microorganisms

MDPI

Article

Multiscale Patterns of Bacterial and Protist Diversity Across Red Sea Coral Reefs

Christopher A. Hempel and Larissa Frühe *

Marine Science Program, Biological and Environmental Science and Engineering Division (BESE), King Abdullah University of Science and Technology (KAUST), Thuwal 23955-6900, Saudi Arabia
* Correspondence: larissa.fruehe@kaust.edu.sa

Abstract

Coral reef microbial communities play pivotal roles in ecosystem functioning but remain understudied, particularly across spatial gradients and domains. Here, we use environmental DNA (eDNA) metabarcoding of 16S and 18S rRNA genes to profile bacterial and protistan communities in surface sediments from six coral reefs along the central Red Sea. At each reef, we sampled both exposed (seaward) and sheltered (shoreward) sites, enabling a multiscale analysis of diversity and community composition. We found significant differences in alpha and beta diversity between reefs and between exposure sites within reefs for both microbial groups. Redundancy analysis (RDA) and PERMANOVA revealed reef identity and exposure category as key structuring factors. Indicator species analysis and Random Forest classification identified microbial taxa predictive of exposure gradients, with several exact sequence variants (ESVs) serving as robust bioindicators in both methods. Bacterial and protistan communities exhibited overlapping but distinct patterns, highlighting their complementary ecological roles. Our results underscore the importance of fine-scale habitat heterogeneity in shaping reef microbial assemblages and support the integration of multi-domain eDNA data into coral reef monitoring frameworks.

Keywords: environmental DNA; coral reef microbes; marine protists; marine bacteria; supervised machine learning; indicator species; coral reef monitoring

check for updates

Academic Editor: Esti Kramarsky-Winter

Received: 16 August 2025
Revised: 10 September 2025
Accepted: 18 September 2025
Published: 7 November 2025

Citation: Hempel, C.A.; Frühe, L. Multiscale Patterns of Bacterial and Protist Diversity Across Red Sea Coral Reefs. *Microorganisms* **2025**, *13*, 2549. https://doi.org/10.3390/microorganisms13112549

Copyright: © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

Coral reefs are among the most biologically diverse and ecologically and economically valuable ecosystems on Earth [1,2]. While the visible biodiversity of corals and reef fishes has been extensively studied, these habitats also support vast and complex communities of microorganisms, including bacteria and protists, which play fundamental roles in reef productivity and function [3]. Microbial communities contribute to primary production, nutrient cycling, host symbioses, and food web dynamics, ultimately shaping the health and resilience of reef ecosystems. Despite their importance, reef-associated microbial communities—particularly microeukaryotic protists—remain underrepresented in biodiversity and monitoring studies [4].

Many studies investigating the role of microbes in coral reefs focus on the coral microbiome [5–9] or bacterioplankton assemblages [10]. Yet the highly diverse benthic bacterial communities in reef systems are especially key-players in the benthic–pelagic link [11] and provide a close-knitted interaction between the coral themselves by degrading sinking coral mucus and contributing to the small-scale nitrogen cycle of coral reefs [12,13].

Tropical reef ecosystems sustain a diverse and dynamic microbial community, both associated with corals and the water column [14]. The microbial communities and processes of coral reefs have been investigated in depth, along various reef types, impact statuses, and oceans, for more than 50 years [15–18]. Especially in shallow reefs, the entirety of microbial communities consists of three sub-communities, benthic bacteria, the bacterioplankton of the water column, and the coral-associated microbiome, each influencing each other [19,20].

Bacterial communities in reefs have been increasingly profiled using high-throughput sequencing, revealing patterns related to reef zone, host association, and environmental conditions [3,21,22]. Protists, by contrast, are often omitted or pooled under broad microbial categories, though they include diverse photoautotrophs, grazers, and parasites [4]. Studies that simultaneously examine both bacteria and protists are rare, yet such comparisons are essential for understanding community assembly processes across microbial domains. Furthermore, few investigations have assessed microbial diversity at multiple spatial scales, leaving open questions about how microbial communities vary not just between reefs but within them.

The shallowest and most thermally tolerant [23] reef systems globally are located in the Red Sea, a semi-enclosed basin characterized by high salinities [24] and elevated temperatures [25], within steep environmental gradients [26–28] across the whole Red Sea. However, microbial studies in the Red Sea remain sparse, and even fewer focus on free-living benthic microbial communities independent of host organisms. There is a particular need to understand how microbial communities respond to spatial variation in reef structure and exposure, especially given growing interest in microbial indicators of reef condition.

Here, we use environmental DNA (eDNA) metabarcoding to investigate bacterial and protist community structures across six coral reefs in the central Red Sea. At each reef, we sampled exposed (seaward) as well as sheltered (shoreward) sites, allowing us to test for microbial differentiation at two nested spatial scales: between reefs as well as between exposure categories within reefs. Sheltered and exposed reef sites of reef structures often differ in hydrodynamic forcing, temperature ranges, and nutrient availability [29,30], offering two distinct niches within the same ecosystem. We calculated diversity metrics, performed community composition analyses, and utilized both indicator value analysis and Random Forest classification models to identify key taxa associated with specific reefs and exposure sites (sheltered vs. exposed). These methods are increasingly applied in the monitoring of anthropogenic impacts [31–33] but remain underutilized in reef microbial ecology and for the monitoring of reef structures and health. These analyses provide insights into reef microbial biodiversity and aid in the understanding and ultimately protection of these unique communities.

2. Materials and Methods

2.1. Sampling Stations

We sampled six coral reefs (Tahla, Al-Fahal, Shib Nazar, Qita' Al-Gersh, Abu Shosha, and King Abdullah Economic City (KAEC)) in the central Red Sea along the Saudi Arabian coast (Figure 1). At each reef, six stations were sampled in total: in detail, three at sheltered (shoreward) and exposed (seaward) sites, respectively. Metadata including GPS location and observed environmental parameters are listed in Supplementary Table S1. At each station, surface sediment (max. 5 cm depth) was sampled by scuba divers using 50 mL sterile prelabelled VWR® polypropylene centrifuge tubes. The tubes remained closed underwater until sampling and were closed immediately after sediment was sampled by pushing the sterile tube approx. 5 cm into the sediment and scooping upwards. Divers avoided touching the inside of the tubes during the sampling process. Afterwards, tubes

were placed in individual labeled bags and kept in coolers for transport (approx. 2 h). Samples were further processed in a molecular laboratory at King Abdullah University of Science and Technology (KAUST), and seawater was first discarded, followed by the addition of 4 mL of LifeGuard solution (Qiagen, Hilden, Germany) and storage at 4 °C until further processing.

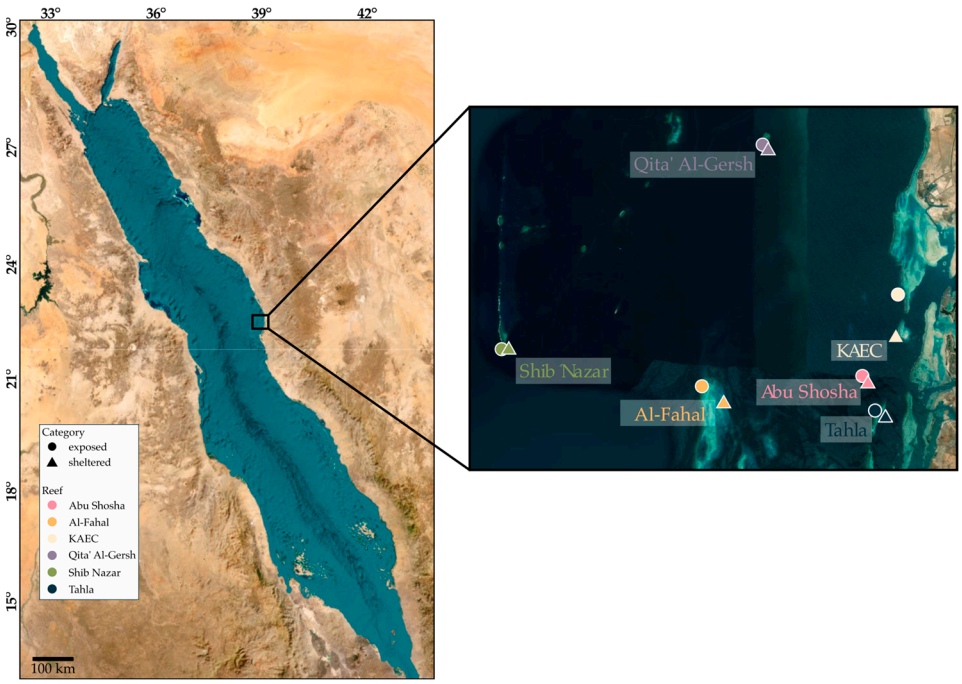


Figure 1. Overview of coral reef sampling locations at the Saudi Arabian Eastern Red Sea coast. Individual dots highlight the locations of the six reefs, and six stations were sampled in triplicate within each reef (three exposed sites and three sheltered sites), totaling 108 individual samples. The average sampling depth was 9 m at Abu Shosha, 7 m at Al-Fahal, 8.5 m at KAEC, 24.5 m at Qita' Al-Gersh, 21 m at Shib Nazar, and 5.5 m at Tahla. The sampling depth varied between exposed and sheltered sites; for more details, see the Supplementary Table S1.

