

# Systematic review of cnidarian microbiomes reveals insights into the structure, specificity, and fidelity of marine associations

Received: 6 September 2022

Accepted: 30 June 2023

Published online: 14 August 2023

 Check for updatesM. McCauley<sup>1,2,3</sup>✉, T. L. Goulet<sup>2</sup>, C. R. Jackson<sup>2</sup> & S. Loesgen<sup>1</sup>

Microorganisms play essential roles in the health and resilience of cnidarians. Understanding the factors influencing cnidarian microbiomes requires cross study comparisons, yet the plethora of protocols used hampers dataset integration. We unify 16S rRNA gene sequences from cnidarian microbiome studies under a single analysis pipeline. We reprocess 12,010 cnidarian microbiome samples from 186 studies, alongside 3,388 poriferan, 370 sea-water samples, and 245 cultured Symbiodiniaceae, unifying ~6.5 billion sequence reads. Samples are partitioned by hypervariable region and sequencing platform to reduce sequencing variability. This systematic review uncovers an incredible diversity of 86 archaeal and bacterial phyla associated with Cnidaria, and highlights key bacteria hosted across host sub-phylum, depth, and microhabitat. Shallow (< 30 m) water Alcyonacea and Actinaria are characterized by highly shared and relatively abundant microbial communities, unlike Scleractinia and most deeper cnidarians. Utilizing the V4 region, we find that cnidarian microbial composition, richness, diversity, and structure are primarily influenced by host phylogeny, sampling depth, and ocean body, followed by microhabitat and sampling date. We identify host and geographical generalist and specific *Endozoicomonas* clades within Cnidaria and Porifera. This systematic review forms a framework for understanding factors governing cnidarian microbiomes and creates a baseline for assessing stress associated dysbiosis.

Cnidarians (e.g., corals, sea anemones, jellyfish) face anthropogenic stressors that are increasing in severity and frequency. Globally, the ecosystem services that corals and other cnidarians provide have fallen by approximately 50%<sup>1</sup>, with further decreases forecasted in the future<sup>2</sup>. Since many organisms reside within and around cnidarians, loss of cnidarian species has resulted in a 63% decline in coral reef associated biodiversity, including macroinvertebrates and fish<sup>1</sup>.

In addition to macroorganisms, cnidarians host a multitude of microorganisms, which together constitute their microbiome<sup>3–6</sup>. Research on the mutualistic symbiosis between cnidarians and

dinoflagellate algae (family Symbiodiniaceae), including the bleaching phenomenon (loss of Symbiodiniaceae and/or their chlorophyll content), overshadows the knowledge on the Archaea, Bacteria, Fungi, microalgae, protists, and viruses of the cnidarian microbiome<sup>7–11</sup>. Nevertheless, the non-Symbiodiniaceae members of cnidarian microbiomes play important roles within the holobiont<sup>12</sup> (host and microbiota). To understand cnidarians under current climate conditions, and cnidarian capacity to withstand environmental change, it is imperative to understand the fundamental factors influencing the microbiome composition of cnidarians.

<sup>1</sup>Department of Chemistry, Whitney Laboratory for Marine Bioscience, University of Florida, St. Augustine, FL, USA. <sup>2</sup>Department of Biology, University of Mississippi, University, MS, USA. <sup>3</sup>Present address: U.S. Geological Survey, Wetland and Aquatic Research Centre, Gainesville, FL, USA.

✉ e-mail: [mrkmccauley@gmail.com](mailto:mrkmccauley@gmail.com)

The cnidarian microbiome functions in host nutrient cycling (C, N, P, S) (reviewed in<sup>13–16</sup>), homeostasis (reviewed in<sup>17,18</sup>), protection, including antimicrobial production and competitive exclusion (reviewed in<sup>19–21</sup>), development (reviewed within<sup>22</sup>), health (reviewed in<sup>23,24</sup>), and response to environmental fluctuations (reviewed in<sup>25,26</sup>), in addition to contributing to reef processes (reviewed in<sup>27,28</sup>), and ecosystem resilience (reviewed in<sup>16,29</sup>). These findings were made possible with the increasing access to DNA sequencing and prompted by the increased concern for coral reef health<sup>30–32</sup>.

The majority of cnidarian microbiome research has utilized 16S rRNA amplicon sequencing of scleractinian coral microbiomes to identify primarily bacterial assemblages. As the main frame builders of coral reefs, scleractinian corals (class Hexacorallia, sub-phylum Anthozoa<sup>33</sup>), have been the main emphasis of cnidarian microbiome research. This focus has framed our current understanding of cnidarian microbiomes. Research into non-scleractinian cnidarians, however, has identified significantly different structuring of cnidarian microbial symbioses<sup>34–36</sup>. Unfortunately, the lack of standardized sampling, storage, processing, and analytical protocols across 16S rRNA gene amplicon studies, hinders the ability to synthesize published data to create a baseline for cnidarian microbiomes<sup>37–39</sup>. Thus, the scientific community often relies on literature reviews in lieu of meta-analyses or systematic reviews<sup>37,40,41</sup>. The methodological variability not only challenges data integration but can lead to different conclusions due to identifying significantly different microbial communities from the same coral sample<sup>42</sup>.

In addition to variability resulting from different processing protocols, the majority of cnidarian datasets published to date have been processed by clustering (typically to 97% similarity) their sequences into operational taxonomic units (OTUs), with the remaining datasets denoised to produce amplicon sequence variants (ASVs). Clustering within datasets limits detailed examination of trends across available literature, as OTUs from one study are inherently different from those from another. Further, while both OTU and ASV approaches have their strengths and weaknesses<sup>43,44</sup>, the denoising steps in ASV pipelines can increase the precision of measuring environmental specificity<sup>45</sup> and, importantly, allow for direct comparison of sequences across studies.

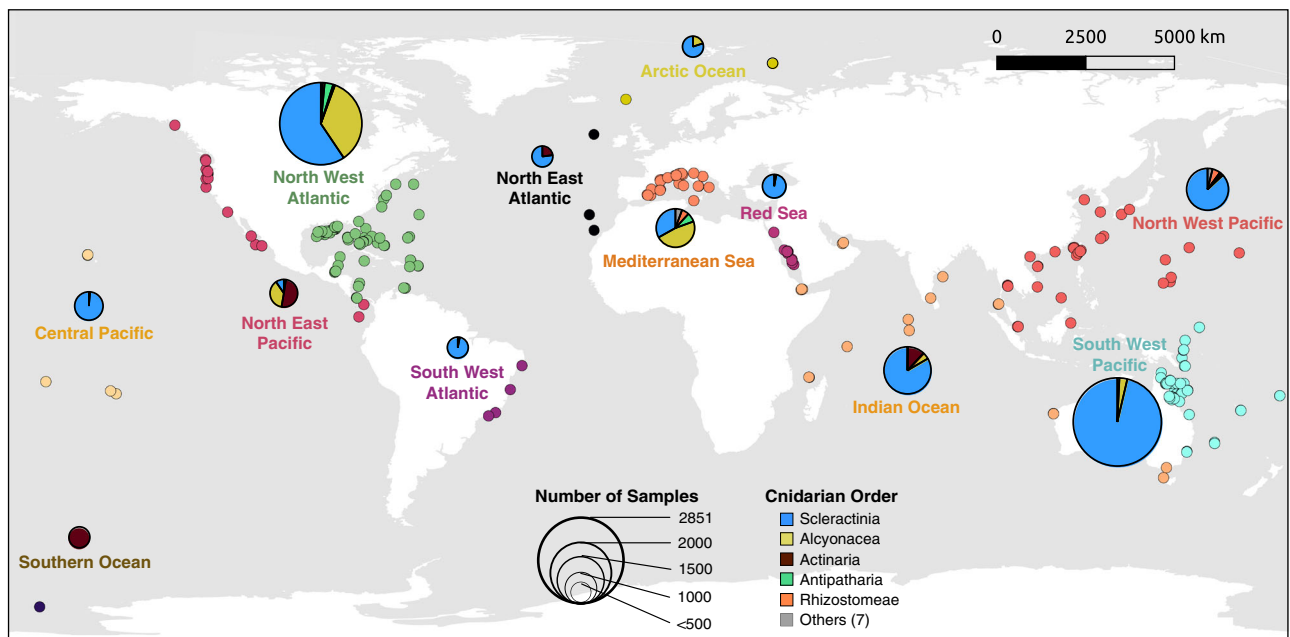
However, the production of ASVs across different denoising pipelines, or different parameters, can also introduce sequence variation.

In this study we unify available cnidarian 16S rRNA sequences (Fig. 1, Supplementary Data 1) under a single analysis pipeline, thereby creating a global baseline for cnidarian microbiomes, and elucidating patterns and factors that govern the assemblage of cnidarian microbiomes. Importantly, this database is freely available (Figshare) to serve as a community resource for further exploration into cnidarian microbiomes, and to support global conservation strategies. We assemble all available 16S rRNA sequences from nearly 200 studies, each of which preserved, extracted, amplified, processed, and analyzed samples using distinct techniques. Partitioning these samples according to host health, sampling effort, hypervariable region, and sequencing platforms, we reduce the methodological and sequencing variability present for cross-study analyses. We compare the diversity, structure, and richness of scleractinian coral prokaryotic microbial communities to other members of the phylum Cnidaria, to determine how phylogeny, geography, and depth influence cnidarian microbiomes. We investigate the relative abundance of microbiota as a function of depth, and identify the taxonomic ‘core microbiome’ (microbiota present in all samples within a species per site), as this microbial component often serves as a proxy for holobiont health and resilience when facing anthropogenic stressors<sup>20,46</sup>. Further, we examine differences in the microbial relative abundance in three different coral microhabitats (tissue, skeleton, and mucus) by using unique sequence and clustering-based techniques. Finally, we document microbial genera (including *Endozoicomonas*) that are highly prevalent and abundant within the different microhabitats of cnidarian species, serving as core members across cnidarian classes.

## Results

### Available 16S rRNA sequences

A total of 12,010 cnidarian 16S rRNA sequences were processed, alongside 3388 poriferan, 370 seawater, and 245 cultured Symbiodiniaceae, to produce a final database of 16,012 samples. This database was partitioned according to sample metadata, including host health status, sampling effort per site, amplified hypervariable region, and



**Fig. 1 | Global distribution of the 12,010 cnidarian microbial (16S rRNA gene) samples included in the systematic review.** Regional pie chart diameters indicate sample abundance, with colors representing the relative abundance of five

cnidarian orders and an ‘other’ category which includes seven additional cnidarian orders. Smaller circles represent studies with at least one cnidarian microbiome sample, with colors specific to water bodies. Made with QGIS.

