REPORT

# Reducing host DNA contamination in 16S rRNA gene surveys of anthozoan microbiomes using PNA clamps

Alicia M. Reigel[1] · Sarah M. Owens[2] · Michael E. Hellberg[1]

**Abstract** Efforts to study the microbial communities associated with corals can be limited by inefficiencies in the sequencing process due to high levels of host amplification by universal bacterial 16S rRNA gene primers. Here, we develop an inexpensive peptide nucleic acid (PNA) clamp that binds to a target sequence of host DNA during PCR and blocks amplification. We then test the ability of this PNA clamp to mitigate host contamination and increase overall microbial sequence coverage on samples from three coral species: the gorgonians *Eunicea flexuosa* and *Gorgonia ventalina,* and the scleractinian *Porites panamensis*. The 20-bp PNA clamp was designed using DNA from *E. flexuosa*. Adding the PNA clamp during PCR increased the percentage of microbial reads in *E. flexuosa* samples more than 11-fold. Microbial community diversity was similar without- and with-PNA clamps, as were the relative frequencies of the ten most abundant ASVs (amplicon sequence variants), indicating that the clamps successfully blocked host DNA amplification while simultaneously increasing microbial DNA amplification proportionally across the most abundant taxa. The reduction of *E. flexuosa* DNA correlated with an increase in the abundance of rarer taxa. The clamp also increased the average percentage of microbial reads in another gorgonian, *G. ventalina,* by 8.6-fold without altering the microbial community beta diversity, and in a distantly related scleractinian coral, *P. panamensis,* by nearly double. The reduction of host contamination correlated with the number of nucleotide mismatches between the host amplicon and the PNA clamp. The PNA clamp costs as little as $0.48 per sample, making it an efficient and cost-effective solution to increase microbial sequence coverage for high-throughput sequencing of coral microbial communities.

## Introduction

Microbial diversity within a coral host can exceed thousands of species, with individual colonies hosting as many as 650 unique OTUs (operational taxonomic units; Bayer et al. 2013; Hernandez-Agreda et al. 2018). Coral- associated microbial symbionts likely play key roles in host biology, such as organic molecule cycling (Wegley et al. 2007; Kimes et al. 2010; Bourne et al. 2016) and production of antibiotic compounds that may prevent pathogenic infections (Castillo et al. 2001; Ritchie 2006). Despite the proposed importance of microorganisms in corals and the recent advances in high-throughput sequencing technology that allow for deep sequencing of bacteria and archaeal amplicons, our knowledge of the composition and function of these microbial communities remains limited.

Topic Editor Carly Kenkel

✉ Alicia M. Reigel
   reige012@gmail.com

1  Department of Biological Sciences, Louisiana State University, 202 Life Sciences Bldg, Baton Rouge, LA 70803, USA

2  Biosciences Division, Argonne National Laboratory, 9700 S. Cass Avenue, Lemont, IL 60439, USA

These limitations stem largely from difficulties in extracting and amplifying microbial DNA from coral tissue. The extraction of coral microbial DNA is hindered by the presence of the mesoglea, a gelatinous layer of collagen fibers found in coral polyps that can decrease the effectiveness of lysis steps (Weber et al. 2017). Traditional microbial cell lysis methods can lyse coral cells, swamping samples with eukaryotic DNA (Galkiewicz and Kellogg 2008; Weber et al. 2017). Additionally, corals cells are also rife with heavy pigmentation and nucleases, both of which are PCR inhibitors (ten Lohuis et al. 1990; Price and Linge 1999). These issues are problematic for coral microbiome studies because coral DNA is easily amplified by the bacterial specific primers employed by large-scale microbial sequencing projects (e.g., Earth Microbiome Project; Galkiewicz and Kellogg 2008; Weber et al. 2017).

Host-associated microbial communities are commonly identified by sequencing short ($< 300$ bp), variable regions of the 16S rRNA gene (Case et al. 2007). Caporaso et al. (2012) developed a standardized sequencing method that utilizes universal primers (515F-806R) to amplify the 16S rRNA V4 region of DNA in bacteria and archaea. This method has been successfully implemented in host–microbiome studies of vertebrates (David et al. 2014; Loudon et al. 2014; Song et al. 2020), invertebrates (Easson and Thacker 2014; Powell et al. 2014; Pollock et al. 2018), and plants (Marzinelli et al. 2015; Shi et al. 2016). However, shared ancestry between bacteria and eukaryotic host organelles, both mitochondria and chloroplast, has led to instances of high sequence similarity between the primer-annealing sites of the targeted bacterial 16S rRNA gene and homologous sequence regions in the organelles that can cause these primers to amplify the host DNA (Sagan 1967; Allen 2015; Fitzpatrick et al. 2018).

Non-specific amplification of eukaryotic host DNA can result in massive data collecting inefficiencies, with the majority of amplicons often derived from the host rather than from targeted bacteria and archaea (Lundberg et al. 2012; Sakai and Ikenaga 2013; Fitzpatrick et al. 2018). Lundberg et al. (2013) found that when sequencing plant-associated microbial communities, host plastid and mitochondrial sequences can make up $> 90\%$ of sequence reads in a sample. Sequencing inefficiencies, as a result of non-specific PCR amplification, of this degree limit the number of samples that can be used to maintain downstream microbial sequence read depth and can make next-generation sequencing technologies cost prohibitive for large-scale projects. Further, higher microbial read coverage allows a more complete view of the microbial community by increasing diversity, often by increasing the potential for rare taxa to be sequenced (Lemos et al. 2011).

As the use of high-throughput sequencing to assess coral-associated microbes has increased, researchers have noted similar sequencing inefficiencies as in plants and other eukaryotes due to host contamination (J. Price personal comm.; Bayer et al. 2013; Pollock et al. 2018). The diversity of coral microbiomes is comparable to that of the human gut microbiome, which houses $> 1000$ OTUs per individual (Claesson et al. 2009; Lozupone et al. 2012). To characterize such highly complex microbiomes using 16S rRNA genes at the 97% OTU level may require 50,000 bacterial reads per sample (Jovel et al. 2016). Obtaining this many reads is feasible with current technologies, but high (80–90%) rates of loss due to host contamination (Fig. 1; Fitzpatrick et al. 2018) may logistically limit the number of samples that can be run while maintaining the required read counts per sample. This can become a problem for ecological studies of coral microbiomes that aim to compare treatments, host species, or environments.

