Changes in the Coral Microbiome in Response to Restoration Attempts

Abstract: Coral reefs are the most diverse marine environments on the planet and yet are one of the most endangered. Specifically in the Caribbean, 50% of all coral cover has been lost due to anthropogenic actions such as pollution. Some reefs, however, have been shown to be resilient in the face of pollution, like Varadero Reef in Columbia. Varadero is exposed to a constant plume of turbidity and yet has considerable coral cover. To test whether the microbiome changes in response to pollution events, coral was harvested from two reefs (Varadero (medium pollution levels) and Rosario (low pollution levels)) and transplanted across reefs and to an area of high pollution, Cartagena Bay. This project aims to re-analyze 16s rRNA sequence data collected from this transplant using an updated pipeline and reference database to illustrate the differences between bacterial families identified. In addition, the merits of using OTU and ASV identification of sequences will be compared. The reanalysis of metagenomic data using new and improved pipelines can provide major insights on changes in the microbiome over time and can inform conservation efforts in the present.

Introduction

Coral reefs are one of the most diverse marine environments on the planet. Though they make up less than 1% of the world's oceans, they are the habitat for at least 25% of all marine life (US Department of Commerce n.d.). Reefs are considered oceanic oases of life and can host many species, some of which are endangered and depend on reef environments for food sources (US EPA 2017). Besides providing an essential habitat for marine life, coral reefs also provide a home to economically important species of fish that depend on coral reefs for parts of their lifestyle, specifically shelter for many species juvenile stages (US Department of Commerce n.d.). Reefs support the local economy by providing natural resources and locations for tourism (US Department of Commerce n.d.). In addition, reefs also help to reduce storm surges and minimize damage like flooding and erosion (Patrick et al. 2020).

Despite their great importance to the environment and to humans, reefs have been suffering huge losses and degradation. This loss is caused by anthropogenic influence. From climate change, pollution, overfishing, oil drilling, tourism, and mechanical stress caused by boat traffic, almost every human activity is harmful to reefs (US EPA 2017). This is especially true in areas with large marine tourism-based economies, like the Caribbean. The coral reefs in the Caribbean have suffered a loss of 50% of their coral cover since the 80s (Gardner et al. 2003). Environmental stressors driving these losses include high sediment levels, increased turbidity, and low light availability (Roitman et al. 2020). These stressors increase the likelihood of disease and coral bleaching, making the decline of these reefs almost certain. However, there have been unusual occurrences of turbid estuary-like reefs in Australia inland of the Great Barrier Reef and outside of Barrow Island (Browne et al. 2013, Fisher et al. 2019). The existence of these reefs brings into question the susceptibility of types of corals to environmental stressors and why some are resilient in areas with high disturbance rates.

Varadero Reef is an example of one such resilient reef. Located at the entrance of Cartagena Bay, in the Colombian Caribbean, Varadero Reef is the small remaining reef of a once dominant reef system spreading across the bay and surrounding areas (Roitman et al. 2020). Due to the discharge from the nearby Canal del Dique, Varadero reef is exposed to a constant plume of salinity and turbidity (Roitman et al. 2020). Despite these conditions, Varadero is home to many diverse species and high levels of coral cover, making it a perfect candidate for investigating the relationship between the coral, its microbiome, and its resilience to constant exposure to stressors (Roitman et al. 2020).

Examining the response to stressors in coral requires a multifaceted approach. Due to the relationship between the coral, its symbiotic algae, and the microbiome, it is essential to

investigate each part of the system and its response to the stressors. Knowledge of the interaction between each part of this system is a determining factor in the effectiveness of conservation methods and the ability to restore already degraded reefs (Peixoto et al. 2017).

Focusing on the microbial part of the coral, algae, and microbiome system the original study aims to investigate the change in the microbiome after transplantation to areas with stressors (salinity and turbidity) and asses the role it plays in the survivorship of the coral (Roitman et al. 2020). To do so, the coral microbiomes of three different sites with a gradient of stress exposure were compared (Roitman et al. 2020). In addition, corals were transplanted across sites to determine the change in microbiome after an acclimation period (Roitman et al. 2020). These results provide information on how restoration practices affect the coral microbiome and its role in determining the survivorship of corals in stressed environments, allowing the development of more effective reef restoration practices.