**Table 1 | The number of studies, species, samples, and amplicon sequence variants (ASVs) included in the various datasets analyzed in the systematic review**

	Studies	Species	Samples	ASVs
Entire Database	186	495	16,012	
-Cnidaria	177	212	12,010	
--V1-V2 Library	19	46	1949	7934
--V3 Library	43	102	2106	212,316
--V4 Library	85	233	5979	360,006
--V4_p Library	86	467	9737	493,157
--V5-V6 Library	44	52	3751	128,686
--V7-V8 Library	9	13	469	11,893
---V4_MiSeq_healthy_cnidarian_unrarified	63	196	4243	167,227
---V4_MiSeq_healthy_cnidarian_rarified	49	150	2964	131,761
---V4_MiSeq_healthy_scleractinian_unrarified	53	149	3546	163,004
---V4_MiSeq_healthy_scleractinian_rarified	34	103	2275	127,384
---V4_MiSeq_healthy_nonscleractinian_unrarified	24	47	697	31,550
---V4_MiSeq_healthy_nonscleractinian_rarified	18	37	495	24,400
----Relative Abundance Analyses	143	131	7067	588,776

sequencing chemistry (Table 1). Primarily, this included six ASV libraries that span the amplified hypervariable regions: V1-V2, V3, V4, V5-V6, V7-V8, alongside overlapping V4 cnidarian and poriferan samples (V4\_p; Table 1). The cnidarian V4 library was further partitioned by sequencing chemistry and cnidarian classes, to reduce the effects of the MiSeq processing and scleractinian sampling bias. Additionally, the relative abundances of microbial communities were investigated within cnidarian tissue, skeletal, and mucus samples across all hypervariable regions, whereby a minimum of four ‘healthy’ (no record of observed stress) cnidarian individuals were sampled per site (Table 1).

### Taxonomic variability in cnidarian 16S rRNA sequences

Across the phylum Cnidaria, 12 archaeal and 74 bacterial phyla (22 and 185 classes, respectively) were identified from 720,835 ASVs present in libraries of five regions of 16S rRNA cnidarian genes (V1-V2, V3, V4, V5-V6, V7-V8). Out of 697,284 ASVs classified as bacterial, the family Endozoicomonadaceae (Gammaproteobacteria) exhibited the greatest prevalence in each of the four cnidarian classes: Anthozoa, Cubozoa, Hydrozoa, and Scyphozoa. Vibrionaceae (Gammaproteobacteria) was prevalent in three of the four cnidarian classes, with the exception of Hydrozoa (visualized with the V4 library, Fig. 2). Both of these bacterial families also presented the highest average relative abundance in the microbial samples that they were present in. Strikingly, the vast majority of cnidarian associated microbiota were rare (<1% prevalence) and contributed little to the relative abundance of their host microbiome (visualized with the V4 library, Fig. 2).

Proteobacteria (currently ascribed as Pseudomonadota) was the dominant bacterial phyla (40.3%), followed by Firmicutes (currently ascribed as Bacillota) constituting 16.1%, and Bacteroidetes (currently ascribed as Bacteroidota) at 8.3% (Fig. 3c). Archaea were dominated by Nanoarchaeota (59.6%), followed by Crenarchaeota (22.6%).

### The diversity of cnidarian microbiomes ( $\alpha$ -diversity)

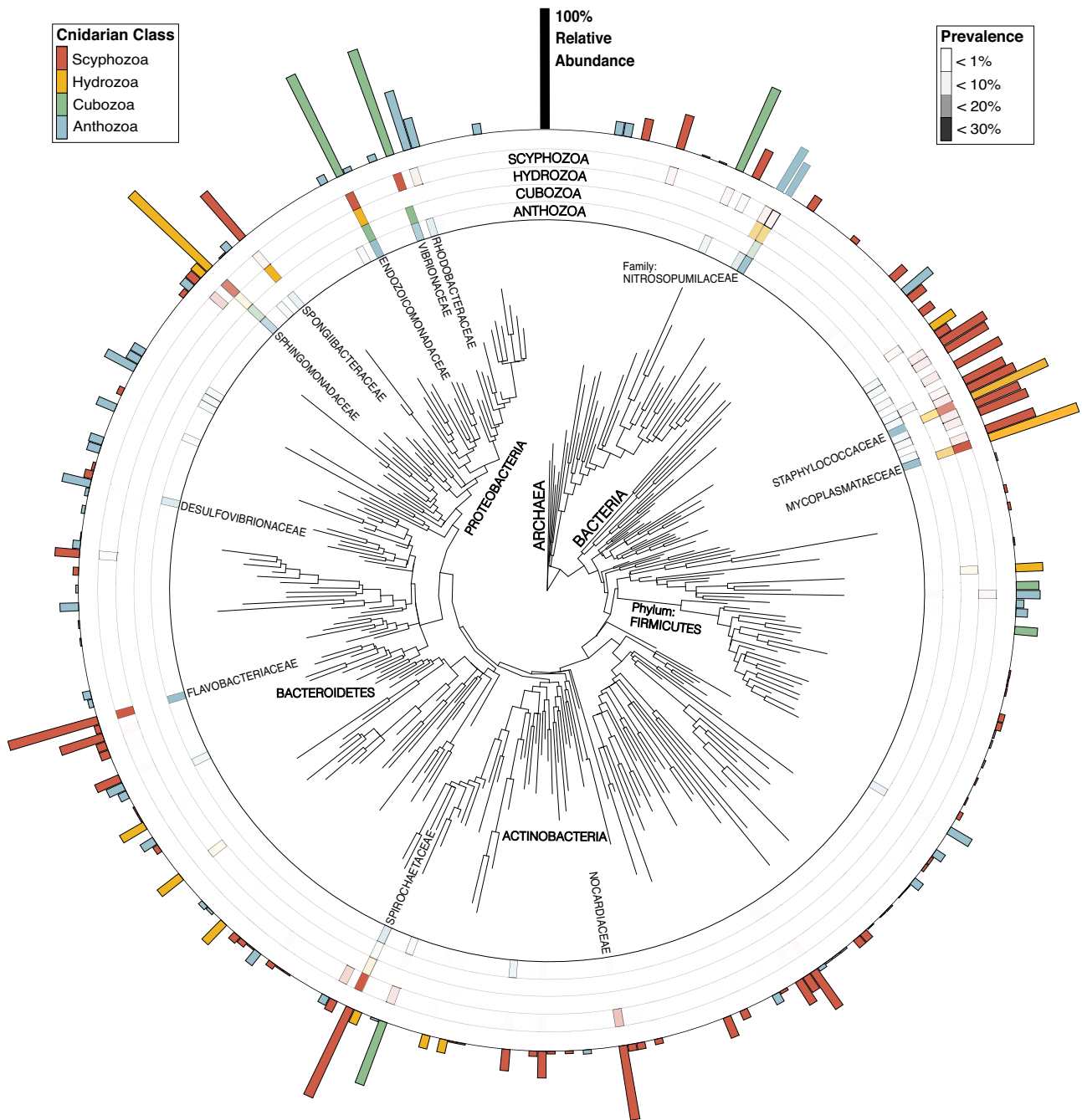
For the subset of ‘healthy’ Illumina MiSeq samples (V4\_MiSeq\_healthy\_cnidarian library), microbial diversity was significantly correlated with cnidarian phylogeny, life stage, hosting Symbiodiniaceae, as well as collection depth, and location (observed ASVs, Shannon Diversity and Inverse Simpson index; Supplementary Tables 1–2). Consistent across unrarified and rarified data sets, sampling water body was often the strongest influence, explaining up to 43% of the microbial diversity within scleractinian corals, and up to 33% in non-scleractinians (Supplementary Tables 1–2).

With and without rarefaction, class Hexacorallia contained a significantly greater microbial diversity than Discomedusae, Hydrozoa, and Octocorallia ( $p$  adj. < 0.02, Supplementary Fig. 1). Cnidarian orders were separated into three groups based on the diversity indices of microbial ASVs identified within the Illumina MiSeq V4 library. The first, representing only Scleractinia, yielded the greatest average (894) and maximum (3794) number of ASVs per sample. Scleractinia also hosted the greatest range of microbial richness, Shannon and Inverse Simpson index values, although this is partially a result of sampling bias (Fig. 3a, b, Supplementary Figs. 2–3). The second group contained both Alcyonacea and Antipatharia, hosting an average of 427 ASVs per sample and a maximum of 2000 ASVs, with index values generally higher than the remaining seven cnidarian orders (Actiniaria, Anthoathecata, Corallimorpharia, Helioporacea, Rhizostomeae, Semaestomeae, and Zoantharia), although these differences were not as consistently significant as observed with Scleractinia (Fig. 3a, Supplementary Fig. 2). The third group consisted of the seven least diverse remaining cnidarian orders, with an average of 103 ASVs per sample.

Increased microbial diversity was significantly correlated with shallower sampling sites in alcyonaceans, scleractinians, and zoanthids, (98.3%, 82.7%, 100% depth metadata available per samples, respectively) whereas the opposite trend was present in Antipatharia (99.6% depth metadata available,  $p$  adj. < 0.02, Fig. 3b, Supplementary Figs. 4–7). Scleractinians from the ‘complex’ lineage (often creating more porous skeletal structures)<sup>47</sup> hosted a significantly ( $p$  adj. < 0.02) more diverse microbiome in comparison to species of the ‘robust’ lineage (often creating more solid skeletal structures) (Supplementary Fig. 8). A trend toward a greater microbial diversity in scleractinian mucus and skeleton compared to tissue was also identified, suggesting a highly species-specific relationship ( $p$  = 0.09). The microhabitat trend, however, did not occur in non-scleractinian cnidarians, potentially indicating a more niche dependent microbiome in scleractinian species (Supplementary Tables 1–2).

### The structure of cnidarian microbiomes ( $\beta$ -diversity)

Dispersion of microbial dissimilarities within the V4\_MiSeq\_healthy\_cnidarian library was significantly different between scleractinian and non-scleractinian cnidarians, which may be partially due to the oversampling of the former ( $p$  <  $1 \times 10^{43}$ , Supplementary Tables 3–6). Subsequently, scleractinian corals were separated from non-scleractinian cnidarians, and all diversity metrics were repeated on the separated data sets (Supplementary Tables 1–2). All tested parameters including, host phylogeny, life-stage, collection depth,