Current methods to combat coral DNA contamination include the use of microbial extraction kits and PCR protocols that have been shown to increase microbial DNA yields (Al-Soud and Rådström 1998; Galkiewicz and Kellogg 2008; Sunagawa et al. 2010; Weber et al. 2017), but most PCR products still contain high concentrations of coral DNA. Post-PCR techniques to remove host DNA include agarose-gel size separation of post-PCR amplicon bands (McDevitt-Irwin et al. 2019), sequencing of less standard 16S rRNA regions (Bayer et al. 2013), and the use of longer reads on less efficient sequencing technology (i.e., 454 or Sanger; Galkiewicz and Kellogg 2008; Ainsworth et al. 2015; Jensen et al. 2019). However, these approaches can be inefficient and costly. Accordingly, developing an effective and low-cost method for decreasing the amplification of coral DNA without introducing bias would benefit continued efforts to characterize coral-associated microbial communities.
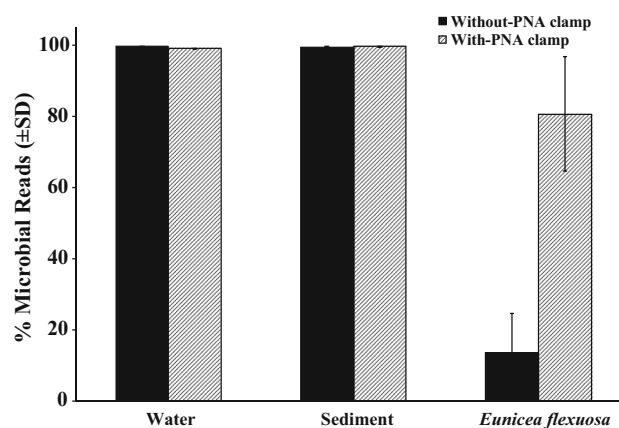


**Fig. 1** Average percent ($\pm$ SD) of all sequence reads classified as microbial DNA for water, sediment, and *Eunicea flexuosa* samples without- (black bars) and with- (diagonal line bars) PNA clamps. Error bars for sediment and water samples are too small to be visible

Peptide nucleic acid oligos, or PNA clamps, provide such a method for reducing host contamination. PNA clamps are DNA mimics made up of short nucleotide sequences with a pseudopeptide backbone that have a strong affinity for double-stranded DNA and form highly stable PNA–DNA bonds (Nielsen et al. 1991; Nielsen 1999). When added to a PCR, PNA clamps selectively bind to target DNA and block its amplification (Hyrup and Nielsen 1996). In host–microbiome studies, this selective blocking of host DNA by the PNA clamp results in amplicons composed largely of bacterial and archaeal sequences (Lundberg et al. 2013). PNA clamps have been used successfully to reduce host DNA contamination in microbial studies of plants, arthropods, and diatoms (Vestheim and Jarman 2008; Lundberg et al. 2013; Sakai and Ikenaga 2013; Ikenaga and Sakai 2014; Belda et al. 2017; Fitzpatrick et al. 2018). For example, adding a PNA clamp reduced host contamination from 23–65% across 32 plant species (Fitzpatrick et al. 2018). PNA clamps also successfully reduced 18S rRNA contamination from mosquito host DNA by > 80% (Belda et al. 2017) and by up to 100% in krill hosts (Vestheim and Jarman 2008).

Here, we design a species-specific PNA clamp to decrease coral host contamination and tested its effectiveness in a high-throughput microbial (henceforth used to imply only non-eukaryotic microbes) sequencing project using tissue samples from the coral *Eunicea flexuosa* (Subclass Octocorallia, Family Plexauridae) and non-coral controls (i.e., water and sediment). We also tested the efficacy of the PNA clamps on two other anthozoans: another gorgonian (*Gorgonia ventalina*, Subclass Octocorallia, Family Gorgoniidae) and a scleractinian (*Porites panamensis*, Subclass Hexacorallia, Family Poritidae). The addition of the PNA clamp significantly decreased host contamination in both gorgonians and did not bias overall community composition. The PNA clamp also decreased host contamination in *P. panamensis*; however, the results were more limited, likely due to several mismatches between the host DNA and the PNA clamp.

## Methods

### Sample collection

To test the applicability of the coral PNA clamp across coral species, we sampled a total of 46 coral colonies from three species. The 32 soft coral samples, 16 each of *E. flexuosa* and *G. ventalina*, were collected in August 2016 from four shallow reef sites in southwest Puerto Rico (Atravesado Reef: 17° 55′ 48.00″ N, 67° 5′ 18.00″ W; Terremoto Reef: 17° 55′ 44.10″ N, 66° 58′ 27.50″ W; Cayo Ron: 18° 5′ 14.40″ N, 67° 17′ 10.44″ W; and Enrique: 17°

57′ 15.36″ N, 67° 2′ 37.86″ W). *E. flexuosa* contains two genetically isolated ecotypes (Prada and Hellberg 2013); all *E. flexuosa* corals sampled here were the shallow ecotype, as confirmed by sequencing of the diagnostic *MSH* gene (France and Hoover 2001; Prada and Hellberg 2013). Tissue samples of *E. flexuosa* and *G. ventalina* were clipped using scissors from the axial branch tips of coral colonies located at depths of 3–5 m via SCUBA. Each branch tip was placed in a labeled bag at depth, and upon return to the surface rinsed with DI water, wiped of mucus, and placed in a sterile Whirl–Pak© bag. The sample was then immediately flash frozen at − 80 °C using a precharged Bio-Bottle© Ultra-Freeze MAX and placed in a − 80 °C freezer for long-term storage.

Four samples of reef sediment and water were also collected at each Puerto Rican location to serve as noncoral controls. Each water and sediment sample was collected at the same depth as the sampled corals. Sediment was collected from the top layer of a bare reef patch with a small shovel and put into a labeled bag. At the surface, excess water was drained, and the remaining sediment was placed into a sterile Whirl–Pak© bag, flash frozen, and stored at − 80 °C. Water samples were collected in a bleach-rinsed and 100% ethanol-sterilized 1.5 L Nalgene© bottles at depth. Upon return to the surface, 1-L was filtered through a 2 μm glass microfiber prefilter and then through a 0.22 μm Sterivex filter to collect microbes. The glass microfiber and Sterivex filters were flash frozen and stored at − 80 °C.

All 14 samples of *P. panamensis* were collected in February 2018 from Bahía Concepciòn, Baja California Sur, Mexico (26° 41′ 14.00″ N, 111° 51′ 35.30″ W). The samples of *P. panamensis* were collected from depths of ∼ 4 m using a hammer and chisel via SCUBA and placed in a labeled bag at depth. These samples were placed in 95% EtOH at the surface and stored at − 20 °C.

### DNA extraction

For soft coral samples (*E. flexuosa* and *G. ventalina*), tissue was removed from the skeleton using sterilized razors under a bio-hood (Baker Co., Stanford, ME) to minimize contamination. Hard coral samples (*P. panamensis*) were first drained of excess ethanol, rinsed 3 × with sterile $H_2O$, and then ground up using a sterilized polypropylene pellet pestle. Microbial DNA from all samples was extracted using the Qiagen DNeasy PowerSoil Kit. Sediment and water sample extractions were completed following standard manufacturer protocols. Two kit control samples were also generated by following manufacturer protocols without the addition of any PCR product. Microbial DNA was extracted from coral tissue samples using the modified protocol of Sunagawa et al. (2010). The initial heated

incubation step suggested in Sunagawa et al. (2010) was extended from 60 min to ~ 2–4 h to increase DNA yields for coral samples. For hard coral samples, we added an additional centrifugation step following lysis to concentrate skeletal fragments and then transferred the supernatant to a new tube for extraction.