2.2. Laboratory Processing

Environmental DNA was extracted from each sample using the PowerSoil Extraction Kit (Qiagen, Hilden, Germany) following a modified version of the manufacturer's protocol, in which initial bead beating lysis was increased to 20 min. The quality and quantity of extracted eDNA were measured using a Qubit fluorometer (Invitrogen, Waltham, MA, USA) and Nanodrop spectrophotometer (Thermo Fisher Scientific, Waltham, MA, USA). Subsequently, eDNA was amplified using PCR (polymerase chain reaction) with two sets of primers that targeted different barcodes (16S27F534R and 18SV9M, Table 1). Amplification and sequencing were carried out by eDNAtec at the Centre for Environmental Genomics Applications (CEGA), St. Johns, NL, Canada. The datasets were amplified separately and then paired-end sequenced on an Illumina NovaSeq6000 platform with 2 × 150 bp (18SV9M) and 2 × 250 bp (16S27F534R) read lengths.

Table 1. Target regions and primers used for eDNA amplification.

Primer	Gene	Target Taxa	References
16S27F534R	16S	Bacteria (BAC)	Muyzer et al. 1993, Weisburg et al. 1992 [34,35]
18SV9M	18S	Protists (PRO)	Stoeck et al. [36]

2.3. Bioinformatics Processing

Demultiplexing, adaptor trimming, and primer trimming were carried out by the sequencing facility. We processed both datasets into ESVs with the metabarcoding pipeline Apscale v1.6.3 [37], which involves vsearch v2.20.0 [38], cutadapt v4.4 [39], and a python adaptation of LULU [40]. We ran the Apscale pipeline with a mean expected error cutoff value of 2 (as opposed to the default value of 1, which keeps more erroneous reads during the quality filtering step) and otherwise default parameters.

Taxonomic assignments of ESVs were obtained using BLAST v2.15.0 [41] against the SILVA v138.1 SSU Ref NR99 reference database [42] for bacteria and the PR2 v5.0.0 SSU reference database [43] for protist sequences with an E-value threshold of 1×10^{-5}. We filtered BLAST v2.15.0 hits as follows: (1) all hits with a bitscore <150 and an alignment length of <100 were excluded; (2) for every ESV, all hits whose bitscore did not fall within a 2% margin of the highest bitscore were excluded; (3) if multiple taxa occurred among the remaining hits, only their Lowest Common Ancestor was retained; and (4) taxonomic lineages were trimmed based on percentage identity to the query; specifically, we trimmed at 100%, 95%, 90%, 85%, 80%, and 75% for the species, genus, family, order, class, and phylum levels, respectively, meaning that only hits with a percentage identity score of >95% were assigned to the genus level and so forth. This stringent filtering approach ensured the minimization of false-positive detections.

2.4. Data Analysis

All data analysis was performed in R_4.3.1(2023-06-16)–"Beagle Scouts" using RStudio_2023.06.1+524. All code and supporting files are publicly available and can be found on GitHub (https://github.com/lexscience/Molecular-Coral-Reef-Inventories). Sequence data was deposited in the Sequence Read Archive (SRA) of NCBI under project accession number PRJNA1000118.

The produced ESV-to-sample matrices were stored in a phyloseq object using the phyloseq package v1.44.0 [44]. Using the decontam package v1.20.0 [45], potential contaminant sequences from blanks and negative controls were removed from the datasets. Further, taxonomic hits were checked for each dataset and only sequences with target taxonomy were kept for further analysis (e.g., phylum level hit = "Bacteria" for the bacterial dataset). Exact taxonomic filtering parameters are listed in the code available on GitHub. All sequences assigned to *Homo sapiens* were considered contamination and were removed from all datasets.

2.4.1. Taxonomic Overview Visualization

To visualize microbial community composition across reefs and categories, we generated stacked barplots of relative taxonomic abundance at the phylum level using phyloseq. ESVs were agglomerated at the phylum rank using the tax_glom() function. The data were then transformed to relative abundances using transform_sample_counts() by dividing each taxon's abundance by the total reads per sample. The resulting community profiles were plotted with ggplot2 v3.5.1., grouped by both reef and category. Taxa with very low relative abundance across samples (<1%) were grouped into the category "Other" to improve visual interpretability.

2.4.2. Alpha Diversity Statistics

Alpha diversity measures (observed richness, Shannon diversity, and Simpson diversity) were calculated for bacterial and protist communities using the estimate_richness() function. Boxplots were generated to visualize alpha diversity distributions across reefs and exposure sites. Kruskal–Wallis tests (non-parametric ANOVA-equivalent) were used to test

whether alpha diversity metrics differed significantly among reefs within each exposure site. This was performed separately for each metric and category. Following significant Kruskal–Wallis results, pairwise Wilcoxon rank-sum tests with Benjamini–Hochberg *p*-value adjustments were performed to identify which reefs differed significantly in alpha diversity within each reef exposure site.

2.4.3. Beta Diversity Statistics

To assess differences in microbial community composition, we conducted permutational multivariate analysis of variance (PERMANOVA) using the vegan v2.6-4 [46] and phyloseq packages. We first tested for within-reef differences by grouping samples by reef and running PERMANOVA for each reef separately to evaluate the impact of exposure category on community composition. Next, we performed a between-reef PERMANOVA across all samples to assess differences among reefs irrespective of exposure category. Additionally, to test for combined and interactive effects of reef identity and exposure category (sheltered vs. exposed), we conducted a two-way PERMANOVA using CLR-transformed community data and Euclidean distances. Homogeneity of dispersion was evaluated for each factor using betadisper and permutest, confirming that multivariate dispersion did not differ significantly among groups. All analyses used 999 permutations for significance testing.

2.4.4. Redundancy Analysis

We performed redundancy analysis (RDA) to identify and visualize environmental variables (depth) and categorical variables (exposure category and reef) that significantly impacted community composition. Therefore, ESV abundances were grouped at the phylum level, low-abundance ESVs were removed, and abundances were center log-ratio (CLR) transformed to account for the compositional nature of sequencing data and to enable Euclidean-based ordinations following recommendations by Gloor et al. [47]. Significant environmental variables were identified by performing stepwise selection starting from a null RDA model (with no constraints), using the ordistep() function in both directions with 999 permutations. A full RDA model with all environmental variables as constraints defined the scope of selectable environmental variables. The RDA was then plotted with only the significant variables as constraints.

2.4.5. Indicator ESV Analysis

We used two methods to identify ESVs that significantly indicated either exposed or sheltered reef sites, specifically the indicator value approach [48] and Random Forest (RF) classification [49]. For the indicator value approach, we used the multipatt() function from the indicspecies R package v1.7.15 [48] based on the IndVal.g method with 999 permutations. Taxa with statistically significant associations ($p < 0.05$) were retained. The results were visualized by plotting the top 20 indicator ESVs with the highest indicator values (stat) for bacteria and protists separately, annotated with their full taxonomic classification (phylum and species level), and highlighted by the associated exposure category (sheltered or exposed). For RF classification, ESV tables were extracted from both bacterial and protistan phyloseq objects, and ESVs with zero counts across all samples were removed. RF classification was performed using the randomForest v4.7-1.2 R package [50] with 500 trees and otherwise default parameters. The RF model was trained using exposure category (sheltered or exposed) as the response variable and ESV abundance as the predictor variable. Model performance was assessed using the out-of-bag (OOB) error rate. Variable importance was determined based on the Mean Decrease in Accuracy (MDA). To quantify misclassifications, confusion matrices were generated for both bacterial and protistan datasets.

3. Results

3.1. Community Composition Overview

For the bacterial dataset, bioinformatic processing and data cleaning resulted in a total of 25,605 ESVs, of which 28% were assigned to the phylum Proteobacteria, 23% to Planctomycetota, and 8% to the SAR324 clade. A total number of 10,415 ESVs were obtained from the protistan dataset, with 35% assigned to the Alveolata clade, 23% grouped into the Rhizaria clade, and 18% assigned to the Stramenopiles clades, all three belonging to the SAR supergroup within the eukaryotic domain.

The taxonomic composition of reefs was homogenous within the bacterial communities, with the Abu Shosha and Al-Fahal reefs differing from the four remaining reefs in the relative abundance of the most dominant phyla and classes (Figure 2a). The protistan community compositions were even more homogenous (Figure 2b) and showed similar relative abundance patterns for the most abundant groups. Protistan communities overall showed a slight increase in the abundance of the Rhizaria phylum at exposed sites and a more pronounced shift within the Alveolata phylum for the Tahla and Abu Shosha reefs.