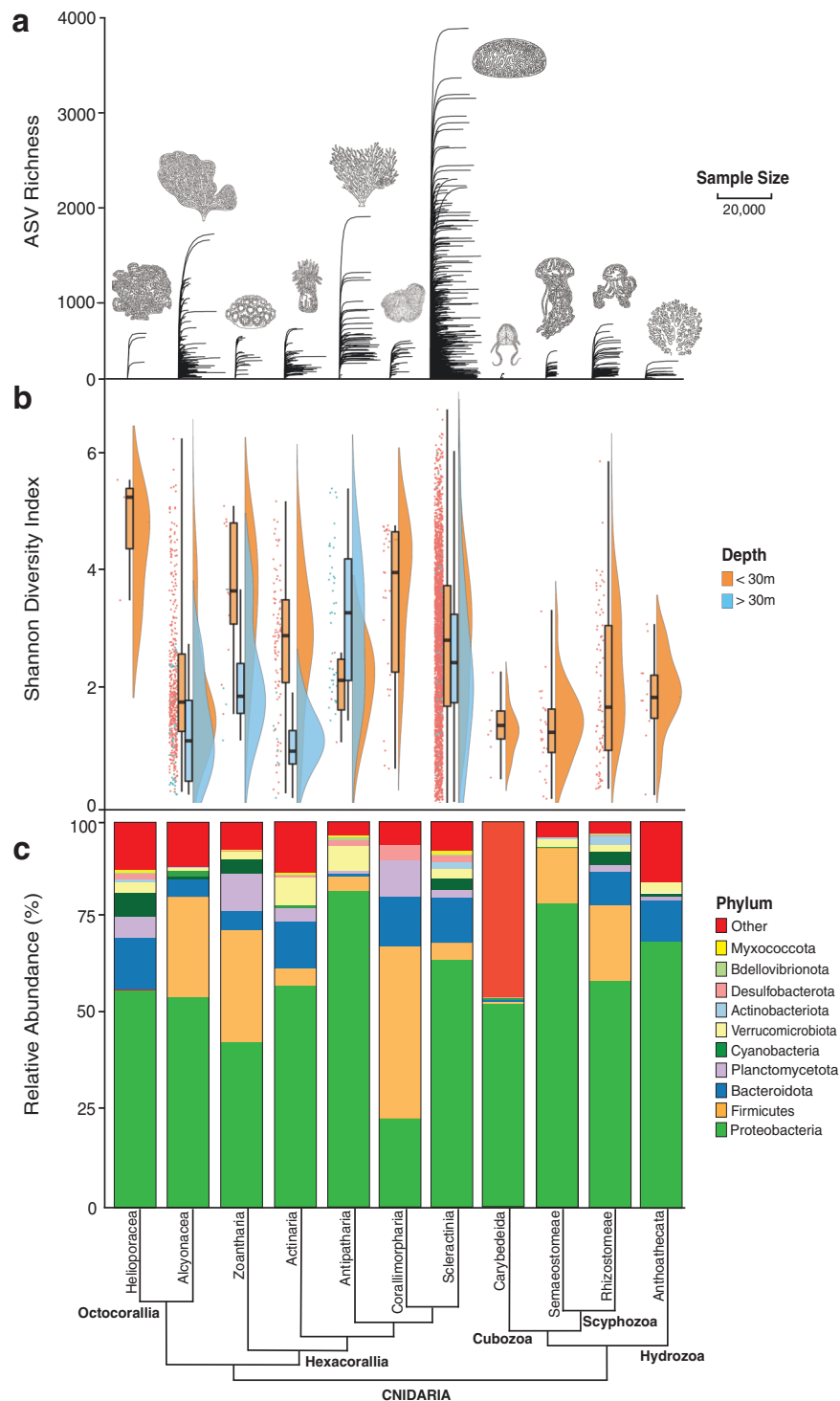
**Fig. 2 | Phylogeny, prevalence, and relative abundance of the most prevalent archaeal and bacterial families identified in the V4 library.** The heat maps represent the prevalence (% of total samples detected in) of different microbial

families across the cnidarian classes, while the outer bar charts display the cnidarian class with the greatest average relative abundance. The most prevalent and relatively abundant microbial families are annotated.

hosting Symbiodiniaceae, original water body, and DNA extraction method were significantly correlated to microbial structure (Bray-Curtis, and Jaccard dissimilarity indices) of scleractinian and non-scleractinian cnidarians ( $p < 0.004$ , Supplementary Tables 1–2). The strongest influencers of both the unrarefied and rarefied Illumina MiSeq subset, were phylogenetic history, followed by original water body, although the strengths of their correlations varied between scleractinian and non-scleractinian data subsets (Supplementary Figs. 9–10, Supplementary Tables 1–2).

By hierarchically clustering the Bray-Curtis dissimilarity index values of the V4\_MiSeq\_healthy\_cnidarian library, species within Octocorallia, Scyphozoa, and Hydrozoa formed numerous clusters that were often geographically distinct, similar to tight clusters of

species within Actinaria, Antipatharia, Corallimorpharia, and Zoantharia. Strong clustering of scleractinian microbiome samples within family and water bodies occurred alongside geographically closer sampling sites also clustered closer together (Supplementary Fig. 11). These geographical signals were observed again within the V4\_p library (combining cnidarian and poriferan samples), with scleractinian microbiome samples spanning almost the entire dissimilarity distance, with numerous small groupings of robust corals, within the wider array of complex species (Supplementary Fig. 12). The majority of high and low microbial abundance sponges represented the extremes of the dissimilarity dendrogram (Supplementary Fig. 12). While some scleractinians and actinarians hosted microbiomes similar to those of low microbial abundance sponges, the majority of cnidarian microbiomes



**Fig. 3 | Microbial diversity and abundance in the phylum Cnidaria.** **a** Rarefaction plot of amplicon sequence variants (ASV) richness (ASVs observed) illustrating the high sequencing coverage across the 11 cnidarian orders within the V4 library. Sequence depth was trimmed to 20,000 for visual purposes only. **b** Sample Shannon index values (dots), averages (bar plot), and distributions (half violin) are

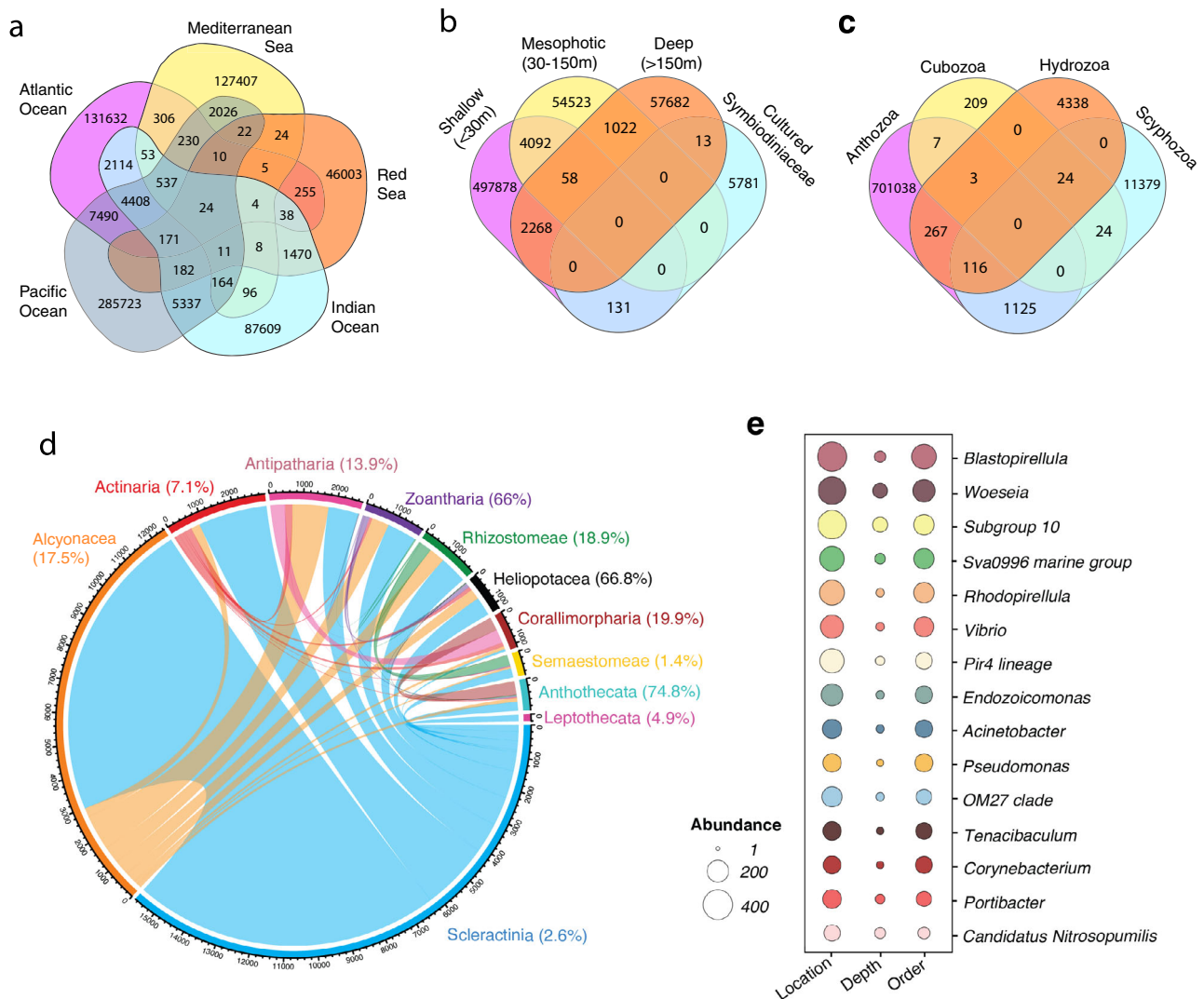
colored by depth per cnidarian order ( $n = 3840$  samples). Box plot whiskers represent the minimum and maximum values, with bounds of the box representing the 25th and 75th percentile, and the median in the center. **c** The relative abundance of the 10 most abundant bacterial phyla per cnidarian order. Cnidarian orders are arranged phylogenetically<sup>119,120</sup>.

fell in-between these sponge types or were more similar to high microbial abundance sponges (Supplementary Fig. 12).

### The specificity of cnidarian microbiomes

Of the 697,284 bacterial ASVs present in the five cnidarian hypervariable libraries (Table 1), fewer than 10% were shared across cnidarian

orders, with the majority (71.1%) identified within Scleractinia, followed by 7.6% identified in Alcyonacea and 5.3% to Actinaria, regardless of geographic location (Table 1; Supplementary Fig. 13). Nevertheless, cnidarian microbiomes were structured strongly across water bodies, with 93.6% of all ASVs identified in a single water body, and less than 3.6% shared across more than two water bodies (Fig. 4a). There was



**Fig. 4 | Distribution of shared microbial sequences within the phylum Cnidaria.** Absolute abundance of archaeal and bacterial amplicon sequence variants (ASVs) identified in, and shared across, **a** major water bodies, **b** sampling depth, **c** cnidarian classes. **d** ASVs identified in at least one cnidarian order, with arc width representing the absolute abundance of shared ASVs. The relative abundance of