All resulting DNA was quantified using a Qubit 2.0 Fluorometer (Life Technologies, Carlsbad, California) and stored at − 20 °C. DNA yields ranged from 0.08–1.73 ng/µl for water, 2.2–19.8 ng/µl for sediment, 0.4–60 ng/µl for *E. flexuosa*, 0.66–96.2 ng/µl for *G. ventalina*, 0.1–7.3 ng/µl for *P. panamensis*, and all kit control samples had DNA yields too low to quantify with the Qubit 2.0. Samples with yields lower than 5 ng/µl were extracted a second time sometimes with improvement, but it can be difficult to extract DNA from coral tissues due to the presence of the mesoglea (Miller and Howard 2004; Weber et al. 2017), which may also explain the variable DNA yields among coral samples. High-yield DNA samples were diluted to 2 ng/µl prior to PCR.

## PNA clamp design and sequencing

Initial sequencing was completed on samples from water (*n* = 4), sediment (*n* = 4), and *E. flexuosa* (*n* = 16); the latter showed high read counts of host DNA (Suppl. Table 1). These results led us to try PNA clamps to mitigate host DNA amplification. We tested three PNA clamp variations (see below). The most successful clamp for *E. flexuosa* DNA mitigation was then applied to *G. ventalina* and *P. panamensis* samples to test its efficacy across coral species.

We first tested the Lundberg et al. (2013) PNA clamp (5′GGCTCAACCCTGGACAG-3′), despite its low similarity with the contaminating *E. flexuosa* sequence, because it reduced host contamination across 14 plant families (Fitzpatrick et al. 2018), was readily available, and had been previously used with success at the Environmental Sample Preparation and Sequencing Facility (ESPSF) at Argonne National Laboratory where our samples were sent for sequencing. However, there was little to no change in host DNA sequence counts (results not shown), so we designed and tested two host-specific PNA clamps based on the contaminating *E. flexuosa* amplicon (likely part of the mitochondrial 12S rRNA gene because of its homology to the 16S rRNA bacterial sequence; Suppl. Table 2) recovered during our initial sequencing.

The process for developing a species-specific PNA clamp requires only the contaminating host sequence, which can often be obtained via public databases, as a by-product of transcriptomic assays, or cloning and sequencing amplicons generated by using the desired microbial primers. After identifying the contaminating sequence, the

PNA clamp can be designed to bind to any portion of the sequence between the primer binding sites as long as it meets the criteria for length (< 30mer) and base content (< 50% purines, no stretches of > 6 purines in a row, and < 30% G content) and has a melting temperature greater than that of double-stranded DNA (pnabio.com).

The two new coral PNA clamps we tested were designed using the PNA Tool (www.pnabio.com) to ensure they met the above criteria. The first (5′-TTAACGCCTAATGCG-3′) was a 15-bp reverse complement to the *E. flexuosa* contaminating sequence (Suppl. Table 2) that had a melting temperature of 72 °C, but it did not block amplification of the coral DNA effectively (results not reported). The second (5′-GACTCCTTACTCCGTTCATG-3′) was a 20-bp exact match to the contaminating host sequence with a melting temperature of 74.2 °C and a lower percentage of purines and G content.

All amplicon libraries targeting the 16S rRNA V4 region (515F-806R) were generated at ESPSF using a barcoded primer set adapted for the Illumina HiSeq 2000 and MiSeq (Caporaso et al. 2012). The forward amplification primer contained a 12-bp barcode sequence. PCR protocols varied without- and with-PNA clamps (Suppl. Table 3). Following PCR, amplicons were quantified using PicoGreen (Invitrogen) and a plate reader (Infinite 200 PRO, Tecan). Once quantified, products were pooled so that each amplicon was represented in equimolar amounts. This pool was then cleaned using AMPure XP Beads (Beckman Coulter) and quantified using a fluorometer (Qubit, Invitrogen). After quantification, the molarity of the pool was determined and diluted down to 2 nM, denatured, and then diluted to a final concentration of 6.75 pM with a 10% PhiX spike for sequencing.

All sequence data were generated at ESPSF using either 2 × 251 bp (*E. flexuosa* samples) or 2 × 151 bp (*G. ventalina* samples) Illumina MiSeq or 2 × 251 Illumina HiSeq 2000 (*P. panamensis* samples) paired-end sequencing. Our original intention was not to test the PNA clamp across multiple coral species, but as the opportunities arose, we determined it would be beneficial to know whether the clamps were useful for other coral hosts; therefore, the three coral species were sequenced on separate runs. All 16S rRNA gene fastq files associated with this project are accessioned at the NCBI Sequence Read Archive under BioProject #: PRJNA643267.

## Sequence analysis and sample quality control

Microbial reads were processed using QIIME2 (Bolyen et al. 2019) and a modified version of the 'Atacama soil microbiome tutorial' bioinformatics pipeline at qiime2.org. Briefly, raw fastq-formatted files were imported into QIIME2 and demultiplexed according to the pipeline. We

then used the QIIME2 wrapper of the 'DADA2' package (Callahan et al. 2016) 'denoise-paired' method to remove low quality (expected error rate > 2.0) and chimeric reads (method: 'consensus') and to truncate reads when necessary. The 3′ end of each read was truncated when the quality score consistently dropped below 20, as determined by assessing the interactive quality control plot (output during the demultiplexing step). Forward and reverse reads from *E. flexuosa* samples were truncated to 240 bp. *G. ventalina* sample reads were truncated to 150 bp. *Porites panamensis* forward reads were truncated to 240 bp, while the reverse reads were truncated to 215 bp.

DADA2 was also used to infer amplicon sequence variants (ASVs). Because the without- and with-PNA clamp samples for each species were sequenced on separate Illumina runs, the sequences were denoised independently with DADA2. The resulting ASV tables from each run and each species were later merged, following the QIIME2 protocol for merging ASV tables, thereby ensuring that identical sequences from the independent runs were combined into a single ASV identity. These merged ASV tables were then used for all downstream analyses.

Taxonomy was assigned to each ASV using the Silva rRNA v.132 reference database (Quast et al. 2012) and was done independently for the merged ASV table of each coral species. Following taxonomic identification, ASVs that did not have at least a phylum level assignment or were identified as mitochondrial, eukaryotic, or chloroplast were filtered from the ASV table. The resulting ASV table and accompanying taxonomic information were exported from QIIME2 for use in the R package 'phyloseq' (McMurdie and Holmes 2013).

## Community diversity analyses

The data were rarefied due to large differences (> 50×) in sequencing depth among our samples (Suppl. Table 1), which can impact diversity analyses (Weiss et al. 2017). We used the rarefying method in QIIME2, which pares each sample down to the same total number of reads (Weiss et al. 2017) by first removing samples below the desired read count cutoff and then resampling the remaining reads without replacement. For *E. flexuosa* samples, we examined rarefaction curves produced by the *rarecurve* command in the R package 'vegan' (Oksanen et al. 2013). Our objective was to compare the without- and with-PNA clamp samples, and because the *E. flexuosa* without-PNA clamp samples had so few microbial reads, we chose to rarefy to 800 reads per sample. This value captured all of the diversity for ¾ of the without-PNA clamp samples and minimized sample loss by discarding only 3 without-PNA clamp samples. We repeated the same process for *G. ventalina* samples but rarefied to 5000 reads.