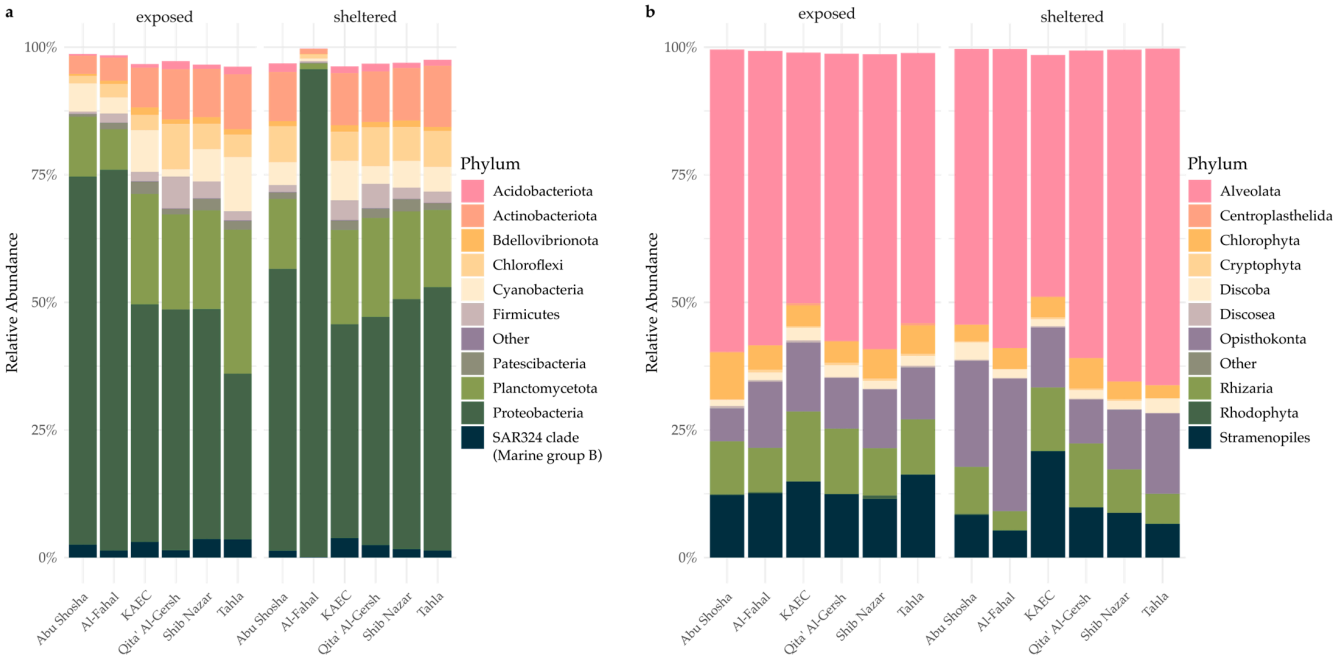


Figure 2. Relative abundance of microbial taxa at the phylum level for (**a**) bacteria and (**b**) protists across reefs and exposure sites. The stacked barplots show the mean relative abundance of phyla grouped by reef and exposure category (sheltered vs. exposed). Only the most abundant phyla are shown; low-abundance taxa (<1% across samples) are grouped under "Other".

Protistan and bacterial communities showed high richness and alpha diversity across Red Sea reefs, both at sheltered and exposed sites (Figure 3). Bacterial communities exhibited generally higher observed richness and alpha diversity compared to protistan communities. Notably, sheltered sites from the Al-Fahal reef showed reduced bacterial alpha diversity, while sheltered sites from the Tahla and Abu Shosha reefs displayed elevated protistan alpha diversity.

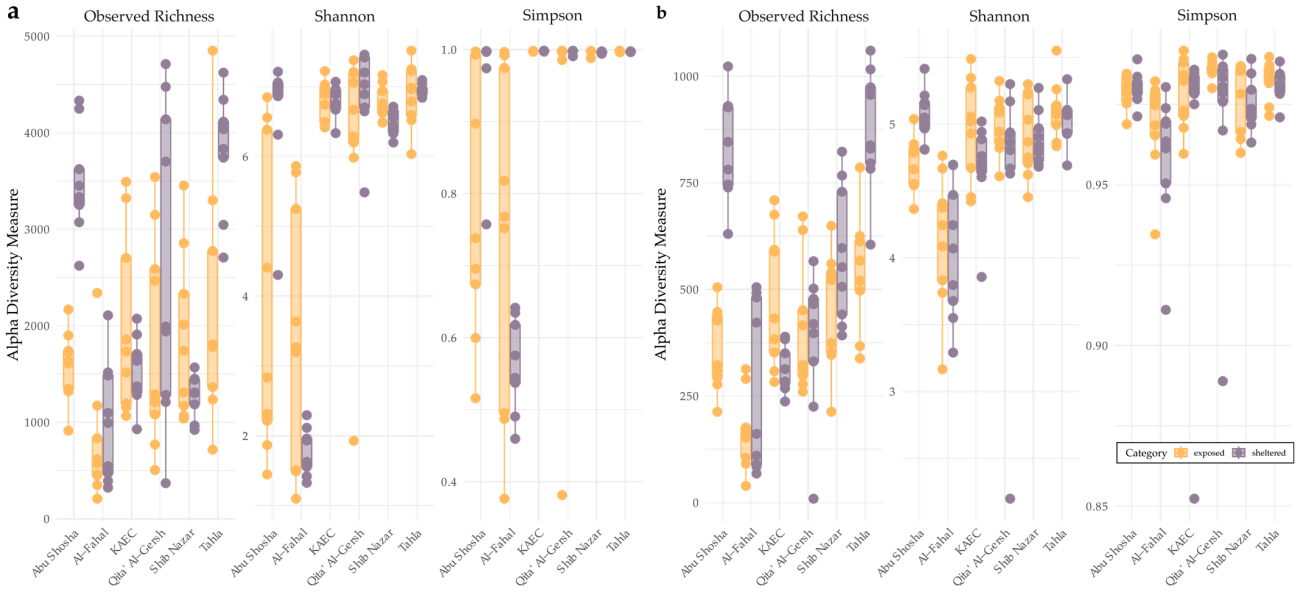


Figure 3. Alpha diversity of bacterial (**a**) and protist (**b**) communities across reefs and exposure sites. The boxplots show observed richness, Shannon diversity, and Simpson diversity indices for (**a**) bacteria and (**b**) protists at each reef, highlighted by sheltered (purple) and exposed (orange) sites. Each dot represents an individual sample.

3.2. Impact of Reef and Exposure Site on Microbial Community Composition

Across the sampled coral reefs, alpha diversity analyses revealed pronounced within- and between-reef heterogeneity for both bacteria and protists. Kruskal–Wallis tests confirmed significant alpha diversity differences between reefs across both exposed (seaward) or sheltered (shoreward) categories. Pairwise Wilcoxon tests further revealed that observed richness and Shannon diversity differed significantly between multiple reef pairs across both microbial groups, particularly among Al-Fahal, Abu Shosha, and Tahla, regardless of whether samples were exposed or sheltered. For example, Al-Fahal (exposed) consistently diverged from other exposed reefs including Tahla, Shib Nazar, KAEC, and Qita' Al-Gersh (adjusted $p < 0.01$). Similarly, Abu Shosha (both exposed and sheltered) showed repeated contrasts with other reefs in both microbial groups. These patterns were corroborated by beta diversity analyses, which detected significant compositional dissimilarities between reefs (between-reef PERMANOVA, bacteria: $R^2 = 0.27$, $F = 7.04$, $p = 0.001$; protists: $R^2 = 0.15$, $F = 3.60$, $p = 0.001$), confirming that reef identity was a major driver of microbial community structure.

Within-reef PERMANOVA also revealed significant differences in microbial community composition between exposed and sheltered samples across all six reefs (Table 2). The effect size varied by reef, with KAEC showing the strongest separation within both microbial groups while Qita' Al-Gersh showed the weakest but still significant separation. Additionally, two-way PERMANOVA revealed significant reef × exposure interactions for both bacterial ($R^2 = 0.111$, $F = 3.29$, $p = 0.001$) and protistan communities ($R^2 = 0.092$, $F = 2.35$, $p = 0.001$), further solidifying that the influence of exposure (sheltered vs. exposed) varied across reefs. Residual variation remained high (65% and 75%, respectively), suggesting that additional unmeasured environmental or biotic factors contribute to community structure.

Table 2. PERMANOVA results for within-reef differences between exposed and sheltered samples for both microbial groups. * = 0.01 < $p \leq 0.05$; ** = 0.001 < $p \leq 0.01$; *** = $p \leq 0.001$.

Reef	Bacteria		Protists	
	R^2	p	R^2	p
KAEC	0.41	0.001 ***	0.34	0.001 ***
Abu Shosha	0.35	0.001 ***	0.25	0.002 **
Tahla	0.28	0.002 **	0.11	0.014 *
Al-Fahal	0.17	0.008 **	0.13	0.003 **
Shib Nazar	0.13	0.001 ***	0.16	0.001 ***
Qita' Al-Gersh	0.11	0.01 **	0.1	0.005 **

Redundancy analysis revealed that reef identity and exposure category significantly influenced bacterial and protistan community compositions (Figure 4). Additionally, for bacteria, depth was also a significant variable. Importantly, the analysis revealed similar patterns for both domains: communities inhabiting the reef Al-Fahal were most distinguishable, regardless of domain, and differed most from communities inhabiting the reef KAEC. Communities from the reef Qita' Al Gersh fell in between the former two, with communities from the reefs Shib Nazar and Tahla being the most different from Qita' Al Gersh communities. This indicates that individual reefs provide distinct ecological niches that shape communities across domains. However, while both RDAs were significant, the proportion of variance explained by the first two ordination axes was only 11.9% (bacteria) and 9% (protists), indicating that the majority of community variance is explained by other, non-measured abiotic and biotic factors.