In this study, the aim is to analyze a subset of the data (the microbiome across transplantation sites) through QIIME2-2024.2. In addition, this study aims to compare analysis methods of amplicon sequences across bioinformatic pipelines with the original study focusing on OTUs using older versions of QIIME and UCLUST/ USEARCH packages, and the reanalysis focusing on ASVs using more recent versions of QIIME and more recently developed packages. Because it is known that ASV identification is more accurate and reproducible than OTU identification ("OTU vs. ASV in Metagenomic Analysis" n.d.), the reanalyzed data should show identification of more species than the original paper. This will include renamed or newly identified families.

Methods

Sequence Acquisition

This data was generated from the paper *Surviving Marginalized Reefs: Assessing the Implications of Microbiome on Coral Physiology and Survivorship* available at NCBI accession under BioProject ID: PRJNA639618**.** Data includes amplicon sequences of the 16S rRNA gene (V4 region) sequenced on the Illumina MiSeq platform.

In brief, corals were collected and transplanted to/from a gradient of stress exposure: Rosario Reef (low disturbance) (10° 11' 12.1" N, 75° 44' 43.0" W), Varadero Reef (medium disturbance) (10° 18' 23.3" N, 75° 35' 0.80" W), and Cartagena Bay (high disturbance) (10° 18' 04.5" N, 75° 34' 38.5" W). Cartagena Bay currently has no coral due to high turbidity and low light availability, though there once was coral there.

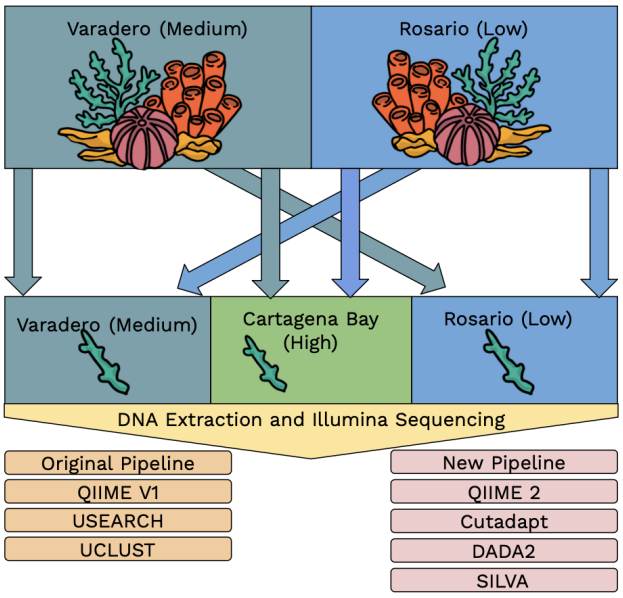

Figure 1: Experimental Design. Coral from reefs Varadero (medium levels of pollution) and Rosario (low pollution levels) were transplanted across reefs and to Cartagena Bay (high pollution levels). DNA was extracted from the corals and sequenced. Then bioinformatic pipelines were compared with the original in orange and the new pipeline in pink.

The coral species *Orbicella faveolata* was chosen as the transplant candidate. Flat fragments of this coral were collected from donor sites (Varadero and Rosario) (n=45). After an acclimation period, these fragments were transplanted to each of the three sites (Rosario Reef- low disturbance, Varadero Reef- medium disturbance, and Cartagena Bay- high disturbance) from their donor sites (Varadero and Rosario) in equal proportions and sampled before and after transplantation (7 months). Extracted

DNA was sent 2x 250 bp paired end sequencing on the Illumina MiSeq sequencing platform at the DNA Services Facility at the University of Illinois at Chicago.

.