Note that *G. ventalina* samples were sequenced separately from *E. flexuosa* on a sequence run with fewer total samples, which may account for their overall higher read counts. We did not rarefy *P. panamensis* samples because beta diversity was not assessed.

All ASV diversity analyses were completed using the R package 'PhyloSeq' (McMurdie and Holmes 2013) and following a modified R script reported in Henson et al. (2018). The Phyloseq command 'estimate richness' was employed to calculate six alpha diversity metrics (McMurdie and Holmes 2013). Beta diversity was examined using non-metric multidimensional scaling (NMDS) ordinations based on the Bray–Curtis dissimilarity matrix. NMDS ordination is a rank-based approach that represents pairwise dissimilarities between samples in low-dimensional space (Buttigieg and Ramette 2014). The *anosim* function from the R package 'vegan' was used to run Analysis of Similarity (ANOSIM) to assess significant differences between groups (i.e., sample types, clamp treatment; Oksanen et al. 2013). The R package 'ggplot2′ was used to plot NMDS. The R function *hclust* ('complete' method) from the vegan package (Oksanen et al. 2018) was utilized to perform hierarchical clustering to compare microbial community composition across sample types and treatments. Rank abundance curves were created and plotted using the R package 'ampvis2′ (Andersen et al. 2018). The 'Phyloseq' package was also utilized to calculate and plot relative abundances of bacterial and archaeal taxa at both the phyla and class level for all rarefied *E. flexuosa* samples (McMurdie and Holmes 2013). Nonparametric t tests comparing microbial reads for specific phyla between groups (without- and with-PNA clamp) were performed in JMP Pro®, Version 14 (SAS Institute Inc., Cary, NC, 1989–2019).

## Results

### Sequencing results and identification of host contamination

Initial sequencing (i.e., the without-PNA clamp samples) was completed on samples from *E. flexuosa*, sediment, and water only. Later sequencing runs included the with-PNA clamp samples and the additional two coral host species. *E. flexuosa*, sediment, and water without-PNA clamp samples retained 635,232 reads ($\sim$ 83%) after quality control (Table 1). More than half of those reads (392,407) were non-specific or unclassifiable taxonomically. The non-specific reads were almost exclusively from *E. flexuosa* samples and were assigned to 115 unique ASVs. We queried all 115 ASV sequences against the NCBI Blastn database. Fourteen of them, accounting for 380,793 (97%)

**Table 1** Information detailing the number of raw reads, post-quality control (QC) reads, reads attributed to microbial DNA, and the number of microbial ASVs in without- and with-PNA clamp samples and the merged datasets used for downstream analyses

| Dataset | Raw reads | Post-QC reads | Microbial reads | Microbial ASVs |
|---|---|---|---|---|
| Without-PNA clamp | | | | |
| Samples: 16 *E. flexuosa,* 4 sediment, 4 water | 767,612 | 635,232 | 252,035 | 2480 |
| With-PNA clamp | | | | |
| Samples: 16 *E. flexuosa,* 4 sediment, 4 water | 1,030,347 | 899,741 | 797,445 | 8113 |
| Merged dataset | | | | |
| (Including both without- and with-PNA clamp data sets) | | | | 8661 |
| *Gorgonia* dataset without- and with-PNA clamp | | | | |
| Samples: 16 *G. ventalina* | 2,078,558 | 1,635,219 | 307,554 | 1043 |
| *Porites* dataset without- and with-PNA clamp | | | | |
| Samples: 14 *P. panamensis* | 19,903,318 | 8,999,269 | 8,247,502 | 19,110 |

of the reads, had high sequence similarity matches (> 95%) to the mitochondrial genome sequences of two soft corals, *Muricea purpurea* and *M. crassa* (Genbank Accession #s LT174653.1 and LT174652.1, respectively; Suppl. Table 2). Both *Muricea* species are from the same family as *E. flexuosa*, suggesting that our contaminating reads were host DNA.

## PNA clamp efficiency and effect on microbial composition across sample types

The 20-bp exact match PNA clamp was added to the *E. flexuosa*, sediment and water samples. Sequence depth varied due to differences between sequencing runs and among samples in the same run (Suppl. Table 1), so we could not directly compare the number of microbial reads between without- and with-PNA clamp treatments even though they are what we sought to investigate. Instead, we assessed microbial reads as proportions of the total post-QC reads within a sample (Suppl. Table 1), which allowed for intra-sample and inter-treatment comparisons. The addition of the PNA clamps increased the percentage of microbial reads 11-fold on average for *E. flexuosa* on a per sample basis (Suppl. Figure 1). Without-PNA clamp samples averaged 13.7% microbial reads, but this increased to 80.8% with the PNA clamp (*t* test, $t_{22} = -10.02$, $p < 0.0001$; Fig. 1). Sediment and water samples averaged > 99% microbial reads (Fig. 1), regardless of whether PNA clamps were used (*t* test—sediment: $t_6 = 3.87$, $p > 0.05$; water: $t_6 = -1.5e{-}13$, $p > 0.05$).

Microbial community composition, both without- and with-PNA clamps, was strongly defined by sample type, with sediment, water, and *E. flexuosa* samples forming distinct clusters (ANOSIM $R^2$: 0.98, $p < 0.0001$, Suppl. Figure 2). Hierarchical clustering confirmed the distinction between microbial communities among sample types

(Suppl. Figure 3). Because the control samples (sediment and water) were distinct from the coral samples, all further analyses were completed using only *E. flexuosa* samples.

## PNA clamp effect on *E. flexuosa* microbial composition

The increase in relative abundance of microbial reads with the addition of the PNA clamp correlated with an increase in alpha diversity for non-rarefied *E. flexuosa* samples (Kruskal–Wallis test, Chi-squared$_1$ = 10.022, $p$ = 0.0015; Suppl. Figure 4), reflecting an increase of > 6000 ASVs (Table 1). Despite this increase, the ten most abundant ASVs (accounting for ∼ 80% of all microbial reads) were taxonomically identical in without- and with-PNA clamp samples, nor did their relative proportions change (Suppl. Figure 5). The dominant genus in the *E. flexuosa* samples was *Endozoicomonas* (Suppl. Figure 5). Three distinct ASVs, all identified as *Endozoicomonas spp.*, accounted for > 30% of the microbial reads in without- and with-PNA clamp samples (Suppl. Figure 5).

We further assessed *E. flexuosa* microbial compositions with the rarefied dataset to mitigate any differences in sequence depth between samples. Alpha diversity remained higher in the with-PNA clamp than in the without-PNA clamp samples after rarefying (Kruskal–Wallis Test, Chi-squared = 16.599, df = 1, $p < 0.0001$, Suppl. Figure 6). This higher diversity coincided with an increase in the overall abundance of rare ASVs (Suppl. Figure 7). This increase in rare ASVs did not have a discernable impact on the overall microbial composition (ANOSIM $R^2$: 0.00013, $p < 0.001$; Fig. 2a). Clustering analyses placed most microbial communities from the same *E. flexuosa* colony amplified without- and with-PNA as closest to one another (Fig. 2b), indicating that the clamp did not cause any large-scale compositional shifts within a sample.