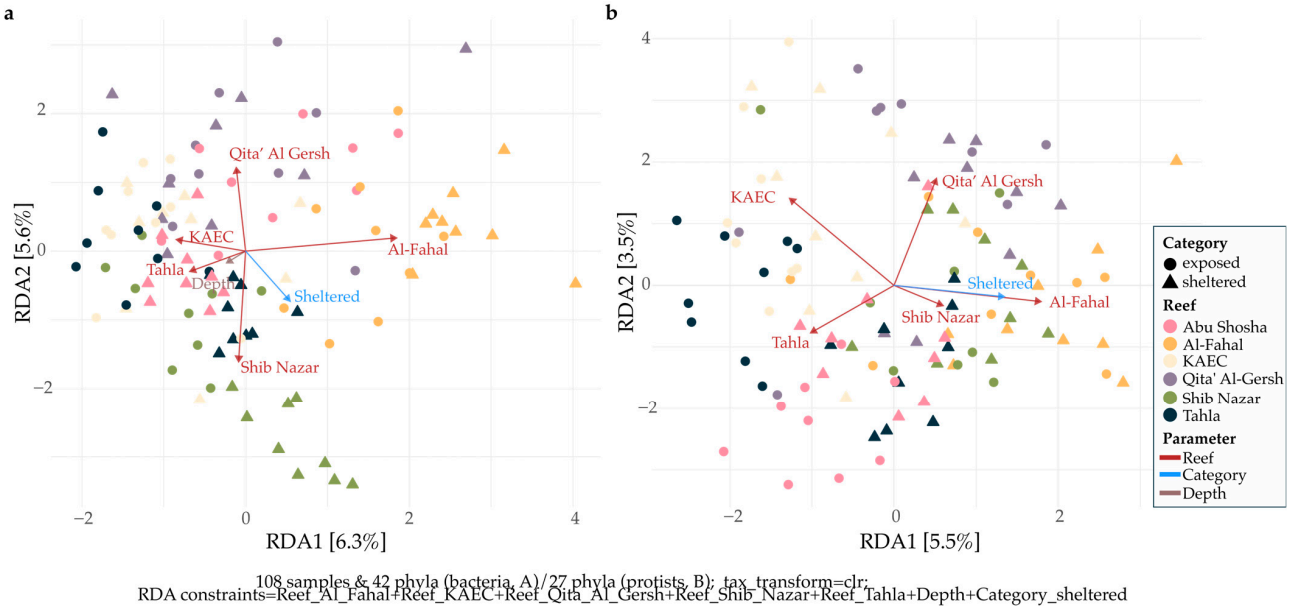


Figure 4. Redundancy analysis (RDA) of microbial community composition across Red Sea coral reefs and exposure categories. Panels show ordination results for bacterial (**a**) and protist (**b**) communities. Community data were CLR-transformed prior to ordination to account for the compositional nature of sequencing data. Points represent individual samples, colored by reef and shaped by category (sheltered vs. exposed). Axis labels indicate the proportion of variance explained by each ordination axis.

Notably, while both alpha and beta diversity analyses converged in identifying reef identity and exposure category as key structuring factors, they also revealed nuanced differences. Beta diversity detected significant compositional shifts even in cases where Simpson diversity indicated no major changes in evenness, suggesting that subtle taxonomic turnover—not necessarily changes in dominant taxa—can drive community dissimilarity. Conversely, reefs showing strong differences in richness (e.g., via observed or Shannon metrics) also exhibited compositional dissimilarity, indicating a general alignment between taxonomic richness and broader community structure.

3.3. Microbial ESVs as Bioindicators

Indicator species analysis revealed distinct sets of bacterial and protistan ESVs strongly associated with either sheltered or exposed sites. For both groups, we identified taxa with high indicator values (stat > 0.6), suggesting strong specificity and fidelity to the respective categories (Figure 5). For bacteria, ESV_158 assigned to the genus Methyloceanibacter showed the highest indicator values, followed closely by ESV_599 and ESV_507. However, no specific taxonomic group stood out as indicators (Figure 5a). For the protistan dataset, two ESVs assigned to the phylum Alveolata, two assigned to Rhizaria, and one to Opisthokonta exhibited the highest indicator values, with a noteworthy drop-off for the following ESVs (Figure 5b).

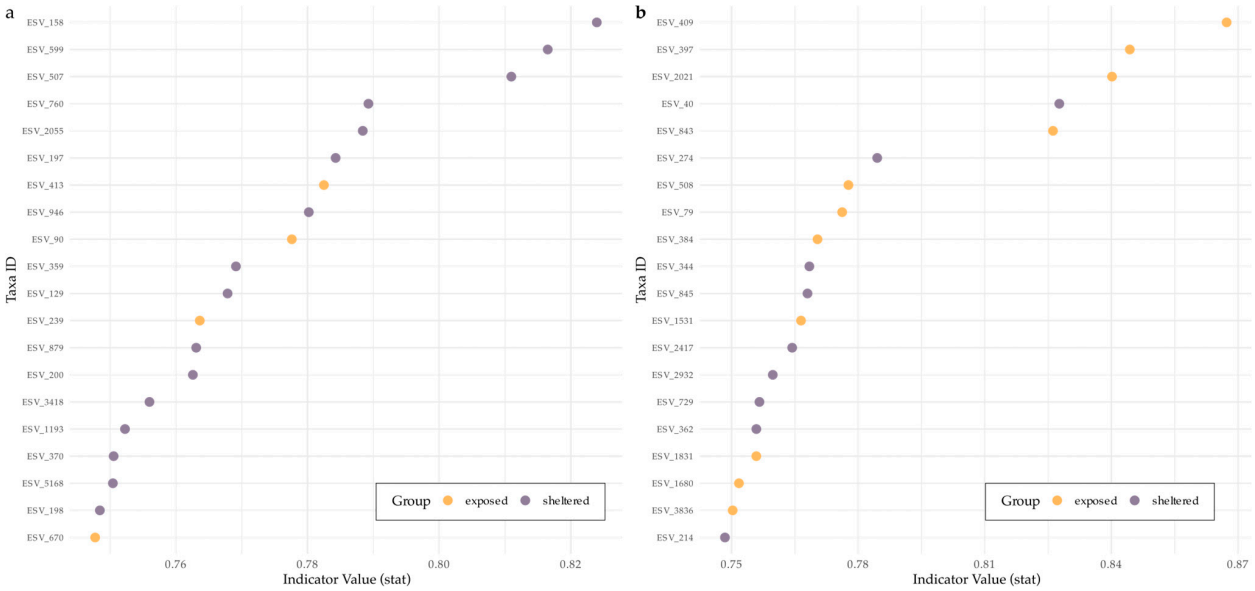


Figure 5. Top 20 indicator species for each exposure category identified by the IndVal method. (**a**) Bacterial ESVs and (**b**) protistan ESVs showing the highest indicator values (stat) based on association with either exposed or sheltered sites. Taxa are annotated at the phylum and species level and colored by the category with which they are significantly associated ($p < 0.05$). Indicator value (IndVal) reflects both the specificity and fidelity of each taxon to a particular exposure category.

Random Forest models performed well for both datasets, with 500 trees built for each dataset and out-of-bag (OOB) errors of 12.04% for bacteria and 16.67% for protists, resulting in the accurate classifications of more than 80% of all samples. Of 54 exposed samples, 52 were classified correctly using the bacterial dataset, whereas only 45 were correctly classified for the protist dataset (Table 3a,b). For the 54 sheltered samples, 43 and 45 were classified correctly for the bacterial and protistan dataset, respectively (Table 3a,b).

Table 3. (**a**) Confusion matrix of the Random Forest model for bacteria; (**b**) confusion matrix of the Random Forest model for protists.

(**a**)			
	Exposed	Sheltered	Class Error
Exposed	52	2	3.7%
Sheltered	11	43	20.4%
(**b**)			
	Exposed	Sheltered	Class Error
Exposed	45	9	16.67%
Sheltered	9	45	16.67%

Model performance was higher when classifying exposed samples compared to sheltered samples for bacteria, whereas protists performed similarly for both exposure categories. Mean Decrease Accuracy (MDA) for bacteria (Figure 6a) and protists (Figure 6b) illustrate the importance of each ESV in classifying the samples, specifically showcasing how much the models' accuracy would decrease with the removal of the respective ESVs from the datasets. For bacteria, specifically ESV_599 (Chloroflexi), ESV_45 (Cyanobacteria-Xenococcus sp.), ESV_413 (SAR324_clade-Marine Group B), and ESV_8 (Proteobacteria-Cohaesibacter sp.) were more influential in distinguishing exposed from sheltered sites. For protists, ESV_2021(Rhizaria-Miliolida), ESV_362 (Alveolata-Gregarinomorphea), ESV_175 (Alveolata-Colpodellidea), and ESV_409 (Alveolata-Vitrellaceae) had a significant contribution to model accuracy.

