-021-01957-8

Dong, X., Lan, H., Huang, L., Zhang, H., Lin, X., Weng, S., Peng, Y., Lin, J., Wang, J., Peng, J., & Yang, Y. (2023b). Metagenomic Views of Microbial Communities in Sand Sediments Associated with Coral Reefs. *Microbial Ecology*, *85*(2), 465–477. https://doi.org/10.1007/s00248-021-01957-8

Dunphy, C. M., Gouhier, T. C., Chu, N. D., & Vollmer, S. V. (2019). Structure and stability of the coral microbiome in space and time. *Scientific Reports*, *9*(1), 6785. https://doi.org/10.1038/s41598-019-43268-6

Ellis, N., Smith, S. J., & Pitcher, C. R. (2012). Gradient forests: Calculating importance gradients on physical predictors. *Ecology*, *93*(1), 156–168. https://doi.org/10.1890/11-0252.1

Eyre, B. D., Glud, R. N., & Patten, N. (2008). Mass coral spawning: A natural large-scale nutrient addition experiment. *Limnology and Oceanography*, *53*(3), 997–1013. https://doi.org/10.4319/lo.2008.53.3.0997

Fountain-Jones, N. M., Kozakiewicz, C. P., Forester, B. R., Landguth, E. L., Carver, S., Charleston, M., Gagne, R. B., Greenwell, B., Kraberger, S., Trumbo, D. R., Mayer, M., Clark, N. J., & Machado, G. (2021). MrIML: Multi-response interpretable machine learning to model genomic landscapes. *Molecular Ecology Resources*, *21*(8), 2766–2781. https://doi.org/10.1111/1755-0998.13495

Glasl, B., Bourne, D. G., Frade, P. R., Thomas, T., Schaffelke, B., & Webster, N. S. (2019). Microbial indicators of environmental perturbations in coral reef ecosystems. *Microbiome*, *7*(1), 94. https://doi.org/10.1186/s40168-019-0705-7

Glasl, B., Robbins, S., Frade, P. R., Marangon, E., Laffy, P. W., Bourne, D. G., & Webster, N. S. (2020). Comparative genome-centric analysis reveals seasonal variation in the function of coral reef microbiomes. *The ISME Journal*, *14*(6), 1435–1450. https://doi.org/10.1038/s41396-020-0622-6

Glasl, B., Smith, C. E., Bourne, D. G., & Webster, N. S. (2019). Disentangling the effect of host-genotype and environment on the microbiome of the coral Acropora tenuis. *PeerJ*, *7*, e6377. https://doi.org/10.7717/peerj.6377

Grupstra, C. G. B., Meyer-Kaiser, K. S., Bennett, M.-J., Andres, M. O., Juszkiewicz, D. J., Fifer, J. E., Da-Anoy, J. P., Gomez-Campo, K., Martinez-Rugerio, I., Aichelman, H. E., Huzar, A. K., Hughes, A. M., Rivera, H. E., & Davies, S. W. (2024). Holobiont Traits Shape Climate Change Responses in Cryptic Coral Lineages. *Global Change Biology*, *30*(11), e17578. https://doi.org/10.1111/gcb.17578

Hoshino, T., Doi, H., Uramoto, G.-I., Wörmer, L., Adhikari, R. R., Xiao, N., Morono, Y.,

D'Hondt, S., Hinrichs, K.-U., & Inagaki, F. (2020). Global diversity of microbial communities in marine sediment. *Proceedings of the National Academy of Sciences*, *117*(44), 27587–27597. https://doi.org/10.1073/pnas.1919139117

Hurtado-López, N., Haines, L. & S. Woodside. (2022). Coral Spawning Predictions, The Bahamas 2022. The Bahamas Coral Innovation Hub, Perry Institute for Marine Science & Cape Eleuthera Island School, The Bahamas.

Kang, D. D., Li, F., Kirton, E., Thomas, A., Egan, R., An, H., & Wang, Z. (2019). MetaBAT 2: An adaptive binning algorithm for robust and efficient genome reconstruction from metagenome assemblies. *PeerJ*, *7*, e7359. https://doi.org/10.7717/peerj.7359

Karp, P. D., Billington, R., Caspi, R., Fulcher, C. A., Latendresse, M., Kothari, A., Keseler, I. M., Krummenacker, M., Midford, P. E., Ong, Q., Ong, W. K., Paley, S. M., & Subhraveti, P. (2019). The BioCyc collection of microbial genomes and metabolic pathways. *Briefings in Bioinformatics*, *20*(4), 1085–1093. https://doi.org/10.1093/bib/bbx085

Keith, S. A., Maynard, J. A., Edwards, A. J., Guest, J. R., Bauman, A. G., van Hooidonk, R., Heron, S. F., Berumen, M. L., Bouwmeester, J., Piromvaragorn, S., Rahbek, C., & Baird, A. H. (2016). Coral mass spawning predicted by rapid seasonal rise in ocean temperature. *Proceedings of the Royal Society B: Biological Sciences*, *283*(1830), 20160011. https://doi.org/10.1098/rspb.2016.0011