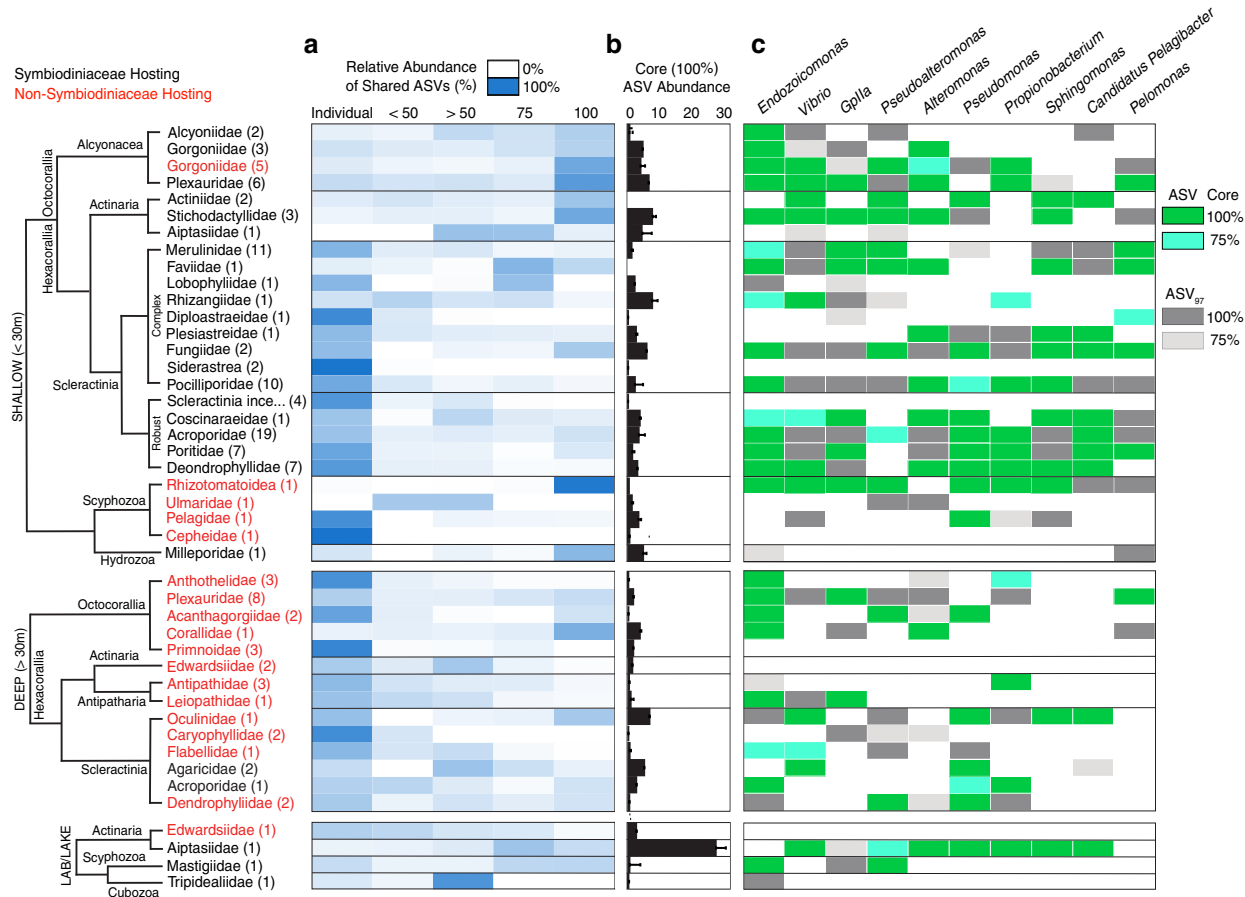
shared sequences (within all sequences identified in that order) is presented as a %. **e** The absolute abundances of the 15 most shared bacterial genera, with ASVs identified across major water bodies (location), depth, and cnidarian order, representing the most globally distributed cnidarian-associated bacteria.

also intense structuring with depth, with just 1% of ASVs shared across depths (Fig. 4b). Of the ASVs identified in shallow waters (<30 m), 98.7% were found only at that depth range, followed by 90.1% of those identified in mesophotic depths (30–150 m) unique to that range, and 94.4% of those identified in deeper waters (>150 m) (Fig. 4b). Further, only 2.6% of ASVs identified in cultured Symbiodiniaceae microbiomes were identified in any cnidarian microbiome, and primarily from those sampled in shallow waters (V3, V4 libraries; Fig. 4b). While Scyphozoa shared the greatest number of sequences with Anthozoa when compared to Cubozoa and Hydrozoa, they constituted less than 0.001% of all ASVs (Fig. 4c).

Over two thirds (67.3%) of the sequences shared across cnidarian families were sequenced from Scleractinia and Alcyonacea (Fig. 4d). While Scleractinia hosted 35.6% of all shared sequences, they accounted for only 2.6% of microbes exclusive to Scleractinia, as opposed to the 31.7% of shared sequences found in Alcyonacea accounting for 17.5% of their entire microbiome. The absolute abundance of shared sequences was not solely a result of biased sampling effort, as a similar number of Actinaria and Alcyonacea sequences were included in this study, with Actinaria hosting fewer than 10% of shared

sequences. The relative abundance of shared sequences varied widely across cnidarian orders from Semaestomeae sharing the least (1.4% of their entire microbiome), to the Anthothecata sharing the most (74.8% of their entire microbiome), even with a similar number of samples in the library. Alcyonaceans shared more sequences with cnidarians outside of their order (17.5%) than within (10.2%), whereas the Scleractinia shared more sequences within their order (12.3%) than outside of it (2.6%).

Across all scleractinian microbiomes, 86.1% of sequences were exclusive to cnidarian tissue samples, 6.7% to mucus, and 5.4% to skeleton, with 1.8% found in other tissue types (Supplementary Data 1). Only 2.4% of the scleractinian microbiome was shared between tissue and mucus samples, which was greater than the 1.2% shared between the skeleton and tissue. Relatively few ASVs were identified in both skeleton and mucus samples (0.2%), while even fewer were identified in tissue, mucus, and skeleton samples (0.1%). ASVs were more likely to be shared between the tissue and mucus from non-Hexacorallia and Hexacorallia samples (0.2% and 0.1%, respectively), than between skeletons (<0.01%). Surprisingly, the identities of shared bacterial genera across cnidarian orders were also those most likely to be shared



**Fig. 5 | Abundance and composition of the cnidarian core microbiome across depth.** **a** The relative abundance of archaeal and bacterial sequences identified within tissue samples per site, averaged to cnidarian family. The number of species present within each family is included in parenthesis. Columns represent the relative abundance of Amplicon Sequence Variants (ASVs) present within a single (Individual), fewer than 50% (<50), more than 50% (>50), at least 75% (75), or present in all (100) cnidarians sampled per site. **b** The average absolute abundance of

ASVs identified in 100% of individuals sampled per site, with standard error. **c** The identity of the 10 most commonly present bacteria in 75 and 100% of individuals sampled per site, as an ASV or an ASV<sub>97</sub> (clustered at 97% similarity). Colored boxes indicate if bacteria were identified as a core member in at least one site per species. Cnidarian species were only included if there was a minimum of four 'healthy' individuals per site.

across depth and geography (Fig. 4e). Because the data are inherently dependent on the techniques used to partition samples from the different microhabitats, many previous studies have inadvertently mixed these location specific microbes during processing. Given the distinct communities present within microhabitats within individual cnidarians, future studies investigating the functional role of these microbiota should utilize methods that minimize overlap in microorganisms derived from the mucus, tissue, and skeleton microhabitats of the coral holobiont<sup>45</sup>.

### Core microbiota of 'healthy' cnidarians

To examine the relative distribution of microbes within Cnidaria, an additional dataset was created from all studies in the original Cnidaria database, specifically including all 'healthy' cnidarians, for which at least four individuals were collected per sampling site (Table 1). Cnidarian order and sampling depth were significantly correlated with the proportion of bacteria present in all cnidarian tissue samples at a site ( $p < 0.001$ ,  $p < 0.05$ , respectively) and those identified only in a single cnidarian sample per site ( $p < 0.005$ ,  $p < 0.05$ , respectively), with deeper cnidarians hosting a more individualized microbiome, whereas shallow water Cnidaria (primarily Alcyonacea and Actinaria) generally hosting a more abundant core

microbiome (Fig. 5). Hosting Symbiodiniaceae significantly influenced the proportion of the microbiota identified in all cnidarians sampled per site ( $p < 0.05$ ), but not that identified in a single cnidarian ( $p > 0.05$ ).

Comparing the average relative abundances of both the clustered and unique sequence analyses revealed cnidarian and microhabitat variability (Supplementary Figs. 14–16). Use of clustered sequences instead of unique sequences, when examining colony-specific microbiota in scleractinian tissue and skeletal samples, resulted in a 61.2% and 63.8% increase in their average relative abundances. A similar comparison with scleractinian mucus samples resulted in only a 45.4% increase (Supplementary Figs. 14–15). Further, using clustered instead of unique sequences yielded a 55% increase in the relative abundance of what constituted a 'core microbiota' in all samples for scleractinian Acroporidae tissues, while clustering resulted in a 21% reduction of the 'core microbiota' in the alcyonacean Plexauridae (Supplementary Fig. 16). The use of clustered sequences may therefore be problematic when comparing microbial communities of different cnidarians or different microhabitats.

Within their tissue microhabitat, scleractinian coral colonies hosted a significantly greater proportion of microbes (ASVs) identified in a single colony compared to shallow water (<30 m) actinarians,

alcyonaceans, and hydrozoans, ( $p < 0.05$ , for all, Fig. 5a). However, there was a significant shift to a greater proportion of microbiota identified only in a single cnidarian sample with actinarian and alcyonaceans sampled deeper than 30 m ( $p < 0.05$ , Fig. 5a). All ambient collected cnidarian samples hosted an average of <10 core (100%) ASVs, although this varied significantly within species sampled at different sites ( $p < 0.05$ , Fig. 5b). The major exception to this were cultured sea anemones, *Exaiptasia pallida*, which hosted on average six times the number of core ASVs of field-collected *E. pallida*.

Core microbiota within mucus samples presented a similar pattern, with the majority of scleractinians hosting a highly colony-specific community, while concurrently hosting a group of commonly shared microbial genera (Supplementary Fig. 14). Of all microhabitats however, skeletal microbiomes demonstrated not only the greatest relative abundance of ASVs present within a single scleractinian colony (68.7%, compared to 48.6% in mucus, and 53.0% in tissues), they also shared few core microbiota across species (Supplementary Fig. 15).

The most commonly occurring microbiota within cnidarian tissues, including Gp11a, *Vibrio*, *Pseudoalteromonas*, *Alteromonas*, *Pseudomonas*, and *Propionibacterium* were generally present in the core microbiome of most cnidarian orders across depth, while *Sphingomonas*, *Pelomonas*, and *Candidatus Pelagibacter* were more commonly abundant in shallow waters (<30 m) (Fig. 5c). Of the 143 studies that examined the relative abundance of microbiota, only 19 included either PCR, and/or extraction kit controls. Nevertheless, after we excluded bacterial sequences identified as potential contaminants we still detected core communities containing *Alteromonas*, *Pseudoalteromonas*, and *Propionibacterium*, suggesting that these taxa are not kit contaminants<sup>49,50</sup>

*Endozoicomonas* was the most commonly identified core bacterial genus across all cnidarians, including shallow water non-Symbiodiniaceae hosting cnidarian species, and cnidarians living deeper than 30 m (Fig. 5c). Additionally, while Gp11a, *Candidatus Pelagibacter*, *Vibrio*, and *Pseudoalteromonas* were identified as common core microbiota across both tissue and mucus samples, *Endozoicomonas* was the only bacterial genus that commonly occurred across tissue, mucus, and skeletal microhabitats (Fig. 5c, Supplementary Figs. 14–16).

### Endozoicomonas case study

To further examine the bacterial genus that was the most prevalent across cnidaria, globally ubiquitous, and also the most commonly identified in the core microbiome, we examined *Endozoicomonas* ASVs within all cnidarian libraries (VI-V2, V3, V4, V5-V6, and V7-V8), as well as within both Cnidaria and Porifera (V4\_p; Table 1). Sequence variation in these ASVs suggested 15 possible major clades of this genus, that revealed host and geographical distribution (Fig. 6, Supplementary Fig. 17).

Major clades were generally identified across all libraries, with a few minor clades expanding and collapsing depending on sampling effort. Neither major nor minor clades were singularly attributable to sequencing platform, abundance available reads or extraction techniques. 98.6% of *Endozoicomonas* ASVs identified in Porifera and non-anthozoan cnidarians were present in only a few closely related major clades, whereas all major clades included *Endozoicomonas* ASVs hosted within Anthozoans (Fig. 6a). Within Cnidaria, *Endozoicomonas* ASVs were identified in eight orders (Actinaria, Alcyonacea, Antipatharia, Corallimorpharia, Helioporacea, Rhizostomeae, Scleractinia, Zoantharia), including many non-Symbiodiniaceae hosting species (Fig. 6b). Further, non-scleractinian *Endozoicomonas* ASVs were often closely related and displayed distinct geographical separation (Fig. 6c).