16S Sequence Analysis

Sequences were reprocessed using QIIME2-2024.2. The primers and adapters were trimmed using *Cutadapt,* and low-quality reads were removed before merging via DADA2 (Callahan *et al.* 2016). Reads were assigned to ASVs through DADA2 and using SILVA (Quast et al. 2013). Original analysis was performed in QIIME; v1.9 using the following pipeline: chimeric sequence removal and OTUs similarity analysis (97%) in USEARCH 6.1 and assignment of taxonomy using UCLUST (Edgar 2010).

Statistical analysis was completed in R (v4.4.1; R Core Team 2024). Shannon Diversity Index was used to determine diversity pre- and post-transplant and was visualized via boxplot and assessed by ANOVA. To compare community composition between sites and determine which site had the most variation in community composition, a PERMANOVA was run using the phyloseq package in R (Oksanen et al. 2013, v4.4.1; R Core Team 2024). In addition, changes to community composition were visualized by PCoA. To determine how species' abundances change with or without transplant, a differential abundance analysis was performed using ANCOMBC (Lin and Peddada 2020, v4.4.1; R Core Team 2024). Results were visualized using a relative abundance chart.

Results

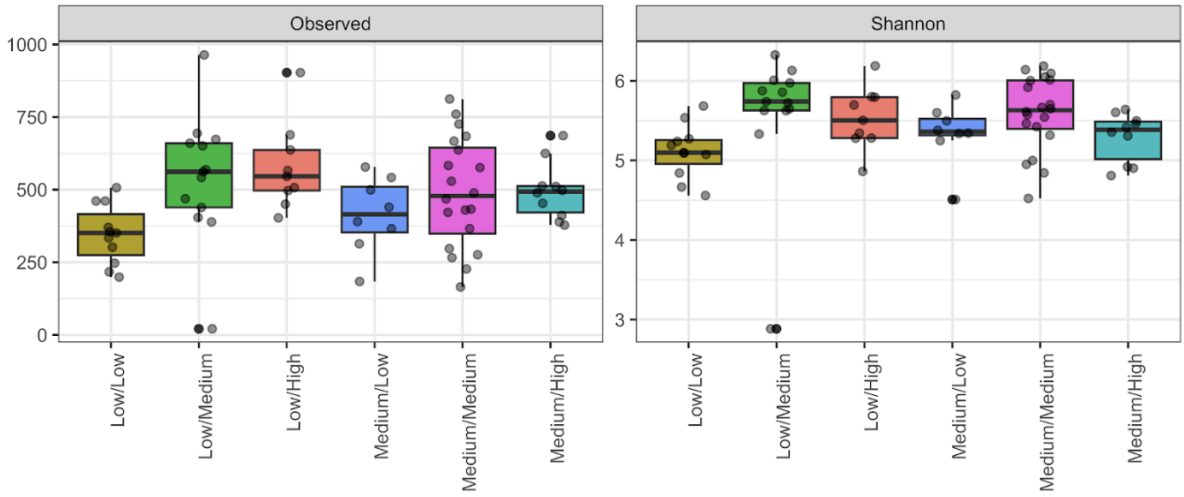

Figure 2: Alpha Diversity. Alpha diversity of transplantation candidates organized by pollution levels of original location and transplantation location. All locations had similar abundances except for Low/Low.

Observed and Shannon diversity indices were calculated using R. Low/Low was the only transplantation status that had a different diversity than the rest of the transplants. It had a lower diversity, possibly indicating that exposure to pollution may increase the diversity of the microbiome.

Table 1: ANOVA Results. ANOVA analysis of alpha diversity for both observed richness and Shannon diversity.

Richness					
	Df	Sum Sq	Mean Sq	F Value	Pr (>F)
Transplantation	5	371874	74375	2.778	0.0246
Residuals	65	1739968	26769		
Shannon					
	Df	Sum Seq	Mean Sq	F Value	Pr (>F)
Transplantation	5	2.264	0.4528	1.72	0.143

Residuals	65	17.112	0.2633		

However, an ANOVA was calculated to validate the results of the alpha diversity indices, which confirmed that only the observed diversity was significant. This indicates that though there were more species present in the samples exposed to pollution, they were at low abundance.