Knowlton, N., & Rohwer, F. (2003). Multispecies microbial mutualisms on coral reefs: The host as a habitat. *The American Naturalist*, *162*(4 Suppl), S51-62. https://doi.org/10.1086/378684

Krausfeldt, L. E., Lopez, J. V., Bilodeau, C. M., Won Lee, H., & Casali, S. L. (2023). Change and stasis of distinct sediment microbiomes across Port Everglades Inlet (PEI) and the adjacent coral reefs. *PeerJ*, *11*, e14288. https://doi.org/10.7717/peerj.14288

Kriefall, N. G., Kanke, M. R., Aglyamova, G. V., & Davies, S. W. (2022). Reef environments shape microbial partners in a highly connected coral population. *Proceedings of the Royal Society B: Biological Sciences*, *289*(1967), 20212459. https://doi.org/10.1098/rspb.2021.2459

Li, D., Liu, C.-M., Luo, R., Sadakane, K., & Lam, T.-W. (2015). MEGAHIT: An ultra-fast single-node solution for large and complex metagenomics assembly via succinct de Bruijn graph. *Bioinformatics*, *31*(10), 1674–1676. https://doi.org/10.1093/bioinformatics/btv033

Littman, R., Willis, B. L., & Bourne, D. G. (2011). Metagenomic analysis of the coral holobiont during a natural bleaching event on the Great Barrier Reef. *Environmental Microbiology Reports*, *3*(6), 651–660. https://doi.org/10.1111/j.1758-2229.2010.00234.x

Lu, J., Breitwieser, F. P., Thielen, P., & Salzberg, S. L. (2017). Bracken: Estimating species abundance in metagenomics data. *PeerJ Computer Science*, *3*, e104. https://doi.org/10.7717/peerj-cs.104

Ma, L., Becker, C., Weber, L., Sullivan, C., Zgliczynski, B., Sandin, S., Brandt, M., Smith, T. B., & Apprill, A. (2022). Biogeography of reef water microbes from within-reef to global scales. *Aquatic Microbial Ecology*, *88*, 81–94. https://doi.org/10.3354/ame01985

Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.Journal*, *17*(1), Article 1. https://doi.org/10.14806/ej.17.1.200

Oksanen, J., Simpson, G. L., Blanchet, F. G., Kindt, R., Legendre, P., Minchin, P. R., O'Hara, R. B., Solymos, P., Stevens, M. H. H., Szoecs, E., Wagner, H., Barbour, M., Bedward, M., Bolker, B., Borcard, D., Carvalho, G., Chirico, M., Caceres, M. D., Durand, S., … Weedon, J. (2024). *vegan: Community Ecology Package* [Computer software]. https://CRAN.R-project.org/package=vegan

Parker, K. E., Ward, J. O., Eggleston, E. M., Fedorov, E., Parkinson, J. E., Dahlgren, C. P., & Cunning, R. (2020). Characterization of a thermally tolerant Orbicella faveolata reef in Abaco, The Bahamas. *Coral Reefs*, *39*(3), 675–685. https://doi.org/10.1007/s00338-020-01948-0

Parks, D. H., Imelfort, M., Skennerton, C. T., Hugenholtz, P., & Tyson, G. W. (2015). CheckM: Assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. *Genome Research*, *25*(7), 1043–1055. https://doi.org/10.1101/gr.186072.114

Patten, N. L., Mitchell, J. G., Middelboe, M., Eyre, B. D., Seuront, L., Harrison, P. L., & Glud, R. N. (2008). Bacterial and viral dynamics during a mass coral spawning period on the Great Barrier Reef. *Aquatic Microbial Ecology*, *50*(3), 209–220. https://doi.org/10.3354/ame01179

Pinzón, J. H., Kamel, B., Burge, C. A., Harvell, C. D., Medina, M., Weil, E., & Mydlarz, L. D. (2015). Whole transcriptome analysis reveals changes in expression of immune-related genes during and after bleaching in a reef-building coral. *Royal Society Open Science*, *2*(4), 140214. https://doi.org/10.1098/rsos.140214

Reshef, L., Koren, O., Loya, Y., Zilber-Rosenberg, I., & Rosenberg, E. (2006). The coral probiotic hypothesis. *Environmental Microbiology*, *8*(12), 2068–2073. https://doi.org/10.1111/j.1462-2920.2006.01148.x

Roff, G. (2021). Evolutionary History Drives Biogeographic Patterns of Coral Reef Resilience. *BioScience*, *71*(1), 26–39. https://doi.org/10.1093/biosci/biaa145

Rohwer, F., Breitbart, M., Jara, J., Azam, F., & Knowlton, N. (2001). Diversity of bacteria associated with the Caribbean coral Montastraea franksi. *Coral Reefs*, *20*(1), 85–91. https://doi.org/10.1007/s003380100138