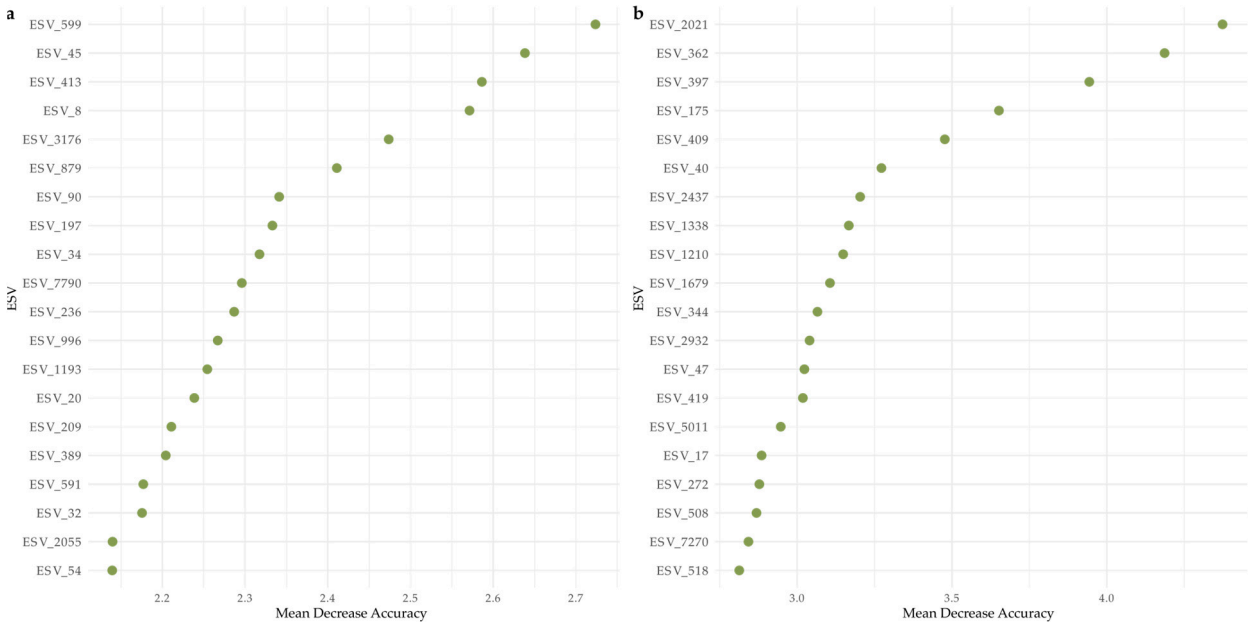


Figure 6. Variable importance based on Mean Decrease Accuracy (MDA) from Random Forest models classifying sheltered vs. exposed sites. Panel (**a**) shows the top bacterial ESVs, and panel (**b**) shows the top protistan ESVs ranked by their contribution to model accuracy. Higher MDA values indicate a greater loss in classification accuracy when the feature is permuted.

Overall, the indicator value approach and Random Forest classification shared several ESVs classified as "indicators". For bacteria, five ESVs were identified by both approaches, whereas eight ESVs were identified by both approaches for the protistan dataset (Table 4a,b).

Table 4. (**a**) Overview of bacterial ESVs identified by both the indicator value approach and Random Forest classification. Empty fields indicate incomplete taxonomic assignment; (**b**) overview of protistan ESVs found by both the indicator value approach and Random Forest classification. Empty fields indicate incomplete taxonomic assignment.

(**a**)						
ESV	Domain	Phylum	Class	Order	Family	Genus
ESV_158	Bacteria	Proteobacteria	Alphaproteobacteria	Rhizobiales	Methyloligellaceae	*Methyloceanibacter*
ESV_599	Bacteria	Chloroflexi	KD4-96	Uncult. bacterium		
ESV_2055	Bacteria	Actinobacteriota	Acidimicrobiia	Actinomarinales	Uncult. bacterium	
ESV_413	Bacteria	SAR324 clade(Marine Group B)	Uncult. bacterium			
ESV_879	Bacteria	Proteobacteria	Alphaproteobacteria	Rhodobacterales	Rhodobacteraceae	*Marinovum*

(**b**)							
ESV	Domain	Phylum	Class	Order	Family	Genus	Species
ESV_409	Eukaryota	Alveolata	Colpodellidea	Vitrelladida	Vitrellaceae		
ESV_397	Eukaryota	Alveolata	Litostomatea	Haptoria_5	Pleurostomatida	*Epiphyllum*	*E. shenzhenense*
ESV_2021	Eukaryota	Rhizaria	Tubothalamea	Miliolida			
ESV_40	Eukaryota	Opisthokonta					
ESV_843	Eukaryota	Rhizaria					
ESV_1531	Eukaryota	Rhizaria	Globothalamea	Rotaliida			
ESV_2932	Eukaryota	Rhizaria	Tubothalamea	Miliolida	Moliolidae		
ESV_362	Eukaryota	Alveolata	Gregarinomorphea				

4. Discussion

Our study revealed clear multiscale patterns in bacterial and protist communities across coral reefs in the central Red Sea. By sampling both sheltered and exposed sites of six reefs, we demonstrated that microbial diversity and composition are structured not only by broad geographic differences between reefs, but also by finer-scale environmental variation within reefs. Specifically, sheltered (shoreward) and exposed (seaward) reef sites represent distinct microhabitats with exposed sites being characterized by higher water flow, greater light penetration, and reduced sediment accumulation, while sheltered sites exhibit calmer hydrodynamics, enhanced sedimentation, and greater temperature variability. These physicochemical contrasts provide distinct ecological microniches that strongly influence microbial community assembly.

4.1. Reef-Specific Bacterial and Protist Assemblages in the Red Sea

Beta diversity analyses showed significant compositional differences between reefs for both bacterial and protist communities, suggesting that each reef harbors a distinct microbial signature. These patterns likely reflect a combination of environmental filtering, stochastic colonization, and possibly host-associated community spillover from corals or macroalgae. The Red Sea is known for its pronounced local environmental gradients, including differences in wave energy, temperature microclimates, and benthic cover, which may create unique selective pressures at each exposure category.

This exposure specificity is consistent with findings from other marine ecosystems [51–56], where microbial communities often exhibit biogeographic structure despite the potential for wide dispersal [57].

For instance, studies in the Caribbean [58] and Indo-Pacific [59] have reported strong spatial turnover in reef microbial assemblages. Our results extend these observations to include protists, a group still understudied in reef microbiome research [4] but equally responsive to environmental heterogeneity.

4.2. Within-Reef Community Differences Reveal Microhabitat-Driven Divergence

Importantly, we also observed consistent differences between sheltered and exposed sites within the same reef. Alpha diversity was often significantly higher in one exposure category compared to the other, and beta diversity analyses showed that exposure drives community divergence even at this fine spatial scale. These results highlight the ecological relevance of reef microhabitats in structuring microbial communities. In sheltered sites, reduced water exchange and higher sedimentation can create environments enriched in organic matter that favor heterotrophic and anaerobic-associated taxa. In contrast, exposed sites with high water movement and light penetration select for photoautotrophic and mixotrophic taxa, alongside aerobic bacteria adapted to rapid organic matter turnover.

Sheltered and exposed reef sites differ in water quality, light penetration, sedimentation, and temperature variability owing to physical wave movement—all factors known to influence microbial composition [60,61]. Previous studies have documented shifts in microbial communities across depth and exposure gradients [10,62–64] including differences in, e.g., cyanobacterial abundance [65] and mixotrophic protists [66]. Our results reinforce the need to treat within-reef variation as an ecologically meaningful unit of study, particularly when designing microbial monitoring programs.

4.3. Cross-Domain Comparisons Reveal Overlapping but Non-Redundant Patterns

While bacterial and protist communities responded similarly to spatial structuring, they also exhibited domain-specific patterns. For example, certain reefs showed strong protist differentiation but weaker bacterial divergence, and vice versa. This highlights the complementary value of analyzing multiple microbial domains: while both groups are shaped by environmental conditions, their ecological niches and life strategies differ. Protists often function as grazers, phototrophs, or parasites [67], while bacteria dominate roles in decomposition, nutrient cycling, parasitism, and host symbiosis [63,68–70]. Combining both groups offers a more complete picture of microbial ecosystem dynamics covering multiple levels of the trophic cascade.

4.4. Indicator Taxa and Classification Models Support Community-Based Monitoring

Using indicator value analysis and Random Forest classification, we identified indicator taxa that reliably distinguish samples between reefs and exposure categories. These approaches are increasingly employed in microbial ecology to move beyond global patterns and toward actionable indicators. Many of the indicator taxa we identified belong to groups known to respond quickly to environmental change—including dinoflagellates [71] and ciliate protists [72,73].

Our results highlight habitat-specific microbial assemblages that may reflect environmental filtering or niche specialization. For example, bacterial taxa such as *Methyloceanibacter* and Chloroflexi clade members were associated with low-flow sheltered habitats, whereas members of the SAR324 clade were linked to exposed sites. Similarly, protistan indicators such as the alveolate classes Gregarinomorphea and rhizarian foraminifera (Miliolida) characterized sheltered microhabitats, while alveolate lineages like Vitrellaceae or ciliated *E. shenzhenense* were enriched in exposed sites. These patterns demonstrate how physicochemical conditions structure distinct microbial assemblages across reef microhabitats. Multiple indicator taxa were identified by both methods, making them strong bioindicator candidates. Additionally, the ability to classify samples into exposure categories using bioindicator taxa further highlights the differences in environmental characteristics between sheltered and exposed sites, such as water residency time and particulate matter suspension, which are likely to affect microbial communities [30].

While some indicator taxa were shared across microbial domains, others were specific to particular conditions, highlighting the potential of using multi-domain signatures for environmental assessment. Notably, the Random Forest models achieved higher classification accuracy compared to the indicator value approach, which corroborates with results from Frühe et al. [33], supporting the robustness of these taxa and methods as predictors of spatial structure. Such tools could be integrated into future reef monitoring frameworks to track ecosystem shifts or detect early signs of degradation.