**Fig. 2** Similarity of microbial community composition for rarefied *Eunicea flexuosa* samples without- (square) and with- (circle) PNA clamps. Samples from the same individual coral share the same color. **a** NMDS plot of beta diversity (Bray–Curtis) **b** Hierarchical clustering dendrogram of *E. flexuosa* microbial community composition
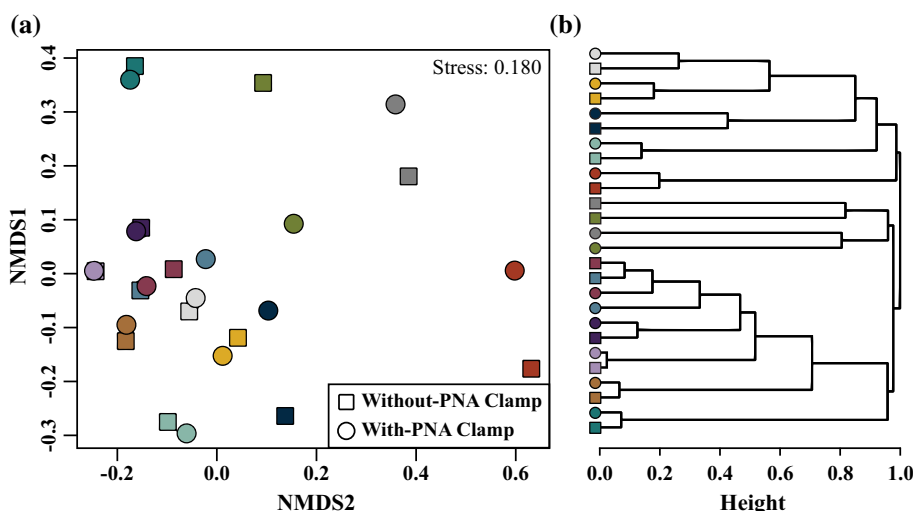


Relative abundance of microbial taxa (at phylum level) was largely unchanged within a sample with the addition of the PNA clamp (Fig. 3), with the exception of one of the rarer phyla, Bacteroidetes. Taxa in this phylum increased from an average relative abundance of 1% in without-PNA clamp samples to 5% in with-PNA clamp samples (Wilcoxon signed-ranks test: $T_S = 304$, df = 1, $p = 0.0147$; Fig. 3). Two other rare phyla (Verrucomicrobia and Firmicutes) increased slightly with the addition of the clamp (Fig. 3), but these increases were neither significant nor consistent across samples. Relative abundances at the class level mirrored those seen at the phylum level (data not shown).

## Testing PNA clamp on additional anthozoan species

The PNA clamps were based upon the *E. flexuosa* amplicon sequence. As anthozoan mitochondrial DNA generally evolves slowly (Hellberg 2006), we also tested the efficacy of the clamp on two other corals: *G. ventalina*, an octocoral which diverged from *E. flexuosa* > 70 mya (Park et al. 2012), and *P. panamensis*, a scleractinian (Suppl. Table 2). Hexacorals (to which the scleractinians belong) and octocorals diverged > 600 mya (Park et al. 2012).

For *G. ventalina*, the PNA clamp increased the average percentage of microbial reads 8.6-fold on a per sample basis (Suppl. Figure 8a). The without-PNA clamp samples averaged 11.3% microbial reads, and the addition of the

**Fig. 3** Relative abundance of microbial taxa at the phylum level in rarefied *Eunicea flexuosa* samples. The same samples from the same colony that were amplified without- (black triangles) and with- (gray circles) PNA clamps are grouped together
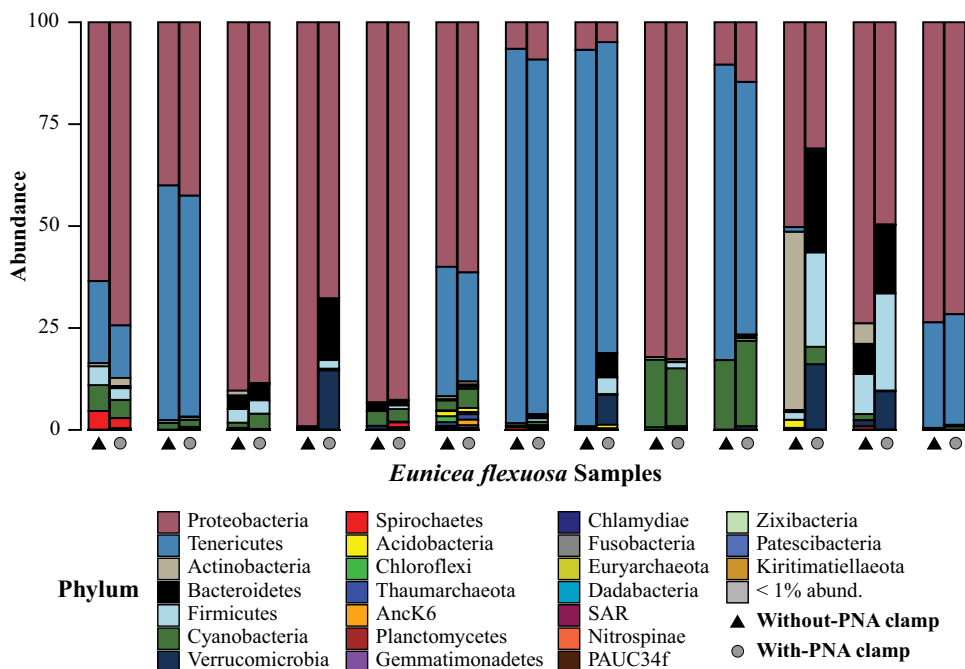
PNA clamp increased this to 68% ($t$ test, $t_{1,30}$ = − 3.29, $p$ = 0.0026; Suppl. Figure 8a). Microbial community alpha diversity of *G. ventalina* samples increased slightly with the PNA clamp (Kruskal–Wallis Test, Chi-squared$_{1}$ = 6.7727, $p$ = 0.009; Suppl. Figure 8b). Beta diversity was not altered by the PNA clamp (ANOSIM $R^2$ = − 0.0184; Suppl. Figure 8c); microbial composition instead seems to have been driven largely by sampling location (Suppl. Figure 8c). A hierarchical analysis did not show any clustering by treatment status, and most paired treatment samples were very similar to one another (Suppl. Figure 8d).

*Porites panamensis* had less host sequence contamination (∼ 14% of all reads) without clamps than did the two octocorals. Adding clamps further decreased host contamination by 50% (Suppl. Figure 9), although this was not significant.

## Discussion

Low microbial sequence coverage constrains the characterization of host-associated microbes, particularly in highly diverse microbiomes or when many taxa are already rare (Lemos et al. 2011; Zaheer et al. 2018). Extraction protocols have been developed that increased microbial DNA yields from coral samples (Galkiewicz and Kellogg 2008; Sunagawa et al. 2010; Weber et al. 2017), but they still fail to completely remove host DNA. Cross-amplification of the remaining coral DNA results in low microbial sequence recovery, making it expensive to characterize diverse coral microbiomes for large-scale sampling projects (Lundberg et al. 2012; Sakai and Ikenaga 2013; Fitzpatrick et al. 2018). Here, we have shown that adding PNA clamps is an efficient, cost-effective way to mitigate coral host contamination leading to higher microbial read counts without introducing bias.