Rosales, S. M., Huebner, L. K., Clark, A. S., McMinds, R., Ruzicka, R. R., & Muller, E. M. (2022). Bacterial Metabolic Potential and Micro-Eukaryotes Enriched in Stony Coral Tissue Loss Disease Lesions. *Frontiers in Marine Science*, *8*. https://doi.org/10.3389/fmars.2021.776859

Scott, C. B., Ostling, A., & Matz, M. V. (2024). Should I stay or should I go? Coral bleaching from the symbionts' perspective. *Ecology Letters*, *27*(5), e14429. https://doi.org/10.1111/ele.14429

Scott, C. B., Schott, R., & Matz, M. V. (2024). *Cryptic genetic structure of the coral host is the primary driver of holobiont assembly in massive Porites* (p. 2024.01.09.574877). bioRxiv. https://doi.org/10.1101/2024.01.09.574877

Sharp, K. H., Ritchie, K. B., Schupp, P. J., Ritson-Williams, R., & Paul, V. J. (2010). Bacterial Acquisition in Juveniles of Several Broadcast Spawning Coral Species. *PLOS ONE*, *5*(5), e10898. https://doi.org/10.1371/journal.pone.0010898

Siboni, N., Abrego, D., Puill-Stephan, E., King, W. L., Bourne, D. G., Raina, J.-B., Seymour, J. R., & Harder, T. (2020). Crustose coralline algae that promote coral larval settlement harbor distinct surface bacterial communities. *Coral Reefs*, *39*(6), 1703–1713. https://doi.org/10.1007/s00338-020-01997-5

Sieber, C. M. K., Probst, A. J., Sharrar, A., Thomas, B. C., Hess, M., Tringe, S. G., & Banfield, J. F. (2018). Recovery of genomes from metagenomes via a dereplication, aggregation and scoring strategy. *Nature Microbiology*, *3*(7), 836–843. https://doi.org/10.1038/s41564-018-0171-1

Simpson, C. J., Cary, J. L., & Masini, R. J. (1993). Destruction of corals and other reef animals by coral spawn slicks on Ningaloo Reef, Western Australia. *Coral Reefs*, *12*(3), 185–191. https://doi.org/10.1007/BF00334478

Sneed, J. M., Sharp, K. H., Ritchie, K. B., & Paul, V. J. (2014). The chemical cue tetrabromopyrrole from a biofilm bacterium induces settlement of multiple Caribbean corals. *Proceedings of the Royal Society B: Biological Sciences*, *281*(1786), 20133086. https://doi.org/10.1098/rspb.2013.3086

Sparagon, W. J., Arts, M. G. I., Quinlan, Z. A., Wegley Kelly, L., Koester, I., Comstock, J., Bullington, J. A., Carlson, C. A., Dorrestein, P. C., Aluwihare, L. I., Haas, A. F., & Nelson, C. E. (2024). Coral thermal stress and bleaching enrich and restructure reef microbial communities via altered organic matter exudation. *Communications Biology*, *7*(1), 1–14. https://doi.org/10.1038/s42003-023-05730-0

Stothart, M. R., McLoughlin, P. D., & Poissant, J. (2023). Shallow shotgun sequencing of the microbiome recapitulates 16S amplicon results and provides functional insights. *Molecular Ecology Resources*, *23*(3), 549–564. https://doi.org/10.1111/1755-0998.13713

Sun, F., Yang, H., Tan, F., Zhang, X., Wang, M., Wang, G., & Shi, Q. (2025). Impact of coral bleaching on the archaeal community and carbon metabolic functions. *Coral Reefs*, *44*(2), 721–730. https://doi.org/10.1007/s00338-025-02641-w

Takemura, A. F., Chien, D. M., & Polz, M. F. (2014). Associations and dynamics of Vibrionaceae in the environment, from the genus to the population level. *Frontiers in Microbiology*, *5*. https://doi.org/10.3389/fmicb.2014.00038

Terzin, M., Laffy, P. W., Robbins, S., Yeoh, Y. K., Frade, P. R., Glasl, B., Webster, N. S., & Bourne, D. G. (2024). The road forward to incorporate seawater microbes in predictive reef monitoring. *Environmental Microbiome*, *19*(1), 5. https://doi.org/10.1186/s40793-023-00543-4

Therkildsen, N. O., & Palumbi, S. R. (2017). Practical low-coverage genomewide sequencing of hundreds of individually barcoded samples for population and evolutionary genomics in nonmodel species. *Molecular Ecology Resources*, *17*(2), 194–208. https://doi.org/10.1111/1755-0998.12593

Tout, J., Siboni, N., Messer, L. F., Garren, M., Stocker, R., Webster, N. S., Ralph, P. J., & Seymour, J. R. (2015). Increased seawater temperature increases the abundance and alters the structure of natural Vibrio populations associated with the coral Pocillopora damicornis. *Frontiers in Microbiology*, *6*. https://doi.org/10.3389/fmicb.2015.00432

Turnlund, A. C., O'Brien, P. A., Rix, L., Webster, N., Lurgi, M., & Vanwonterghem, I. (2025).