Of the 4138 *Endozoicomonas* ASVs present in all cnidarian libraries, the majority 74.3% were identified in Scleractinia, 15.8% identified in Alcyonacea, 3.8% to Actinaria, with 3.5% identified in the remaining orders and only 2.6% shared across orders. Of all *Endozoicomonas* ASVs

identified within scleractinian corals, 88.5% were exclusively identified in tissue samples, with 5.5% found only in mucus samples, 2.1% in skeletal samples, while 3.9% were shared across body compartments.

Within the V4 library specifically, 1,489 *Endozoicomonas* ASVs were present with 58.4% identified only in a single cnidarian sample, and the majority of minor clades were highly specific to either water body or host family (Fig. 6c). There was a range of host specificity, with some highly specific clades primarily amplified only from specific cnidarian families, including Clade 2 identified only in Indian Ocean Acroporidae colonies, Clade 13 in Pacific Pocilloporidae colonies, and Clade 14 primarily identified in Atlantic Rhizostomeae (Fig. 6b, c). On the other hand, highly generalist clades including Clades 5, 10, and 12 were identified in globally distributed Cnidaria and Porifera. Further, these host generalist clades were more likely to be present across multiple tissue types, compared to the most specialist clades which were generally identified primarily in scleractinian tissue. Clade 11 was highly distinctive, having the greatest average abundance within samples, as well as being identified in the greatest number of cnidarian samples (Supplementary Figs. 18–19). Clade 11 was also a geographic generalist, found globally, and a host specialist, being detected primarily in Octocorallia (Fig. 6c, Supplementary Figs. 18–19).

### Nonbacterial members of cnidarian microbiota

Fewer than 5% of the 720,835 ASVs identified across the five hypervariable libraries were classified as non-bacterial (Supplementary Fig. 13), this is likely because the majority of studies analyzed in our data set aimed to specifically amplify bacteria, and nonbacterial members of the cnidarian microbiota are underreported. Host cnidarian sequences were common, with 47.2% of these identified using the modified Silva dataset and the remaining through BLAST searches. Numerous highly ubiquitous non-bacterial symbionts were more prevalent than previously documented, including Apicomplexan Related Lineages V (ARL-V), that were present in seven cnidarian orders globally, and *Ostreobium* present not only in many scleractinian corals and octocorals, but also in black corals, Antipatharia. A total of 11,995 ASVs (1.5%) were obtained from cnidarians that could not be identified across all hypervariable libraries, even after BLAST searches (Supplementary Fig. 13).

## Discussion

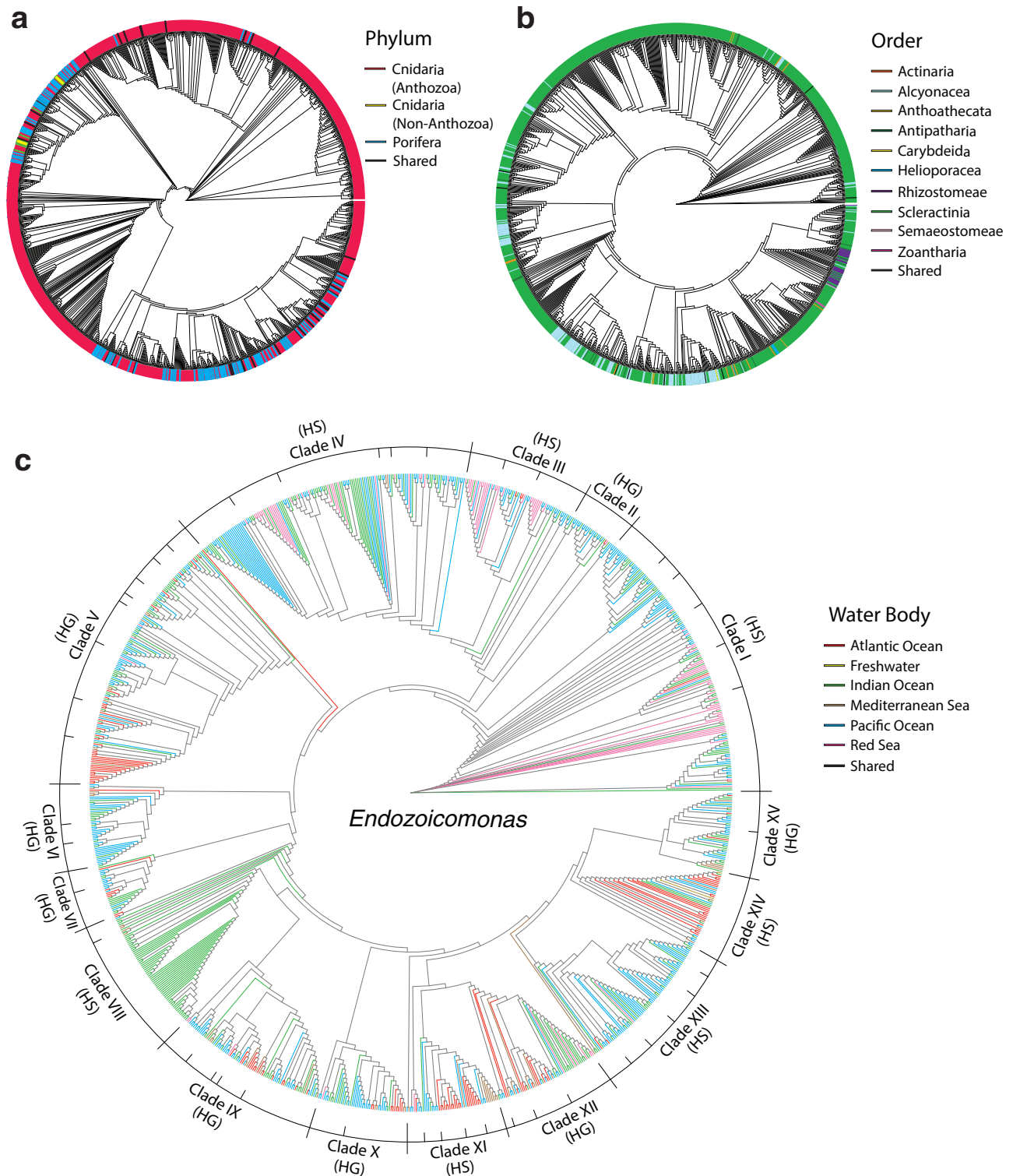
### Cnidarian microbiota diversity

This systematic review revealed that cnidarians host an extraordinary array of microbial phyla (12 archaeal and 74 bacterial phyla, 22 and 185 classes, respectively, Fig. 2). This substantially expands previous totals of the 42 microbial phyla identified in the Coral Microbial Database and the Coral Microbiome Portal<sup>41</sup>, and the 63 bacterial phyla identified in the Sponge Global microbiome dataset<sup>51</sup>. Remarkably, cnidarian microbial phyla diversity also rivals the 85 phyla identified in the Earth Microbiome Project, which examined 2.2 billion free-living and host-associated 16S rRNA sequences<sup>45</sup>. This incredible diversity suggests that cnidarians are one of Earth's most diverse microbial reservoirs, supporting substantial prokaryotic metacommunities.

### Shared Cnidarian microbiota

The integration of microbiome data across Cnidaria, which we performed in this systematic review, is critical for a greater understanding of cnidarian-microbiome structure and plasticity, as well as prokaryotic interactions with other coral reef invertebrates. Establishing a microbial baseline will also allow for the assessment of stressor-associated dysbiosis. For example, even with the current sampling bias of scleractinian corals, Alcyonacea and Scleractinia each hosted approximately 30% of all shared cnidarian microbiota, suggesting that both orders may act as reservoirs for other cnidarian species (Fig. 4d). Conversely, these shared sequences represented <3% of the entire scleractinian microbiome, of which 12.5% was shared within the order, suggesting that scleractinian corals might not only serve as an important reservoir for scleractinians,





**Fig. 6 | Phylogenies of *Endozoicomonas* sequences isolated from the phyla Cnidaria and Porifera. a** *Endozoicomonas* phylogeny of amplicon sequence variants (ASVs) identified within Cnidaria and Porifera (V4\_p library). **b** ASVs identified within Cnidaria (V4 library) with tips colored by cnidarian order. **c** branches colored

by water body. Major clades were identified according to branch length and ASV clustering and labelled as Host Specific (HS) or Host Generalist (HG). Tree branch lengths are normalized for visual purposes only.

but they may also play a key role in the reacquisition of microbiota following stressor-associated dysbiosis.

### Complexity of Cnidarian microbiomes

Across Cnidaria, host phylogeny plays a key role in microbial assembly, influencing the structure, complexity, and diversity of cnidarian

microbiomes. Our cnidarian-wide comparisons greatly expand previous findings of phylogenetic influence on coral microbiomes<sup>52–54</sup>, and are similar to processes identified in poriferan holobionts<sup>55,56</sup>. We posit that increased body plan complexity, e.g., the evolution of skeletons in Alcyonacea, Antipatharia, Helioporacea, and Scleractinia, increased the number of internal niches, driving niche partitioning, and

facilitating increases in the richness and/or diversity of specific microbial communities (Fig. 3). More diverse microbial communities may lead to diversification of the holobiont phenotype, functional capacities, and metabolic capabilities, thereby enabling expansion of the host into novel habitats<sup>57</sup>. This diversification would help explain the significant changes in microbial richness correlated with the radiation of modern scleractinian families<sup>52</sup>. With Cubozoa, Hydrozoa, Scyphozoa, and Octocorallia hosting significantly less complex microbiomes than Hexacorallia, and less diverse microbiomes than Scleractinia, we propose that less complex microbiomes were the ancestral state of Cnidaria, with more complex microbiomes evolving more recently, similar to that of poriferan holobionts<sup>56</sup>.

### Geographic and phylogenetic influence of Cnidarian microbiomes

We found strong clustering by ocean basin of the microbiomes of closely related cnidarians (Supplementary Figs. 11–12), potentially suggesting that phyllosymbiosis (microbial patterns that mirror host phylogeny) may have occurred multiple times in cnidarian holobiont evolution, although this has yet to be tested. While bacterial phyllosymbiosis of Scleractinia and Octocorallia has been documented throughout the South-West Pacific<sup>52–54</sup> we posit that it may have occurred not only throughout Cnidaria, but repeatedly across geographically distinct ocean basins. Additionally, there are significant community signals of cnidarians residing across shallow and deeper depths, with deeper colonies hosting less diverse microbial assemblages. Further, depth strongly influenced microbial dissimilarity not only in Scleractinia, but across Cnidaria.

### Relative abundance of Cnidarian microbiota

Core communities of microbial genera were repeatedly identified across Cnidaria, with their relative abundance primarily influenced by host phylogeny and depth (Fig. 5a). Given that the ‘core microbiome’ has been proposed as important for holobiont resilience<sup>20,58</sup>, we identified smaller proportions of ‘core microbiota’ in shallow water scleractinian corals, that are often more susceptible to environmental changes, than in shallow water Actinarians and Alcyonaceans, which can be more resilient. Shallow Octocorallia exhibited greater microbial richness and higher core relative abundances when compared to phylogenetically older and deeper lineages, suggesting that these core microbiomes were repeatedly assembled across shallow water lineages<sup>57,59</sup>. Surprisingly, tissue core microbiomes (averaged to cnidarian family) consisted of only 0–10 ASVs that were present in all colonies per site (Fig. 5b). The taxonomic identity of these ASVs was highly conserved, exhibiting depth and cnidarian order fidelity, suggesting potentially important functional roles for these microbial families across Cnidaria (Fig. 5c). This expands previous findings that scleractinian microbiomes are highly individualistic<sup>60,61</sup>, and are dominated by rare sequences<sup>23,62</sup>, suggesting that this finding is common for many cnidarians, who potentially exhibit a high degree of functional redundancy across their stable microbial component.