The PNA clamp developed here was an exact match to *E. flexuosa*, but it also reduced host contamination in the two other phylogenetically distant anthozoans we tested. The effectiveness of the clamp, however, decreased as the number of mismatches with the host amplicon sequence increased from a single mismatch between the primer and *G. ventalina* to eight with *P. panamensis* (Suppl. Figure 10). Interestingly, *P. panamensis* had less initial host contamination than either gorgonian. Samples from *P. panamensis* yielded low DNA concentrations, which may suggest that *P. panamensis* cells did not lyse as readily as gorgonian cells, leading to a sample composed mainly of symbiotic microbial DNA. Lysis affinity of cells varies based on cell size, the amount of mesoglea, and the presence of inhibitors (ten Lohuis et al. 1990; Weber et al. 2017). The disparity in initial contamination may also have

been due to a higher number of mismatches between *P. panamensis* mtDNA and the microbial primers used, reducing the amplification of the coral DNA. Regardless, the PNA clamp, while less effective, still halved the remaining *P. panamensis* contamination. All considered, a single PNA clamp that limits host contamination should be possible for congeneric, confamilial, and perhaps conordinal coral species.

One of the major limitations to increasing read counts in microbiome studies is cost. PNA clamps are an inexpensive solution to increase read counts across coral species. Consider this simple example of the cost-saving benefits of utilizing PNA clamps for a coral microbial study. Suppose the aim is to compare 5 locations × 2 treatments per location × 10 biological replicates (colonies) per treatment × 2 collection time points = 200 coral samples. Presently, high-throughput sequencing utilizing the MiSeq Illumina platform costs about $2600 per lane with sample preparation and returns ∼ 13 million reads with an average read retention of 85% after quality control for a total of 11,050,000 reads per run. If 50,000 reads are needed to assess the microbiome of each sample and 80% of the post-quality control reads (as per *E. flexuosa* data, Fig. 1) are lost to host DNA contamination, this sample study would need five MiSeq runs [11,050,000 reads*5 MiSeq runs = 55,250,000 reads ÷ 200 samples = 276,250 reads/sample*0.20 reads retained = 55,250 microbial reads/sample], which would cost $13,000. In contrast, the 20-bp PNA clamp developed for this study costs $0.48 per sample and decreased host contamination from 80 to 13%. For the same mock study, adding the PNA clamps would cost $96, but required sequencing would come down to $5200 [11,050,000 reads*2 MiSeq runs = 22,100,000 reads ÷ 200 samples = 110,500 reads/sample*0.87 reads retained = 96,135 microbial reads/sample], for a total of $5296. In this scenario, using PNA clamps saves $7704.

While the previous example illustrates a best-case scenario for cost savings, we also tested the PNA clamps on the scleractinian *P. panamensis,* where the initial loss of reads to host DNA was much less (14%) than for either gorgonian. PNA clamps halved contamination levels to 7%. Using the same premises as above, only one MiSeq lane would be required, for a savings of $2504. Overall, employing PNA clamps will usually be much less expensive than additional sequencing and will allow coral researchers to maintain higher sample sizes with necessary microbial sequence coverage.

Critically, the addition of PNA clamps did not introduce any amplification or sequencing biases. Previous efforts to mitigate host contamination often include the use of less standardized microbial primers, some of which can introduce biases via selective amplification (Sipos et al. 2007; Galkiewicz and Kellogg 2008) as well as providing

sequences that are harder to compare to data repositories. Treatment with the PNA clamp did not change overall microbial composition (beta diversity) for *E. flexuosa* (Figs. 2a, b and 3) or *G. ventalina* (Suppl. Figure 8c, d). Both the taxonomic identity and percentage of the total reads for the most abundant microbial ASVs (accounting for ∼ 80% of microbial reads) remained consistent with the addition of the clamp (Suppl. Figure 5). The most abundant genus in *E. flexuosa* was *Endozoicomonas,* a well-known and abundant coral symbiont accounting for 10–60% of microbial reads in other soft corals and more than 90% in some scleractinians (Bayer et al. 2013; Pogoreutz et al. 2018; Pollock et al. 2018). *Endozoicomonas* accounted for more than a third of the microbial reads in both without- and with-PNA clamp samples (Suppl. Figure 5). The few other studies that have compared eukaryotic host-associated microbiomes without- and with-PNA clamps have likewise found no taxon-specific amplification biases, nor an effect of the clamp on overall microbial community composition (Lundberg et al. 2013; Belda et al. 2017; Fitzpatrick et al. 2018).

Alpha diversity for both gorgonians increased with addition of the PNA clamps (Suppl. Figures 4, 6 and 8b). These differences were due to an increase in abundance of rarer ASVs (Suppl. Figure 7). One phylum rare in *E. flexuosa* samples, Bacteroidetes, increased significantly with the addition of the PNA clamp (Fig. 3), while two others (Verrucomicrobia and Firmicutes) also increased, albeit not significantly. Similarly, PNA clamps used in a variety of eukaryotic host–microbiome studies have found that decreasing host amplicons increases alpha diversity by making room for rare taxa to amplify during PCR (Lundberg et al. 2013; Ikenaga and Sakai 2014; Belda et al. 2017). Many coral–microbe symbioses include rare phylotypes (Ainsworth et al. 2015), and several core coral microbiome members are from the rare biosphere (Kellogg 2019). The potential importance of such rare bacteria in coral-associated microbial communities enforces the need to reduce host contamination and allow higher amplification of rare microbial amplicons.

As explorations of coral host–symbiont relationships move beyond simple taxonomic characterization and into questions about the impacts of various environmental stressors and physiological conditions, efficient sequencing efforts that stem losses to host DNA contamination and maintain high microbial sequence coverage will be needed. The PNA clamps designed and tested here provide an efficient and unbiased solution to the problem of host DNA contamination in coral microbiome studies. The clamps effectively reduce host contamination during the PCR process and increased microbial read counts in three evolutionarily distant coral species.

**Compliance with ethical standards**

## References

Ainsworth TD, Krause L, Bridge T, Torda G, Raina JB, Zakrzewski M, Gates RD, Padilla-Gamiño JL, Spalding HL, Celia S, Woolsey ES, Bourne DG, Bongaerts P, Hoegh-Guldberg O, Leggat W (2015) The coral core microbiome identifies rare bacterial taxa as ubiquitous endosymbionts. ISME J 9:2261–2274

Allen JF (2015) Why chloroplasts and mitochondria retain their own genomes and genetic systems: Colocation for redox regulation of gene expression. Proc Natl Acad Sci USA 112:10231–10238

Al-Soud WA, Rådström P (1998) Capacity of nine thermostable DNA polymerases to mediate DNA amplification in the presence of PCR-inhibiting samples. Appl Environ Microbiol 64:3748–3753

Andersen KS, Kirkegaard RH, Karst SM, Albertsen M (2018) ampvis2: an R package to analyse and visualise 16S rRNA amplicon data. BioRxiv: 299537

Bayer T, Neave MJ, Alsheikh-Hussain A, Aranda M, Yum LK, Mincer T, Hughen K, Apprill A, Voolstra CR (2013) The microbiome of the Red Sea coral *Stylophora pistillata* is dominated by tissue-associated *Endozoicomonas* bacteria. Appl Environ Microbiol 79:4759–4762