Understanding the role of micro-organisms in the settlement of coral larvae through community ecology. *Marine Biology*, *172*(3), 43. https://doi.org/10.1007/s00227-025-04607-6

van Woesik, R., Scott, W. J., & Aronson, R. B. (2014). Lost opportunities: Coral recruitment does not translate to reef recovery in the Florida Keys. *Marine Pollution Bulletin*, *88*(1), 110–117. https://doi.org/10.1016/j.marpolbul.2014.09.017

Wang, S., Meyer, E., McKay, J. K., & Matz, M. V. (2012). 2b-RAD: A simple and flexible method for genome-wide genotyping. *Nature Methods*, *9*(8), 808–810. https://doi.org/10.1038/nmeth.2023

Weber, L., & Apprill, A. (2020). Diel, daily, and spatial variation of coral reef seawater microbial communities. *PloS One*, *15*(3), e0229442. https://doi.org/10.1371/journal.pone.0229442

Wegley Kelly, L., Nelson, C. E., Haas, A. F., Naliboff, D. S., Calhoun, S., Carlson, C. A., Edwards, R. A., Fox, M. D., Hatay, M., Johnson, M. D., Kelly, E. L. A., Lim, Y. W., Macherla, S., Quinlan, Z. A., Silva, G. G. Z., Vermeij, M. J. A., Zgliczynski, B., Sandin, S. A., Smith, J. E., & Rohwer, F. (2019). Diel population and functional synchrony of microbial communities on coral reefs. *Nature Communications*, *10*(1), 1691. https://doi.org/10.1038/s41467-019-09419-z

Wegley Kelly, L., 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*, *111*(28), 10227–10232. https://doi.org/10.1073/pnas.1403319111

Wild, C., Jantzen, C., Struck, U., Hoegh-Guldberg, O., & Huettel, M. (2008). Biogeochemical responses following coral mass spawning on the Great Barrier Reef: Pelagic–benthic coupling. *Coral Reefs*, *27*(1), 123–132. https://doi.org/10.1007/s00338-007-0298-7

Williams, D., Bright, A., Grazioso, K., & Ladd, M. (2022). Coral Spawning 2022: Activities and Observations. NOAA Fisheries, Southeast Fisheries Science Center.

Wood, D. E., Lu, J., & Langmead, B. (2019). Improved metagenomic analysis with Kraken 2. *Genome Biology*, *20*(1), 257. https://doi.org/10.1186/s13059-019-1891-0

Woodcroft, B. J., Aroney, S. T. N., Zhao, R., Cunningham, M., Mitchell, J. A. M., Blackall, L., & Tyson, G. W. (2024). SingleM and Sandpiper: Robust microbial taxonomic profiles from metagenomic data. *bioRxiv*. https://doi.org/10.1101/2024.01.30.578060

Wright, R. M., Kenkel, C. D., Dunn, C. E., Shilling, E. N., Bay, L. K., & Matz, M. V. (2017). Intraspecific differences in molecular stress responses and coral pathobiome contribute to mortality under bacterial challenge in Acropora millepora. *Scientific Reports*, *7*(1), 2609. https://doi.org/10.1038/s41598-017-02685-1

Young, B. D., Williamson, O. M., Kron, N. S., Andrade Rodriguez, N., Isma, L. M., MacKnight, N. J., Muller, E. M., Rosales, S. M., Sirotzke, S. M., Traylor-Knowles, N., Williams, S. D., & Studivan, M. S. (2024). Annotated genome and transcriptome of the endangered Caribbean mountainous star coral (Orbicella faveolata) using PacBio long-read sequencing. *BMC Genomics*, *25*(1), 226.

https://doi.org/10.1186/s12864-024-10092-w

Zhao, W., Chen, X., Liu, R., Tian, P., Niu, W., Zhang, X.-H., Liu, J., & Wang, X. (2023). Distinct coral environments shape the dynamic of planktonic Vibrio spp. *Environmental Microbiome*, *18*(1), 77. https://doi.org/10.1186/s40793-023-00532-7

Ziegler, M., Grupstra, C. G. B., Barreto, M. M., Eaton, M., BaOmar, J., Zubier, K., Al-Sofyani, A., Turki, A. J., Ormond, R., & Voolstra, C. R. (2019). Coral bacterial community structure responds to environmental change in a host-specific manner. *Nature Communications*, *10*(1), 3092. https://doi.org/10.1038/s41467-019-10969-5

Supplemental Tables

Table S1. Gradient forest model importances for coral-associated community.

Variable	Importance
Genotype	0.0743
dps	0.0283
meanSST	0.0232
Site	0.0076
Model Importance:	0.134

Table S2. Gradient forest model importances for sediment community.