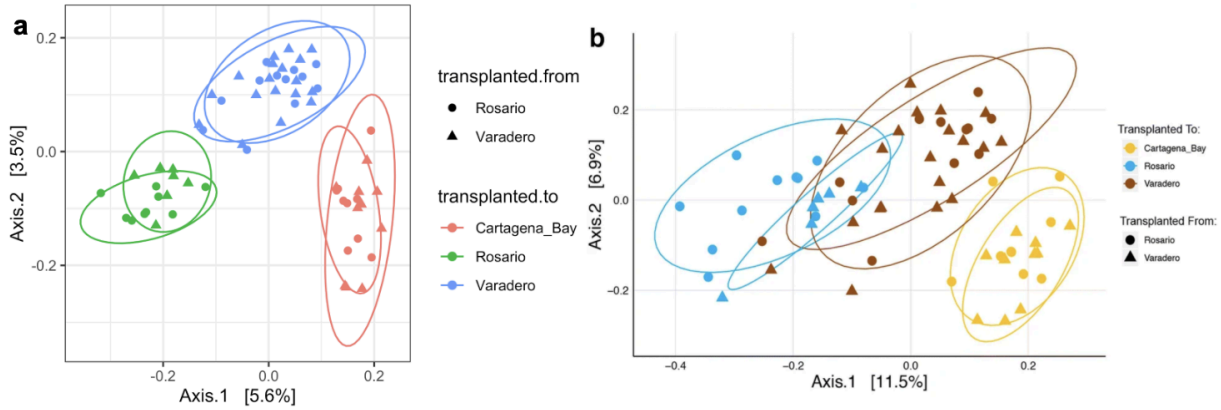

Figure 3: PCoA Analysis. Principal component analysis using bray-curtis matrices. **A. Reanalyzed PCoA.** Samples clustered by transplantation location regardless of origin there is no overlap between groups. **B. Original PCoA.** Samples clustered by transplantation location regardless of origin there is significant overlap between groups.

The reanalyzed sample showed larger differences between groups than the original PCoA analysis. However, like the original analysis, grouping is based on the transplantation site.

Table 2: PERMANOVA Results. PERMANOVA analysis of Bray-Curtis distances between original location (transplanted from) and the transplant location (transplanted to)

	Df	Sum of Sqs	R^2	F	Pr (>F)
Transplanted To	2	0.3655	0.08211	2.7860	0.001
Transplanted From	1	0.0742	0.01666	1.1308	0.080
Transplanted To/From	2	0.1417	0.03183	1.0799	0.092
Residual	59	3.8701	0.86940		

Total	64	4.4514	1.00000		

A PERMANOVA was run to validate the results of the PCoA. From the PERMANOVA of the reanalysis samples, it was found that there is more variation between sites than within a single site. In addition, it was found that 24% of the variation between samples was explained by transplantation location, however, the original location has less of an effect on microbial composition than the transplant location. In the original paper, it was found that there is a convergence of similar microbiomes regardless of the original site and transplant site (Roitman et al. 2020). Because the reanalysis data was completed using ASVs, it is likely that the difference in these conclusions is because the identification of more species of bacteria is in different families with the newer pipeline.