### Prevalence and specificity of *Endozoicomonas*

Of all microbial families detected across Cnidaria, *Endozoicomonas* was not only the most prevalent, it was also one of the most ubiquitous globally and consistently present at different depths (Figs. 2–5). Therefore, *Endozoicomonas* may play an important role in healthy cnidarian microbiomes. *Endozoicomonas* are metabolically versatile<sup>63–65</sup> and may translocate vitamins to their host<sup>66</sup>. While *Endozoicomonas* occur in close proximity to Symbiodiniaceae in coral tissue<sup>67</sup>, 21.4% of the *Endozoicomonas* sequences in our dataset were identified from cnidarian samples that did not host Symbiodiniaceae, with 28.7% of all sequences identified isolated from non-scleractinian samples. Further, comparing the Illumina MiSeq subset of sequences (V4 library), demonstrated that octocorals in the Caribbean,

Mediterranean, and Indo-Pacific not only hosted the greatest absolute abundance of *Endozoicomonas*, but these sequences were shared across more samples, compared to any other cnidarian. Both specialist and generalist strains of *Endozoicomonas* exist within Anthozoa, suggesting a variable genetic component that has the potential for co-evolution with their hosts<sup>41,52–54,62</sup> similar to that of Symbiodiniaceae<sup>68</sup>. We expanded this observation and revealed patterns of switching in host specificity between Anthozoa and non-Anthozoa (including Cubozoa, Hydrozoa, Scyphozoa, and Porifera; Fig. 6a), as well as within Anthozoa, specifically forming multiple highly conserved associations with Alcyonacea and Rhizostomeae (Fig. 6b), in addition to fidelity to body compartment and ocean body (Fig. 6c).

### Potentially beneficial core microbiota

We identified additional common core microbiota across both tissue, skeletal, and mucus samples with potential functional importance. *Pseudoalteromonas*, a genus well-known as a producer of antimicrobials<sup>17</sup>, exhibits co-phylogeny with Scleractinia, as does *Alteromonas*<sup>62</sup>. Spirochaetes dominate some octocoral communities and may be involved in nutrient cycling<sup>69</sup>. While these bacteria have been previously identified as being relatively abundant or core members of many scleractinian and octocoral species, here we demonstrate that other cnidarians host these bacteria, hinting at their potential functional importance across Cnidaria. Additionally, we identified other bacteria that may be of functional importance but require further examination. For example, *Synechococcus* and *Prochlorococcus*, in the cyanobacterial genus Gp11a, commonly occur in cnidarian microbiomes<sup>60</sup>, as well as being abundant in marine water columns<sup>70</sup>, and therefore may be a sampling artifact. It is also important to experimentally determine whether *Propionibacterium* is a naturally abundant microbe or a sequencing kit contamination<sup>49,50</sup>. Although, by comparing these core microbiota to commonly found methodological contaminants i.e., *Propionibacterium*, we have increased the confidence that these bacterial genera may not be contaminants.

### Potentially detrimental core microbiota

In this study, we analyzed microbiomes sampled from putatively ‘healthy’ cnidarians, focusing on original datasets that did not mention observation of stress or anthropogenic disturbance(s), although this does not discount the possibility that some data may have come from stressed corals or emerging dysbiosis. Many of the bacterial genera that appeared as core members of ‘healthy’ tissue and skeletal microbiomes have been connected to anthropogenic stress, or are potential disease-causing microbes, including *Vibrio*, *Mycoplasma*, *Sphingomonas*, *Candidatus Pelagibacter*, and *Pelomonas*<sup>34,71–74</sup>. The presence of these genera was significantly increased with clustered sequence identification, indicating greater variability in these bacteria between samples. The prevalence and abundance of these microbes could potentially indicate non-virulent symbioses that could become rapidly virulent depending on microbe-microbe competition or fluctuation<sup>74,75</sup>, or that many phenotypically ‘healthy’ corals are experiencing underlying dysbiosis<sup>76</sup>.

### Symbiodiniaceae microbiota

Few of the microbiota from cultured Symbiodiniaceae also occurred in the cnidarian samples. *Pseudomonas*, for example, which was identified as a core bacterial genus in cnidarians, also occurred as a core intracellular microbe of Symbiodiniaceae<sup>77</sup>. *Pseudomonas* co-occurs with Symbiodiniaceae globally<sup>78</sup>, indicating a functional, or conserved symbiosis with both cnidarians and Symbiodiniaceae. This could explain the coculturing benefits of Symbiodiniaceae with *Pseudomonas* in laboratory studies<sup>79</sup>. Conversely, most other microorganisms that were identified as being part of the core microbiome of cultured Symbiodiniaceae<sup>77,80,81</sup> were not detected in cnidarian samples. Potentially, the microbiomes of cultured algae differ from those of

non-cultured Symbiodiniaceae, or there may be a minimal introduction of novel bacteria through the acquisition of Symbiodiniaceae on coral reefs. This incongruity reiterates the urgency for exploration into Symbiodiniaceae phycospheres and microbiomes<sup>82,83</sup>, especially given the critical role of Symbiodiniaceae in the survival of many cnidarians.

## Summary

This study provides important and fundamental insights into our understanding of how cnidarian microbiomes are structured globally. Cnidarian microbiomes are incredibly complex and integrating datasets can potentially enable identifying specific prokaryotic communities to focus future research on their functional role in cnidarian holobionts. The analysis also expanded our knowledge of prokaryotic communities across non-scleractinian cnidarians, which complements our understanding of the role and relevance of microbes within scleractinian corals. Because our data analysis demonstrated that clustering-based sequence pipelines are variable and depend on the cnidarian and microhabitat examined, we recommend that future cnidarian microbiome studies utilize unique sequence analysis. The current review focused primarily on a subset of the available data. Utilizing the entire database requires future investigations into the potential bias in sample storage, sequencing, and processing protocols. Given the current threats posed to cnidarians worldwide, our study provides a global, speciose, microbiome baseline for ‘healthy’ corals. With this baseline, we can begin to elucidate the impacts of stress-associated dysbiosis, in addition to addressing a myriad of ecological questions regarding the structuring of cnidarian microbiomes.

## Methods

### Sample and metadata collection

A literature search was conducted to obtain culture-independent 16S rRNA gene sequences available from any cnidarian. The literature search culminated in August 2021 with 257 datasets which identified sequences of the partial or full 16S rRNA genes of the microbial community from at least one cnidarian species (Supplementary Fig. 20). An additional seven projects had partially sequenced the 16S rRNA gene from cultured Symbiodiniaceae, seawater samples, and a standardized collection of globally distributed Porifera. Of the 257 studies, 88 contained sequence data that were either not publicly available, were multiplexed without index tags, lacked critical metadata, or could not be located online. Communication with corresponding authors resulted in 23 of these datasets being made available, which we gratefully acknowledge. From the 192 available data sets, 15 were further excluded due to sequencing files that could not be matched to their original samples, or continued to lack critical metadata.

The remaining 177 datasets provided 12,010 cnidarian associated 16S rRNA sequencing files from 212 cnidarian species (Fig. 1, Supplementary Data 1), with an additional 54 sample groups identified at a classification broader than species. Microbiome sequences from cnidarians included samples from members of all cnidarian classes: Anthozoa, Cubozoa, Hydrozoa, and Scyphozoa, and 12 orders (Actiniaria, Alcyonacea, Anthoathecata, Antipatharia, Carybdeida, Corallimorpharia, Helioporacea, Leptothecata, Rhizostomeae, Scleractinia, Semaestomeae, and Zoantharia). Nevertheless, sampling was highly biased toward Scleractinia (8754 samples), whereas Alcyonacea (1431), Actiniaria (1305) and the remaining orders (520) had far fewer samples (Supplementary Data 1).

Of the sequenced microbiome samples, 7575 were collected from cnidarians at 0–15 m depth, 473 from 16–30 m depth, 505 from 31–150 m depth, and 834 from >150 m depth (Supplementary Data 1). Geographically, microbiome samples were predominantly obtained from the South Pacific Ocean (3777), the North Atlantic Ocean (3376), the Northwest Pacific Ocean (1204), the Indian Ocean (1111), the

Mediterranean Sea (799), and the remaining from nine additional water bodies (Fig. 1, Supplementary Data 1).

To compare cnidarian microbiomes to the microbiomes of their mutualistic endosymbiotic algae (family Symbiodiniaceae), and the environment (water column), we also integrated 4003 microbial sequence files from seawater, poriferan, and cultured Symbiodiniaceae samples, resulting in a final dataset of ~6.5 billion sequences. Furthermore, the inclusion of ‘The Sponge Microbiome Project’<sup>51</sup> allowed us to compare across non-bilaterian marine invertebrates that often cohabit the same reefs.

Up to 33 parameters from metadata, including host, environmental, sequencing, and experimental variables, were collected for each sample, either from the original publication, supplementary information, or sequencing data (Supplementary Data 1).

### Sequence curation

FastQ files were retrieved from online databases, primarily from Genbank (<http://www.ncbi.nlm.nih.gov/genbank>) and the European Nucleotide Archive (<http://www.ebi.ac.uk/ena>), or directly provided by the published papers’ authors. The majority of these were sequenced using an Illumina MiSeq platform (12,808), with most downloaded in their raw unmerged format (8022), followed by 454 (1216), Illumina HiSeq (651), and additional platforms (1338). Idemp<sup>84</sup> was used to demultiplex files where required. Given the variety of sequencing platforms, each file was initially previewed with FastQC (v0.11.19)<sup>85</sup> to determine sequence quality scores and to locate any remaining barcodes, adaptors, or primers, which were subsequently removed with Trimmomatic (v0.39)<sup>86</sup> where present.

### Creation of shared sequence libraries

All edited sequences were visually observed and aligned to a full-length *Escherichia coli* sequence (Accession no. J01859) with GeneiousPrime (build 2020-11-05). From this, overlapping hypervariable regions were mapped across the entire dataset and the ranges for five cnidarian and one that included poriferan sequences (V4\_p) 16S rRNA libraries were generated: V1-V2 (30–360 bp, 1949 samples), V3 (350–500 bp, 2106 samples), V4 (520–780 bp, 5979 samples), V5-V6 (820–1050 bp, 3751 samples), V7-V8 (1040–1400 bp, 469 samples), and V4\_p (550–660 bp, 9737 samples). Sequences from each study were filtered and trimmed separately using the DADA2 R-package (v1.18-1.22)<sup>87</sup> in RStudio (v1.4.1103-4)<sup>88</sup>, allowing for the calculation of a distinct error rate per study. Single and merged Illumina sequences, 454, Ion Torrent, and PacBio sequences were all denoised with unique parameters, according to the DADA2 SOP (<https://benjjneb.github.io/dada2/tutorial.html>). All sequences were processed with the filtering qualifications of (maxN = 0, maxEE = 4, truncQ = 10), merged if appropriate, sequence tables constructed, and chimeras removed. Once sequence tables for each project had been created within a library, they were collapsed together before taxonomic classification with a naïve Bayesian method using DADA2. A total of 720,835 ASVs were present across the five cnidarian libraries, with 493,157 ASVs present in the V4\_p library. Sequences were then merged with their metadata and exported using the ‘phyloseq’ R-package (v1.38.0)<sup>89</sup> for further analysis.