Belda E, Coulibaly B, Fofana A, Beavogui AH, Traore SF, Gohl DM, Vernick KD, Riehle MM (2017) Preferential suppression of *Anopheles gambiae* host sequences allows detection of the mosquito eukaryotic microbiome. Sci Rep 7:3241

Bolyen E, Rideout J, Dillon M, Bokulich N, Abnet C, Al-Ghalith G, Alexander H, Alm E, Arumugam M, Asnicar F, Bai Y, Bisanz J, Bittinger K, Brejnrod A, Brislawn C, Brown C, Callahan B, Caraballo-Rodríguez A, Chase J, Cope E, Da Silva R, Diener C, Dorrestein P, Douglas G, Durall D, Duvallet C, Edwardson C,

Ernst M, Estaki M, Fouquier J, Gauglitz J, Gibbons S, Gibson D, Gonzalez A, Gorlick K, Guo J, Hillmann B, Holmes S, Holste H, Huttenhower C, Huttley G, Janssen S, Jarmusch A, Jiang L, Kaehler B, Kang K, Keefe C, Keim P, Kelley S, Knights D, Koester I, Kosciolek T, Kreps J, Langille M, Lee J, Ley R, Liu Y, Loftfield E, Lozupone C, Maher M, Marotz C, Martin B, McDonald D, McIver L, Melnik A, Metcalf J, Morgan S, Morton J, Naimey A, Navas-Molina J, Nothias L, Orchanian S, Pearson T, Peoples S, Petras D, Preuss M, Pruesse E, Rasmussen L, Rivers A, Robeson M, Rosenthal P, Segata N, Shaffer M, Shiffer A, Sinha R, Song S, Spear J, Swafford A, Thompson L, Torres P, Trinh P, Tripathi A, Turnbaugh P, Ul-Hasan S, van der Hooft J, Vargas F, Vázquez-Baeza Y, Vogtmann E, von Hippel M, Walters W, Wan Y, Wang M, Warren J, Weber K, Williamson C, Willis A, Xu Z, Zaneveld J, Zhang Y, Zhu Q, Knight R, Caporaso J (2019) Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. Nat Biotechnol 37:852–857

Bourne DG, Morrow KM, Webster NS (2016) Insights into the Coral Microbiome: Underpinning the Health and Resilience of Reef Ecosystems. Annu Rev Microbiol 70:317–340

Buttigieg PL, Ramette A (2014) A Guide to Statistical Analysis in Microbial Ecology: a community-focused, living review of multivariate data analyses. FEMS Microbiol Ecol 90:543–550

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

Caporaso JG, Lauber CL, Walters WA, Berg-Lyons D, Huntley J, Fierer N, Owens SM, Betley J, Fraser L, Bauer M, Niall G, Gilbert JA, Smith G, Knight R (2012) Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms. ISME J 6:1621–1624

Case RJ, Boucher Y, Dahllöf I, Holmström C, Doolittle WF, Kjelleberg S (2007) Use of 16S rRNA and rpoB genes as molecular markers for microbial ecology studies. Appl Environ Microbiol 73:278–288

Castillo I, Lodeiros C, Núñez M, Campos I (2001) Evaluación in vitro de sustancias antibacterianas producidas por bacterias aisladas de diferentes organismos marinos. Rev Biol Trop 49:1213–1222

Claesson MJ, O'Sullivan O, Wang Q, Nikkilä J, Marchesi JR, Smidt H, de Vos WM, Ross RP, O'Toole PW (2009) Comparative analysis of pyrosequencing and a phylogenetic microarray for exploring microbial community structures in the human distal intestine. PLoS ONE. https://doi.org/10.1371/journal.pone.0006669

David LA, Maurice CF, Carmody RN, Gootenberg DB, Button JE, Wolfe BE, Ling AV, Devlin AS, Varma Y, Fischbach MA, Biddinger SB (2014) Diet rapidly and reproducibly alters the human gut microbiome. Nature 505:559–563

Easson CG, Thacker RW (2014) Phylogenetic signal in the community structure of host-specific microbiomes of tropical marine sponges. Front Microbiol 5:532

Fitzpatrick CR, Lu-Irving P, Copeland J, Guttman DS, Wang PW, Baltrus DA, Dlugosch KM, Johnson MTJ (2018) Chloroplast sequence variation and the efficacy of peptide nucleic acids for blocking host amplification in plant microbiome studies. Microbiome 6:144–153

France SC, Hoover LL (2001) Analysis of variation in mitochondrial DNA sequences (ND3, ND4L, MSH) among Octocorallia (=Alcyonaria) (Cnidaria: Anthozoa). Bull Biol Soc Wash 10:110–118

Galkiewicz JP, Kellogg CA (2008) Cross-kingdom amplification using bacteria-specific primers: Complications for studies of coral microbial ecology. Appl Environ Microbiol 74:7828–7831

Hellberg ME (2006) No variation and low synonymous substitution rates in coral mtDNA despite high nuclear variation. BMC Evol Biol. https://doi.org/10.1186/1471-2148-6-24

Henson MW, Hanssen J, Spooner G, Fleming P, Pukonen M, Stahr F, Thrash JC (2018) Nutrient dynamics and stream order influence microbial community patterns along a 2914 kilometer transect of the Mississippi River. Limnol Oceanogr 63:1837–1855

Hernandez-Agreda A, Leggat W, Bongaerts P, Herrera C, Ainsworth TD (2018) Rethinking the coral microbiome: Simplicity exists within a diverse microbial biosphere. mBio. https://doi.org/10.1128/mBio.00812-18

Hyrup B, Nielsen PE (1996) Peptide nucleic acids (PNA): synthesis, properties and potential applications. Bioorg Med Chem 4:5–23

Ikenaga M, Sakai M (2014) Application of locked nucleic acid (LNA) oligonucleotide–PCR clamping technique to selectively PCR amplify the SSU rRNA genes of bacteria in investigating the plant-associated community structures. Microbes Environ ME14061 https://doi.org/10.1264/jsme2.ME14061]

Jensen S, Hovland M, Lynch MDJ, Bourne DG (2019) Diversity of deep-water coral-associated bacteria and comparison across depth gradients. FEMS Microbiol Ecol. https://doi.org/10.1093/femsec/fiz091

Jovel J, Patterson J, Wang W, Hotte N, O'Keefe S, Mitchel T, Perry T, Kao D, Mason AL, Madsen KL, Wong GKS (2016) Characterization of the gut microbiome using 16S or shotgun metagenomics. Front Microbiol. https://doi.org/10.3389/fmicb201600459

Kellogg CA (2019) Microbiomes of stony and soft deep-sea corals share rare core bacteria. Microbiome 7:90

Kimes NE, Van Nostrand JD, Weil E, Zhou J, Morris PJ (2010) Microbial functional structure of Montastraea faveolata, an important Caribbean reef-building coral, differs between healthy and yellow-band diseased colonies. Environ Microbiol 12:541–556

Lemos LN, Fulthorpe RR, Triplett EW, Roesch LF (2011) Rethinking microbial diversity analysis in the high throughput sequencing era. J Microbiol Methods 86:42–51