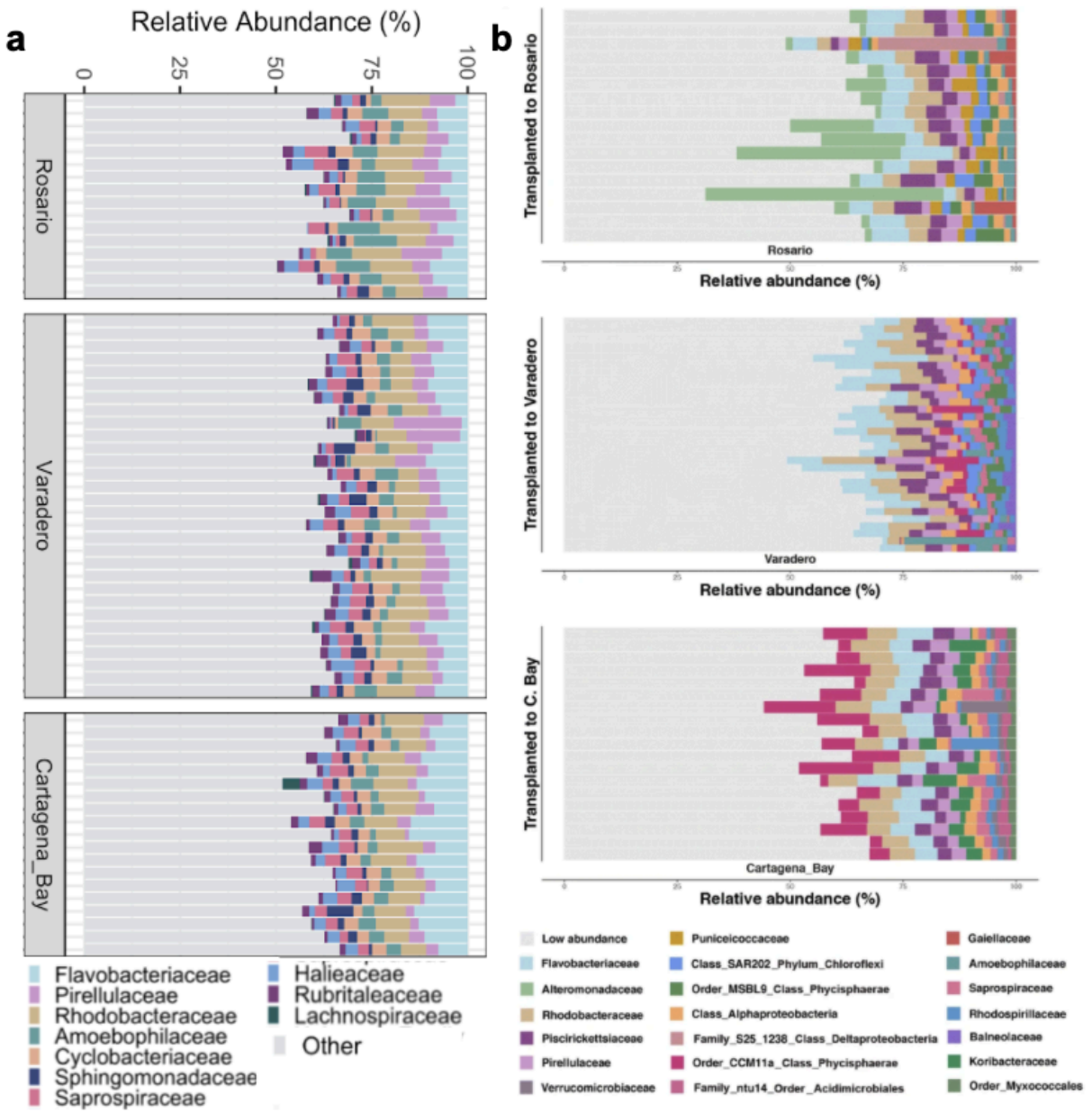

Figure 4: Relative Abundance by Family. Family level bar plots from different transplant locations. **A. Reanalyzed Relative Abundance.** Most abundant ASVs (>3.75% of community) in post-transplant fragments. Fragments that were unidentified or were less than 3.75% of community are identified as other. **B. Original Relative Abundance.** 12 most abundant OTUs in post-transplant fragments. Low abundance fragments were identified as low abundance.

Reanalyzed samples are very dissimilar to fragments identified in the original paper. The only similar families are *Flavobacteriaceae, Pirellulaceae, Rhodobacteraceae, Amoebophilaceae,* and *Saprospiraceae*. In addition, compared to the paper's findings, the only consistent finding was

that the family *Pirellulacea* was more abundant in Rosario and Veradero locations. In the paper, Rosario fragments were found to have higher abundances of OTUs belonging to the *Alteromonadaceae* family (Roitman et al. 2020), which was not detected in the reanalysis. They also found that Varadero and Cartagena Bay fragments had higher abundances of OTUs belonging to the *Saprospiraceae* family and the class *Phycisphaerae* (Roitman et al. 2020). The former was found to be evenly distributed between samples, and the latter was not detected. Trends found in the reanalysis include the Varadero having a low abundance of *Amoebophilaceae* and the very large presence of *Lachnospiraceae* in one Cartagena Bay sample.