Variable	Importance
Depth	0.1588
dps	0.0930
meanSST	0.0767
Site	0.0325
Model Importance:	0.361

Table S3. Gradient forest model importances for water column community.

Variable	Importance
dps	0.0691
meanSST	0.0641
Site	0.0527
Model Importance:	0.185

Table S4. PERMANOVA results for coral-associated community.

Variable	Df	R2	F	Pr(>F)
Genotype	8	0.337	1.617	0.012
Time Point	3	0.121	1.543	0.061
meanSST	1	0.051	1.947	0.057
Time Point:Site	2	0.073	1.397	0.116
	Residual	0.417		

Table S5. PERMANOVA results for the sediment community.

Variable	Df	R2	F	Pr(>F)
Site	1	0.031	1.754	0.001
Depth	1	0.021	1.180	0.037
Time Point	4	0.085	1.207	0.004
meanSST	1	0.021	1.169	0.022
Site:Time Point	2	0.047	1.315	0.002
	Residual	0.795		

Table S6. PERMANOVA results for the water column community.

Variable	Df	R2	F	Pr(>F)
Site	1	0.177	2.759	0.001
Time Point	5	0.301	0.937	0.650
meanSST	1	0.071	1.108	0.294
Site:Time Point	2	0.129	1.001	0.481
	Residual	0.321		

Table S7. Gradient forest model importances for the sediment community at sites CB1-1 and DE1-1.

Variable	Importance
meanSST	0.0893
dps	0.0860
Site	0.0317
Model Importance:	0.207

Table S8. PERMANOVA results for the sediment community at sites CB1-1 and DE1-1.

Variable	Df	R2	F	Pr(>F)
Site	1	0.0577	1.348	0.001
Time Point	4	0.179	1.042	0.258
meanSST	1	0.0440	1.028	0.386
Site:Time Point	2	0.120	1.404	0.002
	Residual	0.599		

Table S9: Gradient forest model importances for the metabolic pathways identified in the metagenomic reads.

Variable	Importance
Type	0.0407
Depth	0.0187
Site	0.0175
meanSST	0.0132
dps	0.0110
Model Importance:	0.101

Table S10: Completeness, contamination, and taxonomic information on reported MAGs. All MAGs were from water column samples.

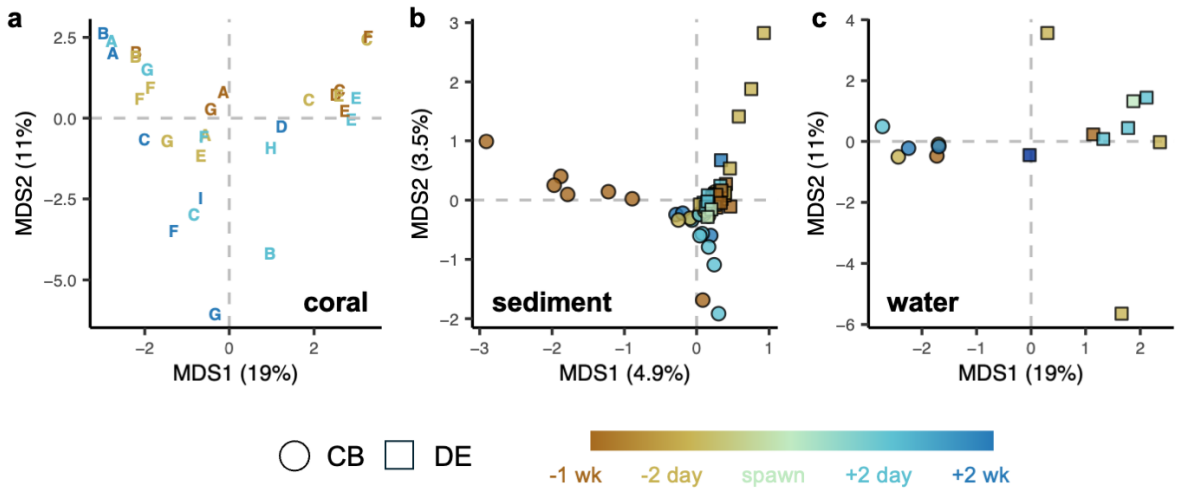
MAG_ID	Completeness	Contamination	Tax Assign
1.water.MAG	98.16	0.74	d__Bacteria; p__Bacteroidota; c__Bacteroidia; o__Flavobacteriales; f__Flavobacteriaceae; g__WTJH01
24.water.MAG	92.35	1.67	d__Bacteria; p__Pseudomonadota; c__Gammaproteobacteria; o__Pseudomonadales; f__Litoricolaceae; g__HIMB30; s__HIMB30 sp000227525
3.water.MAG	91.91	1.57	d__Bacteria; p__Pseudomonadota; c__Alphaproteobacteria; o__Puniceispirillales; f__Puniceispirillaceae; g__WTJO01
11.water.MAG	91.6	1.59	d__Bacteria; p__Bacteroidota; c__Bacteroidia; o__Flavobacteriales; f__Flavobacteriaceae; g__RFYP01; s__RFYP01 sp011526365
5.water.MAG	90.36	1.59	d__Bacteria; p__Bacteroidota; c__Bacteroidia; o__Flavobacteriales; f__Flavobacteriaceae; g__Arcticimaribacter; s__Arcticimaribacter sp011526335
17.water.MAG	87.97	2.84	d__Bacteria; p__Pseudomonadota; c__Alphaproteobacteria; o__Puniceispirillales; f__Puniceispirillaceae; g__WTJO01
7.water.MAG	87.15	0.74	d__Bacteria; p__Bacteroidota; c__Bacteroidia; o__Flavobacteriales; f__Flavobacteriaceae; g__Arcticimaribacter
8.water.MAG	86.56	3.75	d__Bacteria; p__Pseudomonadota; c__Gammaproteobacteria; o__Pseudomonadales; f__Halieaceae; g__Luminiphilus
19.water.MAG	85.33	0.68	d__Bacteria; p__Cyanobacteriota; c__Cyanobacteriia; o__PCC-6307; f__Cyanobiaceae; g__Parasynechococcus; s__Parasynechococcus sp002724845
23.water.MAG	82.87	0.81	d__Bacteria; p__Bacteroidota; c__Bacteroidia; o__Flavobacteriales; f__Schleiferiaceae; g__UBA10364; s__UBA10364 sp002430755
14.water.MAG	82.24	2.43	d__Bacteria; p__Pseudomonadota; c__Alphaproteobacteria; o__Puniceispirillales; f__Puniceispirillaceae; g__JADHLI01
16.water.MAG	81.96	0.8	d__Archaea; p__Thermoplasmatota; c__Poseidoniia; o__Poseidoniales; f__Poseidoniaceae; g__MGIIa-L4
20.water.MAG	78.4	2.1	d__Archaea; p__Thermoplasmatota; c__Poseidoniia; o__Poseidoniales; f__Poseidoniaceae; g__Poseidonia