The indicator taxa identified in our study reveal how sheltered and exposed reef microhabitats select for distinct microbial guilds with characteristic ecological functions. Among bacteria, *Methyloceanibacter* (Alphaproteobacteria) was strongly associated with sheltered sites. Members of this genus are known for methylotrophy, and their distribution patterns have been confirmed to be affected by carbon and nitrogen cycling and salinity, especially in coastal areas [74], traits that may be favored in high-flow, oxygenated environments where dissolved organic carbon is rapidly cycled. Similarly, SAR324 clade members, previously known as Marine Group B, are commonly detected in marine environments with a preference for deeper pelagic layers and broad metabolic versatility [75,76], and indicate exposed reef sites in this study. Chloroflexi (KD4-96), associated with heterotrophic degradation of organic matter in sediments [77], were enriched in sheltered sites with higher water residence time, less flushing, and higher sedimentation rates, indicating a potentially higher concentration of organic matter on the seafloor.

Several protistan lineages emerged as potential indicators of the contrasting microhabitats. Chromerids (e.g., Vitrellaceae) and foraminifera (Miliolida) were indicative of exposed reef sites. Chromerids are photosynthetic alveolates closely related to apicomplexans [78] and were originally isolated from corals [79]; their genomes and physiology endorse a phototrophic lifestyle compatible with high-light, oxygenated reef margins, where flow enhances nutrient exchange and reduces sediment resuspension (thus maintaining light availability). Likewise, miliolid foraminifera are classic inhabitants of high-light, carbonate-dominated shallow reefs, with their morphology designed for thriving under stable illumination and water movement [80]-conditions typical of exposed reef flats and crests. Their occurrence has long been used as a bioindicator of healthy reef systems [81,82]. Notably, another ESV assigned to the foraminiferal order of Miliolida in this study showed the highest indicator values for sheltered sites.

Taken together, these associations suggest that hydrodynamics and sedimentation not only structure microbial diversity but also filter for distinct metabolic strategies and morphological adaptions: rapid aerobic carbon cycling and phototrophy in exposed reefs versus particle-associated heterotrophy and mixotrophy in sheltered reefs. Such taxa may therefore serve as baseline bioindicators of microhabitat conditions, offering a functional lens through which to interpret microbial community change in coral reef ecosystems.

4.5. Implications and Future Directions

Together, our findings demonstrate that both bacteria and protists show structured, multiscale patterns of diversity in Red Sea coral reefs. These results underscore the value of including multiple microbial groups and sampling spatial gradients when studying reef microbiomes. The consistent within-reef differences suggest that small-scale heterogeneity—both ecologically and methodologically—matters and should be incorporated into sampling designs. More broadly, this work contributes to a growing effort of integrating microbial data retrieved from eDNA metabarcoding into coral reef monitoring and management. Future studies should investigate how these microbial patterns relate to environmental parameters (e.g., temperature, nutrients, benthic composition) and reef condition over time. The integration of eDNA, machine learning, and indicator-based frameworks holds

promise for building predictive models that can inform reef conservation strategies in a rapidly changing ocean.

Supplementary Materials: The following supporting information can be downloaded at https://www.mdpi.com/article/10.3390/microorganisms13112549/s1: Table S1: Metadata of sampling locations in the central Red Sea.

Author Contributions: Conceptualization, L.F.; methodology, L.F.; software, C.A.H. and L.F.; validation, C.A.H. and L.F.; formal analysis, C.A.H. and L.F.; investigation, C.A.H. and L.F.; resources, L.F.; data curation, C.A.H.; writing—original draft preparation, L.F.; writing—review and editing, C.A.H. and L.F.; visualization, C.A.H. and L.F.; supervision, L.F.; project administration, L.F.; funding acquisition, L.F. All authors have read and agreed to the published version of the manuscript.

Funding: This study was funded by grant GIF/5/4826-01-01 received by L.F. C.A.H was funded by BAS/1/1071-01-01 baseline funding from KAUST.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: All data analysis was performed in R_4.3.1(2023-06-16)–"Beagle Scouts" using RStudio_2023.06.1+524. All code and supporting files are publicly available and can be found on GitHub (https://github.com/lexscience/Molecular-Coral-Reef-Inventories). Sequence data was deposited in the Sequence Read Archive (SRA) of NCBI under project accession number PRJNA1000118.

Acknowledgments: We would like to thank Carlos Angulo-Preckler and Naira Pluma for their sampling efforts and the team from eDNAtec at the Centre for Environmental Genomics Applications (CEGA), St. Johns, Canada.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Costanza, R.; De Groot, R.; Sutton, P.; Van der Ploeg, S.; Anderson, S.J.; Kubiszewski, I.; Farber, S.; Turner, R.K. Changes in the global value of ecosystem services. *Glob. Environ. Change* **2014**, *26*, 152–158. [CrossRef]
2. Moberg, F.; Folke, C. Ecological goods and services of coral reef ecosystems. *Ecol. Econ.* **1999**, *29*, 215–233. [CrossRef]
3. Bourne, D.G.; Webster, N.S. Coral reef bacterial communities. In *The Prokaryotes*; Springer: Berlin/Heidelberg, Germany, 2013; pp. 163–187.
4. Ainsworth, T.; Fordyce, A.; Camp, E. The other microeukaryotes of the coral reef microbiome. *Trends Microbiol.* **2017**, *25*, 980–991. [CrossRef] [PubMed]
5. Bayer, T.; Neave, M.J.; Alsheikh-Hussain, A.; Aranda, M.; Yum, L.K.; Mincer, T.; Hughen, K.; Apprill, A.; Voolstra, C.R. The microbiome of the Red Sea coral Stylophora pistillata is dominated by tissue-associated Endozoicomonas bacteria. *Appl. Environ. Microbiol.* **2013**, *79*, 4759–4762. [CrossRef] [PubMed]
6. Galand, P.E.; Ruscheweyh, H.-J.; Salazar, G.; Hochart, C.; Henry, N.; Hume, B.C.; Oliveira, P.H.; Perdereau, A.; Labadie, K.; Belser, C. Diversity of the Pacific Ocean coral reef microbiome. *Nat. Commun.* **2023**, *14*, 3039. [CrossRef]
7. Delgadillo-Ordoñez, N.; Raimundo, I.; Barno, A.R.; Osman, E.O.; Villela, H.; Bennett-Smith, M.; Voolstra, C.R.; Benzoni, F.; Peixoto, R.S. Red Sea atlas of coral-associated bacteria highlights common microbiome members and their distribution across environmental gradients—A systematic review. *Microorganisms* **2022**, *10*, 2340. [CrossRef]
8. Osman, E.O.; Suggett, D.J.; Voolstra, C.R.; Pettay, D.T.; Clark, D.R.; Pogoreutz, C.; Sampayo, E.M.; Warner, M.E.; Smith, D.J. Coral microbiome composition along the northern Red Sea suggests high plasticity of bacterial and specificity of endosymbiotic dinoflagellate communities. *Microbiome* **2020**, *8*, 8.
9. Neave, M.J.; Apprill, A.; Aeby, G.; Miyake, S.; Voolstra, C.R. Microbial communities of Red Sea coral reefs. In *Coral Reefs of the Red Sea*; Springer: Berlin/Heidelberg, Germany, 2019; pp. 53–68.
10. Frühe, L.; Klein, S.G.; Angulo-Preckler, C.; Martynova, A.; Alamoudi, T.; García, J.V.A.; Arossa, S.; Breavington, J.; Frappi, S.; Laiolo, E. Particle-Associated Bacterioplankton Communities Across the Red Sea. *Environ. Microbiol.* **2025**, *27*, e70075. [CrossRef]
11. Gaidos, E.; Rusch, A.; Ilardo, M. Ribosomal tag pyrosequencing of DNA and RNA from benthic coral reef microbiota: Community spatial structure, rare members and nitrogen-cycling guilds. *Environ. Microbiol.* **2011**, *13*, 1138–1152. [CrossRef]
12. Naumann, M.S.; Richter, C.; el-Zibdah, M.; Wild, C. Coral mucus as an efficient trap for picoplanktonic cyanobacteria: Implications for pelagic–benthic coupling in the reef ecosystem. *Mar. Ecol. Prog. Ser.* **2009**, *385*, 65–76. [CrossRef]