Discussion

This research outlines the stark differences between methods of sequence analysis (ASVs or OTUs) and bioinformatic pipelines. In the original analysis of these sequences, OTUs were targeted using QIIME; v1.9 (Caporaso et al. 2011) and analyzed using the USEARCH and UCLUST packages (Edgar 2010). These software and packages were released in 2010, while QIIME has had frequent updates. USEARCH and UCLUST packages have not been updated since 2018. In addition, the UCLUST package was last updated in 2018, and even then, its identification of nucleotides was only 75% accurate (Edgar 2017). The taxonomic assignment database was located within the USEARCH package and therefore was not updated since 2018 as well (Edgar 2017). In contrast, the reanalysis of this sequence data occurred using a more recently updated version of QIIME (2024) and current packages, *Cutadapt,* DADA2, and taxonomic classification was done using SILVA (Quast et al. 2013). SLIVA was most recently updated in July 2024 and contains over 9,400,000 reference sequences to align ASVs to.

It has been proven that ASVs are more sensitive for the identification of bacterial strains (Chiarello et al. 2022). In addition, OTUs are more prone to overestimate the richness of bacteria present in samples (Chiarello et al. 2022). While there are certainly major differences between using OTUs and ASVs in this study, the largest factor is that the databases used for taxonomic classification were completely different, and one was not updated since 2018. With SILVA last being updated in 2024, the classification was most likely more representative of the actual sample than the original paper. SILVA would also contain the renamed phyla in the prokaryotic phylogenetic tree (Robitzski 2022). In addition, the use of ASVs over OTUs could have increased the amount of low-abundance bacteria found; however, since there was such a large difference in the relative abundances between the reanalysed sequences and the original analysis, the more likely cause for the difference is a more recently updated taxonomic database.

This research shows the broad changes in sequence identification techniques and the fast-paced rate of discovery of novel sequences. In addition, it highlights the importance of tailoring the bioinformatic pipelines to the unique goals of the study and periodically reanalyzing sequences using the latest technology to create more accurate conclusions about the sequence data collected. The reanalysis of data collected and analyzed years ago can be essential to draw accurate conclusions on the coral microbiome and how it has changed over time. This data can be used to determine how increased anthropogenic interactions have affected corals and potentially provide insight on what a healthy coral microbiome should look like. Though there has not been many papers concerning the reanalysis of specifically coral microbiome, there has been reanalysis of 16S sequence data to draw conclusions on historical data concerning apples, rice, and many applications in the medical field (Jang et al. 2020, Nicola et al. 2018, &

Frankel-Bricker and Frankel 2022). Using current methods to identify data collected in the past is essential to make predictions about the future.

Data Availability

All sequence data can be found in the NCBI Sequence Read Archive under BioProject ID: PRNJA639618. In addition, all code for reproducing the sequence analysis pipeline and statistical analyses can be found on GitHub: https://github.com/egt7347/EnvironmentalGenomicsFinalProject/tree/master.

Works Cited

Browne, N. K., S. G. Smithers, and C. T. Perry. 2013. Spatial and temporal variations in turbidity on two inshore turbid reefs on the Great Barrier Reef, Australia. Coral Reefs 32:195–210.

Callahan, B. J., P. J. McMurdie, M. J. Rosen, A. W. Han, A. J. A. Johnson, and S. P. Holmes. 2016. DADA2: High-resolution sample inference from Illumina amplicon data. Nature Methods 13:581–583.