22.water.MAG	78.26	2.94	d__Bacteria; p__Bacteroidota; c__Bacteroidia; o__Flavobacteriales; f__Flavobacteriaceae; g__Arcticimaribacter
6.water.MAG	75.51	2.56	__Bacteria; p__Bacteroidota; c__Bacteroidia; o__Flavobacteriales; f__UA16; g__UBA11663; s__UBA11663 sp011524845

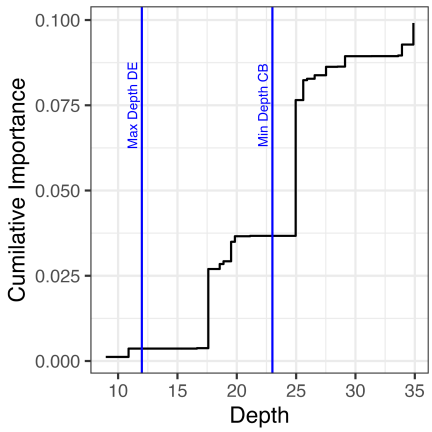
887

888

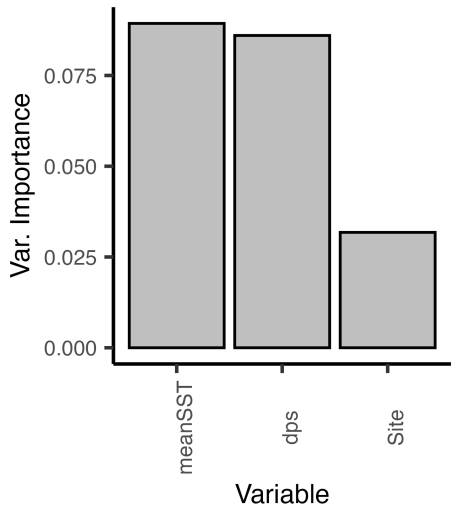
Supplemental Figures



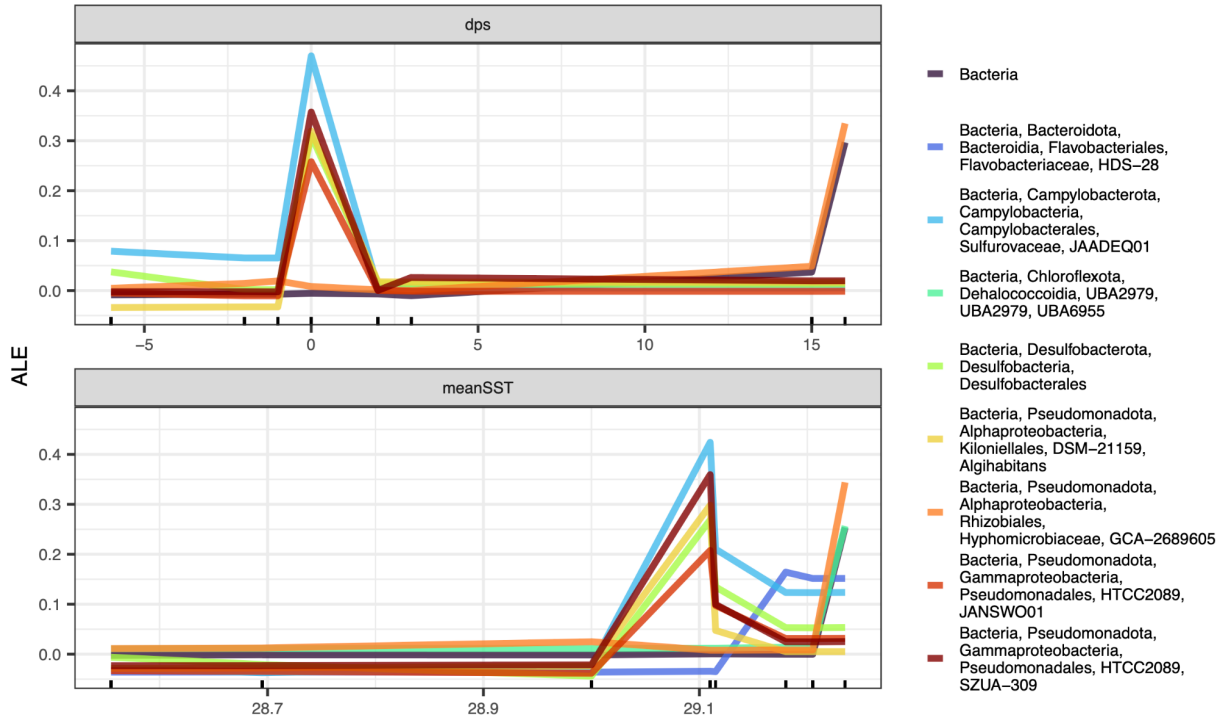
Supplemental Figure 1. PCoA structure of different reef substrate communities colored by time point. (a) PCoA of sediment community structure. Sediment was sampled at three depths at each time point at each reef. (b) PCoA of water column community structure. Water