To increase the detection rate of host mitochondrial sequences, the Silva training dataset (version 138.1<sup>90</sup>) was adapted by inserting 133 partial and full-length mitochondrial sequences representing 82% of the cnidarian and 60% of the poriferan genera found in this study<sup>91</sup>. Additional sequences were added to further identify recently identified cnidarian symbionts, including 120 sequences from eukaryotic ARL (Apicomplexan-Related-Lineages)<sup>92,93</sup>, and 12 from ‘*Candidatus Aquarickettsia rhoweri*’<sup>72</sup>.

### Non-archaeal and bacterial sequences ASVs

ASVs classified as “Eukaryote”, “Mitochondria”, “Chloroplast”, or unassigned at the kingdom level, were removed from all six shared

libraries and from each core microbiome dataset. To test the suitability of this common contamination removal process, the 34,220 ASVs excluded from the five cnidarian libraries, were investigated using BLAST +<sup>94</sup>. A local megablast search was conducted, with the top hit automatically chosen for sequences above 90% sequence similarity (Supplementary Fig. 13).

**Analyzing the structure and diversity of cnidarian microbiomes**  
As the V4 library included the greatest abundance of cnidarian specific sequences (360,006), and the broadest phylogenetic variability (11 cnidarian orders) of samples in our dataset, it was selected for further analysis. Sequences were aligned with the ‘DECIPHER’ R-package v2.22.0<sup>95</sup>, subset to family and used to generate phylogenetic trees. These neighbor-joining trees were optimized with a GTR + G substitution model (model determined by modelTest) to estimate a maximum-likelihood trees, using the ‘phangorn’ R-package (v2.8.1)<sup>96</sup> and merged into phyloseq objects with the ‘ape’ R-package (v5.6-1)<sup>97</sup>. Prevalence of microbial families and relative abundances within the four cnidarian classes were plotted and visualized (Fig. 2) using the ‘ggtree’ R-package (v3.2.1)<sup>98</sup>.

Given the current debate concerning the applicability of unrarefied vs. rarefied data<sup>99,100</sup>, we conducted our analyses on both an unrarefied and rarefied (minimum 5000 reads) V4 library, subset only to ‘healthy’, non-experimental samples sequenced on a MiSeq (4243, and 2964, respectively). From these samples, four additional subsets were created and tested, including only scleractinian, and all non-scleractinian orders.

Shannon diversity, ASVs observed, and Inverse-Simpson metrics were tested on these three data subsets with a linear mixed-effects model, treating the original study as a random blocking effect, using the ‘lme4’ R-package (v1.1-27.1)<sup>101</sup>. R<sup>2</sup> values corrected with false discovery rates were then calculated with the ‘emmeans’ R-package (v1.7.2)<sup>102</sup> and the ‘MuMIn’ R-package (v1.43.17)<sup>103</sup>. Significant factors were further tested with Kruskal Wallis analyses, accompanied by Dunn post hoc tests and Bonferroni corrections (Supplementary Figs. 1–2, 4–8) using the ‘dunn.test’ R-package (v1.3.5)<sup>104</sup>, and the ‘ranacapa’ R-package (v0.1.0)<sup>105</sup>. Rarefaction plots were created using the ‘vegan’ R-package (v2.5-7)<sup>106</sup>, and the Shannon diversity was plotted (Fig. 3), with the ‘ggplot2’ R-package v3.3.5<sup>107</sup>. Extrapolation of diversity was conducted using the ‘iNEXT’ R-package (v3.0.0)<sup>108</sup>.

Bray-Curtis and Jaccard dissimilarity indices and dispersions were calculated for the three unrarefied and rarefied subsets of the V4 library, using a PERMANOVA (999 permutations) with the original study treated as a random blocking effect (Supplementary Tables 1–6), using the ‘vegan’ R-package v2.5-7<sup>106</sup> (Supplementary Figs. 9–10). Both the V4 and V4\_p libraries were hierarchically clustered (complete linkage) using Bray-Curtis dissimilarities and plotted with the ‘stats’ R-package (v4.1.2)<sup>88</sup>.

### Identifying the shared and distinctive microbiota

The absolute abundance of ASVs from each of the five cnidarians libraries that were shared across cnidarian phyla, individual microhabitat, depth, and ocean bodies were calculated and visualized (Fig. 4a–c) with the ‘MiscMetabar’ R-package (v0.21)<sup>109</sup>, while shared sequences between cnidarian orders were calculated and visualized (Fig. 4d) using the ‘circlize’ R-package (v0.4.13)<sup>110</sup>.

### Creation of core microbiome datasets

A total of 143 studies sequenced a minimum of four ‘healthy’ (no record of observed stress) cnidarian (tissue, skeletal, mucus) samples per species per sampling site, collected from ambient conditions (no environmental or experimental stressors documented). The raw untrimmed sequences from these studies were reprocessed separately using DADA2. While the denoising parameters were the same as above, these sequences were not merged with other studies, allowing for the

analysis of their full length of sequences, and sequence tables were classified separately. Sequences were then merged with their metadata and exported using the ‘phyloseq’ R-package for further analysis.

### Calculating variability in the ‘core’ microbiota

The 588,776 ASVs from the core microbiota sequence tables were additionally clustered into 198,710 ASVs<sub>97</sub> sequences with 97% similarity (approximating OTUs) for comparison, using the ‘DECIPHER’ R-package v2.22.0<sup>95</sup>. Both ASV and ASVs<sub>97</sub> sequences from each of the 143 sequence tables were then tested for the relative abundance of shared sequences across each cnidarian species at each sampling site. Sequences were determined as being present across sampled individuals, including every sample per site (100%), more than 75%, more than 50%, less than 50%, and only in a single cnidarian individual (Fig. 5a), using the ‘microbiome’ R-package (v1.16.0)<sup>111</sup>. These values were averaged per site for each cnidarian species and were averaged again within cnidarian families. Kruskal Wallis analyses tested the relative abundances of bacteria against the cnidarian order, symbiotic status, and across sampling depth, and were accompanied by Dunn post hoc tests followed by Bonferroni corrections. The absolute abundance of ASVs present in 100% of samples (Core ASVs) per cnidarian species was averaged across families and plotted (Fig. 5b) using the ‘ggplot2’ R-package<sup>107</sup>. Further, the identities of the ASV and ASVs<sub>97</sub> identified in 100% of individuals, and greater than 75% of individuals of at least one cnidarian species was also visualized (Fig. 5c). Potential PCR and kit contaminants were identified and removed using the ‘decontam’ R-package v1.14.0<sup>112</sup>.

### Endozoicomonas specificity, abundance, and fidelity in Cnidaria

*Endozoicomonas* sequences and abundances were isolated from all six libraries using the ‘phyloseq’ R-package<sup>89</sup>, aligned with the ‘DECIPHER’ R-package v2.22.0<sup>95</sup> and were used to generate phylogenetic trees as previously detailed. The absolute and average abundances of *Endozoicomonas* sequences for each clade were plotted against the average number of samples that *Endozoicomonas* were identified within the phyla Cnidaria and Porifera (Fig. 6a), within cnidarian orders (Fig. 6b), and within water bodies and whether the *Endozoicomonas* was host specific or generalist (Fig. 6c).

### Data availability

All six 16S rRNA gene libraries generated in this study (V1–V2<sup>113</sup>, V3<sup>114</sup>, V4<sup>115</sup>, V5–V6<sup>116</sup>, V7–V8<sup>117</sup>, V4\_p<sup>118</sup>) have been deposited in the ‘Cnidarian Microbiome Database Project’ at Figshare ([https://figshare.com/projects/Cnidarian\\_Microbiome\\_Database/158957](https://figshare.com/projects/Cnidarian_Microbiome_Database/158957)), alongside the cnidarian specific RDP annotation file. All original metadata are provided in Supplementary Data 1, alongside sequence accession numbers for each sample, where available.

### References

- Eddy, T. D. et al. Global decline in capacity of coral reefs to provide ecosystem services. *One Earth* **4**, 1278–1285 (2021).
- Descombes, P. et al. Forecasted coral reef decline in marine biodiversity hotspots under climate change. *Glob. Chang Biol.* **21**, 2479–2487 (2015).
- Rohwer, F., Breitbart, M., Jara, J., Azam, F. & Knowlton, N. Diversity of bacteria associated with the Caribbean coral *Montastraea franksi*. *Coral Reefs* **20**, 85–91 (2001).
- Wegley, L., Edwards, R., Rodriguez-Brito, B., Liu, H. & Rohwer, F. Metagenomic analysis of the microbial community associated with the coral *Porites astreoides*. *Environ. Microbiol.* **9**, 2707–2719 (2007).
- Brandt, K. Ueber das Zusammenleben von thieren und algen. *Arch. Anat. Physiol.* **1881**, 570–574 (1881).
- Kellogg, C. A. Tropical archaea: diversity associated with the surface microlayer of corals. *Mar. Ecol. Prog. Ser.* **273**, 81–88 (2004).