13. Mayer, F.W.; Wild, C. Coral mucus release and following particle trapping contribute to rapid nutrient recycling in a Northern Red Sea fringing reef. *Mar. Freshw. Res.* **2010**, *61*, 1006–1014. [CrossRef]
14. Ducklow, H.W. The biomass, production and fate of bacteria in coral reefs. *Ecosyst. World* **1990**, *25*, 265–289.
15. Rochelle-Newall, E.; Torréton, J.-P.; Mari, X.; Pringault, O. Phytoplankton-bacterioplankton coupling in a subtropical South Pacific coral reef lagoon. *Aquat. Microb. Ecol.* **2008**, *50*, 221–229. [CrossRef]
16. Van Duyl, F.; Gast, G.; Steinhoff, W.; Kloff, S.; Veldhuis, M.; Bak, R. Factors influencing the short-term variation in phytoplankton composition and biomass in coral reef waters. *Coral Reefs* **2002**, *21*, 293–306. [CrossRef]
17. Sorokin, Y.I. Microbial production in the coral-reef community. *Pascal Fr. Bibliogr. Databases* **1978**, *83*, 281–323.
18. Yoshinaga, I.; Fukami, K.; Ishida, Y. Comparison of DNA and protein synthesis rates of bacterial assemblages between coral reef waters and pelagic waters in tropical ocean. *Mar. Ecol. Prog. Ser.* **1991**, *76*, 167–174. [CrossRef]
19. Yahel, R.; Yahel, G.; Genin, A. Daily cycles of suspended sand at coral reefs: A biological control. *Limnol. Oceanogr.* **2002**, *47*, 1071–1083. [CrossRef]
20. Rasheed, M.; Wild, C.; Franke, U.; Huettel, M. Benthic photosynthesis and oxygen consumption in permeable carbonate sediments at Heron Island, Great Barrier Reef, Australia. *Estuar. Coast. Shelf Sci.* **2004**, *59*, 139–150. [CrossRef]
21. Frade, P.R.; Glasl, B.; Matthews, S.A.; Mellin, C.; Serrão, E.A.; Wolfe, K.; Mumby, P.J.; Webster, N.S.; Bourne, D.G. Spatial patterns of microbial communities across surface waters of the Great Barrier Reef. *Commun. Biol.* **2020**, *3*, 442. [CrossRef]
22. Glasl, B.; Bourne, D.G.; Frade, P.R.; Thomas, T.; Schaffelke, B.; Webster, N.S. Microbial indicators of environmental perturbations in coral reef ecosystems. *Microbiome* **2019**, *7*, 94. [CrossRef]
23. Osman, E.O.; Smith, D.J.; Ziegler, M.; Kürten, B.; Conrad, C.; El-Haddad, K.M.; Voolstra, C.R.; Suggett, D.J. Thermal refugia against coral bleaching throughout the northern Red Sea. *Glob. Change Biol.* **2018**, *24*, e474–e484. [CrossRef]
24. Ngugi, D.K.; Antunes, A.; Brune, A.; Stingl, U. Biogeography of pelagic bacterioplankton across an antagonistic temperature–salinity gradient in the Red Sea. *Mol. Ecol.* **2012**, *21*, 388–405. [CrossRef] [PubMed]
25. Yao, F.; Hoteit, I. Rapid red sea deep water renewals caused by volcanic eruptions and the north atlantic oscillation. *Sci. Adv.* **2018**, *4*, eaar5637. [CrossRef] [PubMed]
26. Raitsos, D.E.; Yi, X.; Platt, T.; Racault, M.F.; Brewin, R.J.; Pradhan, Y.; Papadopoulos, V.P.; Sathyendranath, S.; Hoteit, I. Monsoon oscillations regulate fertility of the Red Sea. *Geophys. Res. Lett.* **2015**, *42*, 855–862. [CrossRef]
27. Chaidez, V.; Dreano, D.; Agusti, S.; Duarte, C.; Hoteit, I. Decadal trends in Red Sea maximum surface temperature. *Sci. Rep.* **2017**, *7*, 8144. [CrossRef]
28. Sofianos, S.S.; Johns, W.E. An oceanic general circulation model (OGCM) investigation of the Red Sea circulation, 1. Exchange between the Red Sea and the Indian Ocean. *J. Geophys. Res. Ocean.* **2002**, *107*, 17-1–17-11. [CrossRef]
29. Hench, J.L.; Leichter, J.J.; Monismith, S.G. Episodic circulation and exchange in a wave-driven coral reef and lagoon system. *Limnol. Oceanogr.* **2008**, *53*, 2681–2694. [CrossRef]
30. Nelson, C.E.; Alldredge, A.L.; McCliment, E.A.; Amaral-Zettler, L.A.; Carlson, C.A. Depleted dissolved organic carbon and distinct bacterial communities in the water column of a rapid-flushing coral reef ecosystem. *ISME J.* **2011**, *5*, 1374–1387. [CrossRef]
31. Cordier, T.; Esling, P.; Lejzerowicz, F.; Visco, J.; Ouadahi, A.; Martins, C.; Cedhagen, T.; Pawlowski, J. Predicting the ecological quality status of marine environments from eDNA metabarcoding data using supervised machine learning. *Environ. Sci. Technol.* **2017**, *51*, 9118–9126. [CrossRef]
32. Cordier, T.; Forster, D.; Dufresne, Y.; Martins, C.I.; Stoeck, T.; Pawlowski, J. Supervised machine learning outperforms taxonomy-based environmental DNA metabarcoding applied to biomonitoring. *Mol. Ecol. Resour.* **2018**, *18*, 1381–1391. [CrossRef]
33. Frühe, L.; Cordier, T.; Dully, V.; Breiner, H.W.; Lentendu, G.; Pawlowski, J.; Martins, C.; Wilding, T.A.; Stoeck, T. Supervised machine learning is superior to indicator value inference in monitoring the environmental impacts of salmon aquaculture using eDNA metabarcodes. *Mol. Ecol.* **2021**, *30*, 2988–3006. [CrossRef]
34. Muyzer, G.; De Waal, E.C.; Uitterlinden, A. Profiling of complex microbial populations by denaturing gradient gel electrophoresis analysis of polymerase chain reaction-amplified genes coding for 16S rRNA. *Appl. Environ. Microbiol.* **1993**, *59*, 695–700. [CrossRef]
35. Weisburg, W.G.; Barns, S.M.; Pelletier, D.A.; Lane, D.J. 16S ribosomal DNA amplification for phylogenetic study. *J. Bacteriol.* **1991**, *173*, 697–703. [CrossRef] [PubMed]
36. Stoeck, T.; Bass, D.; Nebel, M.; Christen, R.; Jones, M.D.; Breiner, H.W.; Richards, T.A. Multiple marker parallel tag environmental DNA sequencing reveals a highly complex eukaryotic community in marine anoxic water. *Mol. Ecol.* **2010**, *19*, 21–31. [CrossRef] [PubMed]
37. Buchner, D.; Macher, T.-H.; Leese, F. APSCALE: Advanced pipeline for simple yet comprehensive analyses of DNA metabarcoding data. *Bioinformatics* **2022**, *38*, 4817–4819. [CrossRef] [PubMed]
38. Rognes, T.; Flouri, T.; Nichols, B.; Quince, C.; Mahé, F. VSEARCH: A versatile open source tool for metagenomics. *PeerJ* **2016**, *4*, e2584. [CrossRef] [PubMed]
39. Martin, M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet. J.* **2011**, *17*, 10–12. [CrossRef]