column samples were collected from the same within-reef location at each sampling time point. (c) PCoA of coral holobiont community structure. Like the water column, all coral samples came from the same within-reef location across time points. Genotypes A-C are from CB, genotypes E-G are from DE.



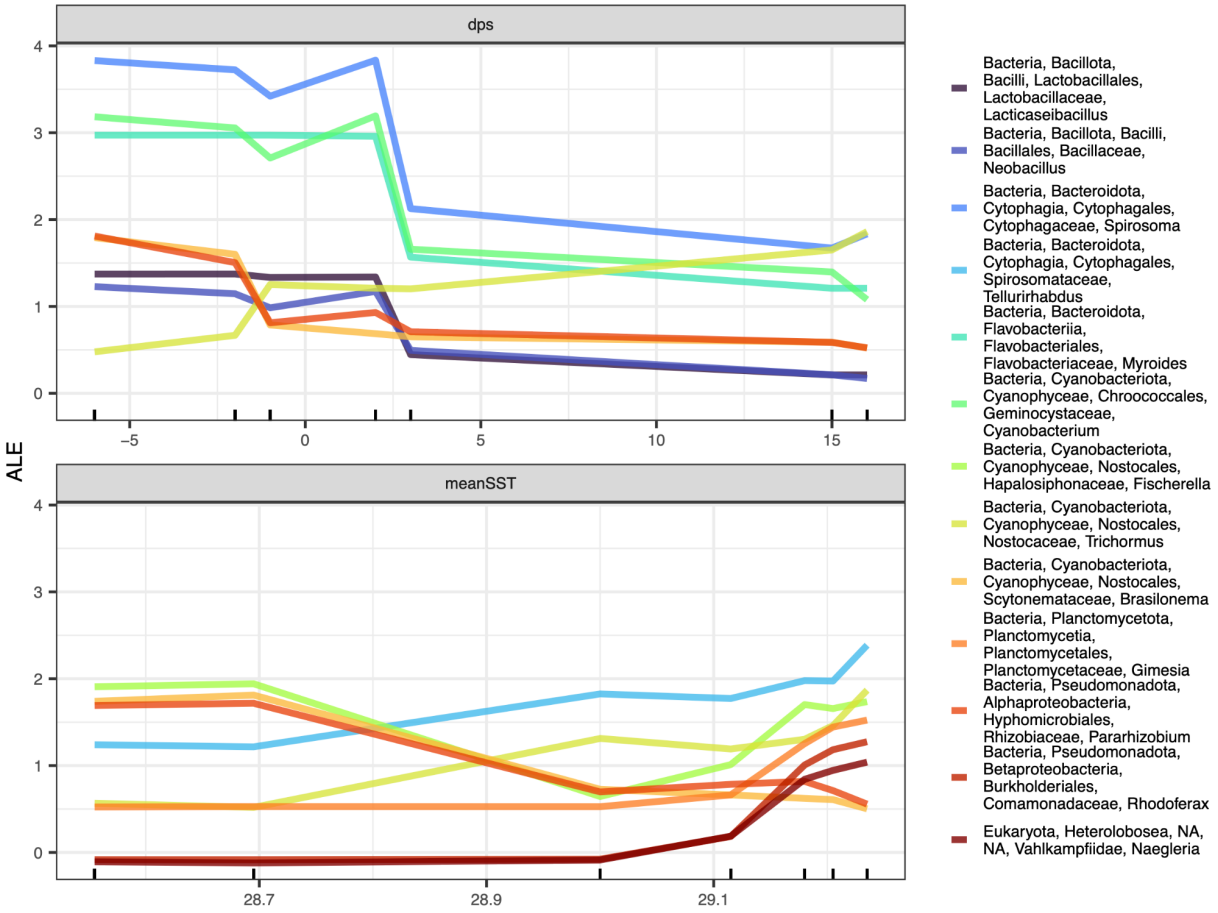
Supplemental Figure 2. Depth cumulative importance to sediment gradient forest model. Cumulative importance of depth to the sediment metagenomic community gradient forest model. CB and DE have mutually exclusive depth ranges; this is indicated by the horizontal lines on the plot giving the maximum sampled depth at DE and minimum sample depth at CB.