Chiarello, M., M. McCauley, S. Villéger, and C. R. Jackson. 2022. Ranking the biases: The choice of OTUs vs. ASVs in 16S rRNA amplicon data analysis has stronger effects on diversity measures than rarefaction and OTU identity threshold. PLoS ONE 17:e0264443.

Edgar, R. C. 2010. Search and clustering orders of magnitude faster than BLAST. Bioinformatics 26:2460–2461

Edgar, R. C. 2017, September 21. Updating the 97% identity threshold for 16S ribosomal RNA OTUs.

Fisher, R., P. Bessell-Browne, and R. Jones. 2019. Synergistic and antagonistic impacts of suspended sediments and thermal stress on corals. Nature Communications 10:2346.

Frankel-Bricker, J., and L. K. Frankel. 2022. Re-Analysis of 16S rRNA Gene Sequence Data Sets Uncovers Disparate Laboratory-Specific Microbiomes Associated with the Yellow Fever Mosquito (Aedes aegypti). Microbial Ecology 83:167–181.

Gardner, T. A., I. M. Côté, J. A. Gill, A. Grant, and A. R. Watkinson. 2003. Long-Term Region-Wide Declines in Caribbean Corals. Science 301:958–960.

Jang, S.-W., M.-H. Yoou, W.-J. Hong, Y.-J. Kim, E.-J. Lee, and K.-H. Jung. 2020. Re-Analysis of 16S Amplicon Sequencing Data Reveals Soil Microbial Population Shifts in Rice Fields under Drought Condition. Rice 13:44.

Lin, H., and S. D. Peddada. 2020. Analysis of compositions of microbiomes with bias correction. Nature Communications 11:3514.

Nicola, L., H. Insam, I. Pertot, and B. Stres. 2018. Reanalysis of microbiomes in soils affected by apple replant disease (ARD): Old foes and novel suspects lead to the proposal of extended model of disease development. Applied Soil Ecology 129:24–33.

Oksanen, J., G. L. Simpson, F. G. Blanchet, R. Kindt, P. Legendre, P. R. Minchin, R. B. O'Hara, P. Solymos, M. H. H. Stevens, E. Szoecs, H. Wagner, M. Barbour, M. Bedward, B. Bolker, D. Borcard, G. Carvalho, M. Chirico, M. De Caceres, S. Durand, H. B. A. Evangelista, R. FitzJohn, M. Friendly, B. Furneaux, G. Hannigan, M. O. Hill, L. Lahti, D. McGlinn, M.-H. Ouellette, E. Ribeiro Cunha, T. Smith, A. Stier, C. J. F. Ter

Braak, J. Weedon, and T. Borman. 2001, September 6. vegan: Community Ecology Package.

Peixoto, R. S., P. M. Rosado, D. C. de A. Leite, A. S. Rosado, and D. G. Bourne. 2017. Beneficial Microorganisms for Corals (BMC): Proposed Mechanisms for Coral Health and Resilience. Frontiers in Microbiology 8.

Quast, C., E. Pruesse, P. Yilmaz, J. Gerken, T. Schweer, P. Yarza, J. Peplies, and F. O. Glöckner. 2013. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. Nucleic Acids Research 41:D590–D596.

Robitzski, D. 2022, January 4. Newly Renamed Prokaryote Phyla Cause Uproar. https://www.the-scientist.com/newly-renamed-prokaryote-phyla-cause-uproar-69578.

Roitman, S., T. López-Londoño, F. Joseph Pollock, K. B. Ritchie, C. T. Galindo-Martínez, K. Gómez-Campo, L. A. González-Guerrero, V. Pizarro, M. López-Victoria, R. Iglesias-Prieto, and M. Medina. 2020. Surviving marginalized reefs: assessing the implications of the microbiome on coral physiology and survivorship. Coral Reefs 39:795–807.

US Department of Commerce, N. O. and A. A. (n.d.). How do corals reproduce? https://oceanservice.noaa.gov/education/tutorial_corals/coral07_importance.html.

US EPA, O. 2017, January 30. Threats to Coral Reefs. Overviews and Factsheets. https://www.epa.gov/coral-reefs/threats-coral-reefs.