40. Frøslev, T.G.; Kjøller, R.; Bruun, H.H.; Ejrnæs, R.; Brunbjerg, A.K.; Pietroni, C.; Hansen, A.J. Algorithm for post-clustering curation of DNA amplicon data yields reliable biodiversity estimates. *Nat. Commun.* **2017**, *8*, 1188. [CrossRef]
41. Altschul, S.F.; Gish, W.; Miller, W.; Myers, E.W.; Lipman, D.J. Basic local alignment search tool. *J. Mol. Biol.* **1990**, *215*, 403–410. [CrossRef]
42. Pruesse, E.; Quast, C.; Knittel, K.; Fuchs, B.M.; Ludwig, W.; Peplies, J.; Glöckner, F.O. SILVA: A comprehensive online resource for quality checked and aligned ribosomal RNA sequence data compatible with ARB. *Nucleic Acids Res.* **2007**, *35*, 7188–7196. [CrossRef]
43. Guillou, L.; Bachar, D.; Audic, S.; Bass, D.; Berney, C.; Bittner, L.; Boutte, C.; Burgaud, G.; De Vargas, C.; Decelle, J. The Protist Ribosomal Reference database (PR2): A catalog of unicellular eukaryote small sub-unit rRNA sequences with curated taxonomy. *Nucleic Acids Res.* **2012**, *41*, D597–D604. [CrossRef]
44. McMurdie, P.J.; Holmes, S. phyloseq: An R package for reproducible interactive analysis and graphics of microbiome census data. *PLoS ONE* **2013**, *8*, e61217. [CrossRef] [PubMed]
45. Davis, N.M.; Proctor, D.M.; Holmes, S.P.; Relman, D.A.; Callahan, B.J. Simple statistical identification and removal of contaminant sequences in marker-gene and metagenomics data. *Microbiome* **2018**, *6*, 226. [CrossRef] [PubMed]
46. Oksanen, J.; Kindt, R.; Legendre, P.; O'Hara, B.; Stevens, M.H.H.; Oksanen, M.J.; Suggests, M. The vegan package. *Community Ecol. Package* **2007**, *10*, 719.
47. Gloor, G.B.; Macklaim, J.M.; Pawlowsky-Glahn, V.; Egozcue, J.J. Microbiome datasets are compositional: And this is not optional. *Front. Microbiol.* **2017**, *8*, 2224. [CrossRef]
48. Cáceres, M.D.; Legendre, P. Associations between species and groups of sites: Indices and statistical inference. *Ecology* **2009**, *90*, 3566–3574. [CrossRef]
49. Breiman, L. Random forests. *Mach. Learn.* **2001**, *45*, 5–32. [CrossRef]
50. Liaw, A.; Wiener, M. Classification and regression by randomForest. *R News* **2002**, *2*, 18–22.
51. Ma, L.; Becker, C.; Weber, L.; Sullivan, C.; Zgliczynski, B.; Sandin, S.; Brandt, M.; Smith, T.B.; Apprill, A. Biogeography of reef water microbes from within-reef to global scales. *Aquat. Microb. Ecol.* **2022**, *88*, 81–94. [CrossRef]
52. Milke, F.; Wagner-Doebler, I.; Wienhausen, G.; Simon, M. Selection, drift and community interactions shape microbial biogeographic patterns in the Pacific Ocean. *ISME J.* **2022**, *16*, 2653–2665. [CrossRef]
53. Shao, Q.; Sun, D.; Fang, C.; Feng, Y.; Wang, C. Biodiversity and biogeography of abundant and rare microbial assemblages in the Western subtropical Pacific Ocean. *Front. Microbiol.* **2022**, *13*, 839562. [CrossRef] [PubMed]
54. Zakem, E.J.; McNichol, J.; Weissman, J.; Raut, Y.; Xu, L.; Halewood, E.R.; Carlson, C.A.; Dutkiewicz, S.; Fuhrman, J.A.; Levine, N.M. Functional biogeography of marine microbial heterotrophs. *Science* **2025**, *388*, eado5323. [CrossRef] [PubMed]
55. Zhang, Z.; Zhang, Q.; Chen, B.; Yu, Y.; Wang, T.; Xu, N.; Fan, X.; Penuelas, J.; Fu, Z.; Deng, Y. Global biogeography of microbes driving ocean ecological status under climate change. *Nat. Commun.* **2024**, *15*, 4657. [CrossRef] [PubMed]
56. Schwob, G.; Segovia, N.I.; González-Wevar, C.; Cabrol, L.; Orlando, J.; Poulin, E. Exploring the microdiversity within marine bacterial taxa: Toward an integrated biogeography in the Southern Ocean. *Front. Microbiol.* **2021**, *12*, 703792. [CrossRef]
57. Martiny, J.B.H.; Bohannan, B.J.; Brown, J.H.; Colwell, R.K.; Fuhrman, J.A.; Green, J.L.; Horner-Devine, M.C.; Kane, M.; Krumins, J.A.; Kuske, C.R. Microbial biogeography: Putting microorganisms on the map. *Nat. Rev. Microbiol.* **2006**, *4*, 102–112. [CrossRef]
58. Weber, L.; González-Díaz, P.; Armenteros, M.; Ferrer, V.M.; Bretos, F.; Bartels, E.; Santoro, A.E.; Apprill, A. Microbial signatures of protected and impacted Northern Caribbean reefs: Changes from Cuba to the Florida Keys. *Environ. Microbiol.* **2020**, *22*, 499–519. [CrossRef]
59. Littman, R.A.; Willis, B.L.; Pfeffer, C.; Bourne, D.G. Diversities of coral-associated bacteria differ with location, but not species, for three acroporid corals on the Great Barrier Reef. *FEMS Microbiol. Ecol.* **2009**, *68*, 152–163. [CrossRef]
60. Schwalbach, M.S.; Brown, M.; Fuhrman, J.A. Impact of light on marine bacterioplankton community structure. *Aquat. Microb. Ecol.* **2005**, *39*, 235–245. [CrossRef]
61. Abirami, B.; Radhakrishnan, M.; Kumaran, S.; Wilson, A. Impacts of global warming on marine microbial communities. *Sci. Total Environ.* **2021**, *791*, 147905. [CrossRef]
62. Brown, S.A.; Balmonte, J.P.; Hoarfrost, A.; Ghobrial, S.; Arnosti, C. Depth-related patterns in microbial community responses to complex organic matter in the western North Atlantic Ocean. *Biogeosciences* **2022**, *19*, 5617–5631. [CrossRef]
63. Moran, M.A. The global ocean microbiome. *Science* **2015**, *350*, aac8455. [CrossRef]
64. Sunagawa, S.; Coelho, L.P.; Chaffron, S.; Kultima, J.R.; Labadie, K.; Salazar, G.; Djahanschiri, B.; Zeller, G.; Mende, D.R.; Alberti, A. Structure and function of the global ocean microbiome. *Science* **2015**, *348*, 1261359. [CrossRef]
65. Harbeitner, R.C.; Wittmers, F.; Yung, C.C.; Eckmann, C.A.; Hehenberger, E.; Blum, M.; Needham, D.M.; Worden, A.Z. Gradients of bacteria in the oceanic water column reveal finely-resolved vertical distributions. *PLoS ONE* **2024**, *19*, e0298139. [CrossRef] [PubMed]
66. Yeh, Y.-C.; Fuhrman, J.A. Contrasting diversity patterns of prokaryotes and protists over time and depth at the San-Pedro Ocean Time series. *ISME Commun.* **2022**, *2*, 36. [CrossRef] [PubMed]

67. Schaechter, M. *Eukaryotic Microbes*; Academic Press: Cambridge, MA, USA, 2011.
68. Arrigo, K.R. Marine microorganisms and global nutrient cycles. *Nature* **2005**, *437*, 349–355. [CrossRef] [PubMed]
69. Austin, B. Bacterial pathogens of marine fish. In *Oceans and Health: Pathogens in the Marine Environment*; Springer: Berlin/Heidelberg, Germany, 2005; pp. 391–413.
70. Petersen, J.M.; Yuen, B. The symbiotic "all-rounders": Partnerships between marine animals and chemosynthetic nitrogen-fixing bacteria. *Appl. Environ. Microbiol.* **2021**, *87*, e02129–20. [CrossRef]
71. Jeong, H.J.; Kang, H.C.; Lim, A.S.; Jang, S.H.; Lee, K.; Lee, S.Y.; Ok, J.H.; You, J.H.; Kim, J.H.; Lee, K.H. Feeding diverse prey as an excellent strategy of mixotrophic dinoflagellates for global dominance. *Sci. Adv.* **2021**, *7*, eabe4214. [CrossRef]
72. Xu, K.; Choi, J.K.; Yang, E.J.; Lee, K.C.; Lei, Y. Biomonitoring of coastal pollution status using protozoan communities with a modified PFU method. *Mar. Pollut. Bull.* **2002**, *44*, 877–886. [CrossRef]
73. El-Tohamy, W.S.; Hopcroft, R.R. Planktonic ciliate communities along an environmental gradient in the Nile Delta (Damietta region, Egypt). *Sci. Rep.* **2024**, *14*, 19950. [CrossRef]
74. Vekeman, B.; Kerckhof, F.M.; Cremers, G.; De Vos, P.; Vandamme, P.; Boon, N.; Op den Camp, H.J.; Heylen, K. New Methyloceanibacter diversity from North Sea sediments includes methanotroph containing solely the soluble methane monooxygenase. *Environ. Microbiol.* **2016**, *18*, 4523–4536. [CrossRef]
75. Swan, B.K.; Martinez-Garcia, M.; Preston, C.M.; Sczyrba, A.; Woyke, T.; Lamy, D.; Reinthaler, T.; Poulton, N.J.; Masland, E.D.P.; Gomez, M.L. Potential for chemolithoautotrophy among ubiquitous bacteria lineages in the dark ocean. *Science* **2011**, *333*, 1296–1300. [CrossRef]
76. Malfertheiner, L.; Martínez-Pérez, C.; Zhao, Z.; Herndl, G.J.; Baltar, F. Phylogeny and metabolic potential of the candidate phylum SAR324. *Biology* **2022**, *11*, 599. [CrossRef]
77. Wasmund, K.; Schreiber, L.; Lloyd, K.G.; Petersen, D.G.; Schramm, A.; Stepanauskas, R.; Jørgensen, B.B.; Adrian, L. Genome sequencing of a single cell of the widely distributed marine subsurface Dehalococcoidia, phylum Chloroflexi. *ISME J.* **2014**, *8*, 383–397. [CrossRef] [PubMed]
78. Oborník, M.; Modrý, D.; Lukeš, M.; Černotíková-Stříbrná, E.; Cihlář, J.; Tesařová, M.; Kotabová, E.; Vancová, M.; Prášil, O.; Lukeš, J. Morphology, ultrastructure and life cycle of Vitrella brassicaformis n. sp., n. gen., a novel chromerid from the Great Barrier Reef. *Protist* **2012**, *163*, 306–323. [CrossRef] [PubMed]
79. Mohamed, A.R.; Cumbo, V.R.; Harii, S.; Shinzato, C.; Chan, C.X.; Ragan, M.A.; Satoh, N.; Ball, E.E.; Miller, D.J. Deciphering the nature of the coral–Chromera association. *ISME J.* **2018**, *12*, 776–790. [CrossRef] [PubMed]
80. Hohenegger, J. Functional shell geometry of symbiont-bearing benthic Foraminifera. *Galaxea J. Coral Reef Stud.* **2009**, *11*, 81–89. [CrossRef]
81. Hallock, P.; Lidz, B.H.; Cockey-Burkhard, E.M.; Donnelly, K.B. Foraminifera as bioindicators in coral reef assessment and monitoring: The FORAM index. *Environ. Monit. Assess.* **2003**, *81*, 221–238. [CrossRef]
82. Prazeres, M.; Martínez-Colón, M.; Hallock, P. Foraminifera as bioindicators of water quality: The FoRAM Index revisited. *Environ. Pollut.* **2020**, *257*, 113612. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.