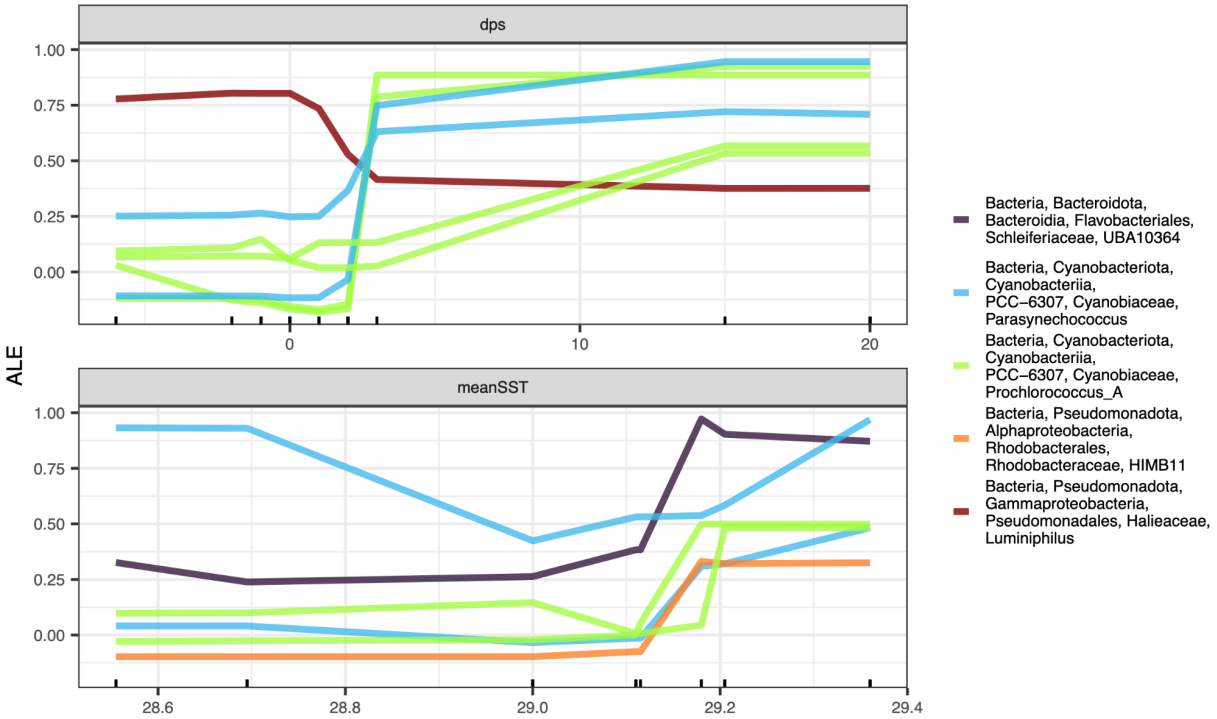
Supplemental Figure 3. Gradient forest model results for the sediment community at sites CB1-1 and DE1-1.



Supplemental Figure 4. ALE over dps and meanSST parameter ranges for the top 1% most variable taxa in the marine sediment.



Supplemental Figure 5. ALE over dps and meanSST parameter ranges for the top 1% most variable genera from coral tissue.



Supplemental Figure 6. ALE over dps and meanSST parameter ranges for the top 1% most variable taxa from water column.