Loudon AH, Woodhams DC, Parfrey LW, Archer H, Knight R, McKenzie V, Harris RN (2014) Microbial community dynamics and effect of environmental microbial reservoirs on red-backed salamanders (Plethodon cinereus). ISME J 8:830–840

Lozupone CA, Stombaugh JI, Gordon JI, Jansson JK, Knight R (2012) Diversity, stability and resilience of the human gut microbiota. Nature 489:220–230

Lundberg DS, Lebeis SL, Paredes SH, Yourstone S, Gehring J, Malfatti S, Tremblay J, Engelbrektson A, Kunin V, Del Rio TG, Edgar RC (2012) Defining the core Arabidopsis thaliana root microbiome. Nature 488:86–90

Lundberg DS, Yourstone S, Mieczkowski P, Jones CD, Dangl JL (2013) Practical innovations for high-throughput amplicon sequencing. Nat Methods 10:999–1002

Marzinelli EM, Campbell AH, Zozaya Valdes E, Vergés A, Nielsen S, Wernberg T, De Bettignies T, Bennett S, Caporaso JG, Thomas T, Steinberg PD (2015) Continental-scale variation in seaweed host-associated bacterial communities is a function of host condition, not geography. Environ Microbiol 17:4078–4088

McDevitt-Irwin JM, Garren M, McMinds R, Vega-Thurber R, Baum JK (2019) Variable interaction outcomes of local disturbance and El Niño-induced heat stress on coral microbiome alpha and beta diversity. Coral Reefs 38:331–345

McMurdie PJ, Holmes S (2013) phyloseq: An R Package for Reproducible Interactive Analysis and Graphics of Microbiome Census Data. PLoS ONE 8:e61217

Miller KJ, Howard CG (2004) Isolation of microsatellites from two species of scleractinian coral. Mol Ecol Notes 4:11–13

Nielsen PE (1999) Peptide nucleic acid. A molecule with two identities. Acc Chem Res 32:624–630

Nielsen PE, Egholm M, Berg RH, Buchardt O (1991) Sequence-selective recognition of DNA by strand displacement with a thymine-substituted polyamide. Science 254:1497–1500

Oksanen J, Blanchet FG, Friendly M, Kindt R, Legendre P, McGlinn D, Minchin PR, O'Hara RB, Simpson G, Solymos P, Stevens MHH, Szoecs E, Wagner H (2018) vegan: Community Ecology Package. https://cranr-projectorg/package=vegan

Oksanen J, Blanchet FG, Kindt R, Legendre P, Minchin P, O'Hara RB, Simpson G, Solymos P, Stevens MHH, Wagner H (2013) Vegan: Community Ecology Package R Package Version 2.0-10

Park E, Hwang DS, Lee JS, Song JI, Seo TK, Won YJ (2012) Estimation of divergence times in cnidarian evolution based on mitochondrial protein-coding genes and the fossil record. Mol Phylogenet Evol 62:329–345

Pogoreutz C, Rädecker N, Cárdenas A, Gärdes A, Wild C, Voolstra CR (2018) Dominance of *Endozoicomonas* bacteria throughout coral bleaching and mortality suggests structural inflexibility of the *Pocillopora verrucosa* microbiome. Ecol Evol 8:2240–2252

Pollock FJ, McMinds R, Smith S, Bourne DG, Willis BL, Medina M, Thurber RV, Zaneveld JR (2018) Coral-associated bacteria demonstrate phylosymbiosis and cophylogeny Nat Commun 9:1–13

Powell JE, Martinson VG, Urban-Mead K, Moran NA (2014) Routes of acquisition of the gut microbiota of the honey bee *Apis mellifera*. Appl Environ Microbiol 80:7378–7387

Prada C, Hellberg ME (2013) Long pre-reproductive selection and divergence by depth in a Caribbean candelabrum coral. Proc Natl Acad Sci 10:3961–3966

Price K, Linge C (1999) The presence of melanin in genomic DNA isolated from pigmented cell lines interferes with successful polymerase chain reaction: a solution. Melanoma Res 9:5–9

Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Yarza P, Peplies J, Glöckner FO (2012) The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. Nucleic Acids Res 41(D1):D590–D596

Ritchie KB (2006) Regulation of microbial populations by coral surface mucus and mucus-associated bacteria. Mar Ecol Prog Ser 322:1–14

Sagan L (1967) On the origin of mitosing cells. J Theor Biol 14:225–274

Sakai M, Ikenaga M (2013) Application of peptide nucleic acid (PNA)-PCR clamping technique to investigate the community structures of rhizobacteria associated with plant roots. J Microbiol Methods 92:281–288

Shi S, Nuccio EE, Shi ZJ, He Z, Zhou J, Firestone MK (2016) The interconnected rhizosphere: high network complexity dominates rhizosphere assemblages. Ecol Lett 19:926–936

Sipos R, Székely AJ, Palatinszky M, Révész S, Márialigeti K, Nikolausz M (2007) Effect of primer mismatch, annealing temperature and PCR cycle number on 16S rRNA gene-targeting bacterial community analysis. FEMS Microbiol Ecol 60:341–350

Song SJ, Sanders JG, Delsuc F, Metcalf J, Amato K, Taylor MW, Mazel F, Lutz HL, Winker K, Graves GR, Humphrey G (2020) Comparative Analyses of Vertebrate Gut Microbiomes Reveal Convergence between Birds and Bats. Bio. https://doi.org/10.1128/mBio.02901-19

Sunagawa S, Woodley CM, Medina M (2010) Threatened corals provide underexplored microbial habitats. PLoS ONE. https://doi.org/10.1371/journal.pone.0009554

ten Lohuis M, Alderslade P, Miller DJ (1990) Isolation and cloning of DNA from somatic tissue of soft corals (Cnidaria: Octocorallia). Mar Biol 104:489–492

Vestheim H, Jarman SN (2008) Blocking primers to enhance PCR amplification of rare sequences in mixed samples – a case study on prey DNA in Antarctic krill stomachs. Front Zool. https://doi.org/10.1186/1742-9994-5-12

Weber L, DeForce E, Apprill A (2017) Optimization of DNA extraction for advancing coral microbiota investigations. Microbiome. https://doi.org/10.1186/s40168-017-0229-y

Wegley L, Edwards R, Rodriguez-Brito B, Liu H, Rohwer F (2007) Metagenomic analysis of the microbial community associated with the coral *Porites astreoides*. Environ Microbiol 9:2707–2719

Weiss S, Xu ZZ, Peddada S, Amir A, Bittinger K, Gonzalez A, Lozupone C, Zaneveld JR, Vázquez-Baeza Y, Birmingham A, Hyde ER, Knight R (2017) Normalization and microbial differential abundance strategies depend upon data characteristics. Microbiome. https://doi.org/10.1186/s40168-017-0237-y

Zaheer R, Noyes N, Ortega Polo R, Cook SR, Marinier E, Van Domselaar GV, Belk KE, Morley PS, McAllister TA (2018) Impact of sequencing depth on the characterization of the microbiome and resistome. Sci Rep. https://doi.org/10.1038/s41598-018-24280-8