7. Glynn, P. W. Widespread coral mortality and the 1982–83 El Niño warming event. *Environ. Conserv* **11**, 133–146 (1984).
8. Jones, R. & Yellowlees, D. Regulation and control of intracellular algae (=zooxanthellae) in hard corals. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **352**, 457–468 (1997).
9. Kawaguti, S. The effect of light on calcium deposition in corals. *Bull. Oceanogr. Inst. Taiwan* **4**, 65–70 (1948).
10. LaJeunesse, T. C. et al. Systematic revision of Symbiodiniaceae highlights the antiquity and diversity of coral endosymbionts. *Curr. Biol.* **28**, 2570–2580.e2576 (2018).
11. Maruyama, S. et al. Heat stress of algal partner hinders colonization success and alters the algal cell surface glycome in a cnidarian-algal symbiosis. *Microbiol Spectr.* **10**, 3 (2022).
12. Margulis, L. Symbiogenesis and symbiogenesis. In: *Symbiosis as a source of evolutionary innovation: Speciation and morphogenesis*. 1–14 (MIT Press, Cambridge, 1991).
13. Radecker, N., Pogoreutz, C., Voolstra, C. R., Wiedenmann, J. & Wild, C. Nitrogen cycling in corals: the key to understanding holobiont functioning? *Trends Microbiol.* **23**, 490–497 (2015).
14. Bourne, D. G., Morrow, K. M. & Webster, N. S. Insights into the coral microbiome: underpinning the health and resilience of reef ecosystems. *Annu Rev. Microbiol.* **70**, 317–340 (2016).
15. Benavides, M., Bednarz, V. N. & Ferrier-Pagès, C. Diazotrophs: overlooked key players within the coral symbiosis and tropical reef ecosystems? *Front. Mar. Sci.* **4**, 10 (2017).
16. Silveira, C. B. et al. Microbial processes driving coral reef organic carbon flow. *FEMS Microbiol. Rev.* **41**, 575–595 (2017).
17. Stabili, L., Parisi, M. G., Parrinello, D. & Cammarata, M. Cnidarian interaction with microbial communities: from aid to animal's health to rejection responses. *Mar. Drugs* **16**, 296 (2018).
18. Boilard, A. et al. Defining coral bleaching as a microbial dysbiosis within the coral holobiont. *Microorganisms* **8**, 1682 (2020).
19. Rosenberg, E., Koren, O., Reshef, L., Efrony, R. & Zilber-Rosenberg, I. The role of microorganisms in coral health, disease and evolution. *Nat. Rev. Microbiol.* **5**, 355–362 (2007).
20. Peixoto, R. S., Rosado, P. M., Leite, D. C., Rosado, A. S. & Bourne, D. G. Beneficial microorganisms for corals (BMC): proposed mechanisms for coral health and resilience. *Front. Microbiol.* **8**, 341 (2017).
21. Modolon, F., Barno, A. R., Villela, H. D. M. & Peixoto, R. S. Ecological and biotechnological importance of secondary metabolites produced by coral-associated bacteria. *J. Appl. Microbiol.* **129**, 1441–1457 (2020).
22. Thompson, J. R., Rivera, H. E., Closek, C. J. & Medina, M. Microbes in the coral holobiont: partners through evolution, development, and ecological interactions. *Front Cell Infect. Microbiol.* **4**, 176 (2014).
23. Ricci, F. et al. Beneath the surface: community assembly and functions of the coral skeleton microbiome. *Microbiome* **7**, 159 (2019).
24. Pernice, M. et al. Down to the bone: the role of overlooked endolithic microbiomes in reef coral health. *ISME J.* **14**, 325–334 (2020).
25. Webster, N. S. & Reusch, T. B. H. Microbial contributions to the persistence of coral reefs. *ISME J.* **11**, 2167–2174 (2017).
26. Voolstra, C. R. & Ziegler, M. Adapting with microbial help: microbiome flexibility facilitates rapid responses to environmental change. *Bioessays* **42**, e2000004 (2020).
27. Wegley Kelly, L., Haas, A. F. & Nelson, C. E. Ecosystem microbiology of coral reefs: linking genomic, metabolomic, and biogeochemical dynamics from animal symbioses to reefscape processes. *mSystems* **3**, e00162–17 (2018).
28. Wilkins, L. G. E. et al. Host-associated microbiomes drive structure and function of marine ecosystems. *PLoS Biol.* **17**, e3000533 (2019).
29. Vanwonderghem, I. & Webster, N. S. Coral reef microorganisms in a changing climate. *iScience* **23**, 100972 (2020).
30. Glasl, B., Herndl, G. J. & Frade, P. R. The microbiome of coral surface mucus has a key role in mediating holobiont health and survival upon disturbance. *ISME J.* **10**, 2280–2292 (2016).
31. Santoro, E. P. et al. Coral microbiome manipulation elicits metabolic and genetic restructuring to mitigate heat stress and evade mortality. *Sci. Adv.* **7**, 33 (2021).
32. Rosado, P. M. et al. Marine probiotics: increasing coral resistance to bleaching through microbiome manipulation. *ISME J.* **13**, 921–936 (2019).
33. McFadden, et al. Systematics of Octocorallia (Cnidaria: Anthozoa) guided by phylogenomics. *Bull. Soc. Syst. Biol.* **1**, 3 (2022).
34. van de Water, J., Allemand, D. & Ferrier-Pages, C. Host-microbe interactions in octocoral holobionts - recent advances and perspectives. *Microbiome* **6**, 1–28 (2018).
35. McCauley, M., Jackson, C. R. & Goulet, T. L. Microbiomes of Caribbean octocorals vary over time but are resistant to environmental change. *Front. Microbiol.* **11**, 1272 (2020).
36. Dube, C. E. et al. Naturally occurring fire coral clones demonstrate a genetic and environmental basis of microbiome composition. *Nat. Commun.* **12**, 6402 (2021).
37. Blackall, L. L., Wilson, B. & van Oppen, M. J. Coral—the world's most diverse symbiotic ecosystem. *Mol. Ecol.* **24**, 5330–5347 (2015).
38. Glasl, B., Bourne, D. G., Frade, P. R. & Webster, N. S. Establishing microbial baselines to identify indicators of coral reef health. *Microbiol. Aust.* **39**, 42–46 (2018).
39. Hernandez-Agreda, A., Leggat, W. & Ainsworth, T. D. A place for taxonomic profiling in the study of the coral prokaryotic microbiome. *FEMS Microbiol. Lett.* **366**, fnz063 (2019).
40. Kellogg, C. A. Microbiomes of stony and soft deep-sea corals share rare core bacteria. *Microbiome* **7**, 90 (2019).
41. Huggett, M. J. & Apprill, A. Coral microbiome database: integration of sequences reveals high diversity and relatedness of coral-associated microbes. *Environ. Microbiol. Rep.* **11**, 372–385 (2019).
42. Epstein, H. E., Hernandez-Agreda, A., Starko, S., Baum, J. K. & Vega Thurber, R. Inconsistent patterns of microbial diversity and composition between highly similar sequencing protocols: a case study with reef-building corals. *Front. Microbiol.* **12**, 740932 (2021).
43. Callahan, B. J., McMurdie, P. J. & Holmes, S. P. Exact sequence variants should replace operational taxonomic units in marker-gene data analysis. *ISME J.* **11**, 2639–2643 (2017).
44. Chiarello, M., McCauley, M., Villegier, S. & Jackson, C. R. Ranking the biases: the choice of OTUs vs. ASVs in 16S rRNA amplicon data analysis has stronger effects on diversity measures than rarefaction and OTU identity threshold. *PLoS ONE* **17**, e0264443 (2022).
45. Thompson, L. R. et al. A communal catalogue reveals Earth's multiscale microbial diversity. *Nature* **551**, 457–463 (2017).
46. Sweet, M. J. & Bulling, M. T. On the importance of the microbiome and pathobiome in coral health and disease. *Front. Mar. Sci.* **4**, 9 (2017).
47. Ying, H. et al. Comparative genomics reveals the distinct evolutionary trajectories of the robust and complex coral lineages. *Genome Biol.* **19**, 175 (2018).
48. Bergman, J. L., Shaw, T., Egan, S. & Ainsworth, T. D. Assessing the coral microbiome at the scale of tissue-specific habitats within the coral meta-organism. *Front. Mar. Sci.* **9**, 985496 (2022).
49. Glassing, A., Dowd, S. E., Galandiuk, S., Davis, B. & Chiodini, R. J. Inherent bacterial DNA contamination of extraction and sequencing reagents may affect interpretation of microbiota in low bacterial biomass samples. *Gut Pathog.* **8**, 24 (2016).
50. Salter, S. J. et al. Reagent and laboratory contamination can critically impact sequence-based microbiome analyses. *BMC Biol.* **12**, 1–12 (2014).

51. Moitinho-Silva, L. et al. The sponge microbiome project. *Giga-science* **6**, 1–7 (2017).
52. Pollock, F. J. et al. Coral-associated bacteria demonstrate phylosymbiosis and cophylogeny. *Nat. Commun.* **9**, 4921 (2018).
53. O'Brien, P. A. et al. Diverse coral reef invertebrates exhibit patterns of phyllosymbiosis. *ISME J.* **14**, 2211–2222 (2020).
54. O'Brien, P. A. et al. Testing cophylogeny between coral reef invertebrates and their bacterial and archaeal symbionts. *Mol. Ecol.* **30**, 3768–3782 (2021).
55. Thomas, T. et al. Diversity, structure and convergent evolution of the global sponge microbiome. *Nat. Commun.* **7**, 11870 (2016).
56. Sabrina Pankey, M. et al. Cophylogeny and convergence shape holobiont evolution in sponge-microbe symbioses. *Nat. Ecol. Evol.* **6**, 750–762 (2022).
57. Henry, L. P., Bruijning, M., Forsberg, S. K. G. & Ayroles, J. F. The microbiome extends host evolutionary potential. *Nat. Commun.* **12**, 5141 (2021).
58. Trevathan-Tackett, S. M. et al. A horizon scan of priorities for coastal marine microbiome research. *Nat. Ecol. Evol.* **3**, 1509–1520 (2019).
59. Sanchez, J. A., Gonzalez-Zapata, F. L., Prada, C. & Duenas, L. F. Mesophotic gorgonian corals evolved multiple times and faster than deep and shallow lineages. *Diversity* **13**, 650 (2021).
60. Hernandez-Agreda, A., Leggat, W., Bongaerts, P., Herrera, C. & Ainsworth, T. D. Rethinking the coral microbiome: simplicity exists within a diverse microbial biosphere. *mBio* **9**, e00812–e00818 (2018).
61. Ainsworth, T. et al. The coral core microbiome identifies rare bacterial taxa as ubiquitous endosymbionts. *ISME J.* **9**, 2261–2274 (2015).
62. Ricci, F. et al. Host traits and phylogeny contribute to shaping coral-Bacterial symbioses. *mSystems* **7**, e0004422 (2022).
63. Ding, J. Y., Shiu, J. H., Chen, W. M., Chiang, Y. R. & Tang, S. L. Genomic insight into the host-endosymbiont relationship of *Endozoicomonas montiporae* CL-33(T) with its coral host. *Front. Microbiol.* **7**, 251 (2016).
64. Neave, M. J., Michell, C. T., Apprill, A. & Voolstra, C. R. *Endozoicomonas* genomes reveal functional adaptation and plasticity in bacterial strains symbiotically associated with diverse marine hosts. *Sci. Rep.* **7**, 40579 (2017).
65. Tandon, K. et al. Comparative genomics: dominant coral-bacterium *Endozoicomonas acroporae* metabolizes dimethylsul-foniopropionate (DMSP). *ISME J.* **14**, 1290–1303 (2020).
66. Pogoreutz, C. et al. Coral holobiont cues prime *Endozoicomonas* for a symbiotic lifestyle. *ISME J.* **16**, 1883–1895 (2022).
67. Bayer, T. et al. Bacteria of the genus *Endozoicomonas* dominate the microbiome of the Mediterranean gorgonian coral *Eunicella cavolini*. *Mar. Ecol. Prog. Ser.* **479**, 75–84 (2013).
68. Gonzalez-Pech, R. A., Bhattacharya, D., Ragan, M. A. & Chan, C. X. Genome evolution of coral reef symbionts as intracellular residents. *Trends Ecol. Evol.* **34**, 799–806 (2019).
69. van de Water, J. A. et al. Spirochaetes dominate the microbial community associated with the red coral *Corallium rubrum* on a broad geographic scale. *Sci. Rep.* **6**, 27277 (2016).
70. Weber, L. & Apprill, A. Diel, daily, and spatial variation of coral reef seawater microbial communities. *PLoS ONE* **15**, e0229442 (2020).
71. Richardson, L. L. et al. Florida's mystery coral-killer identified. *Nature* **392**, 557–558 (1998).
72. Klinges, J. G. et al. Phylogenetic, genomic, and biogeographic characterization of a novel and ubiquitous marine invertebrate-associated Rickettsiales parasite, *Candidatus Aquarickettsia rohweri*, gen. nov., sp. nov. *ISME J.* **14**, 2938–2953 (2020).
73. Ziegler, M. et al. Coral bacterial community structure responds to environmental change in a host-specific manner. *Nat. Commun.* **10**, 3092 (2019).
74. Rubio-Portillo, E. et al. Virulence as a side effect of interspecies interaction in *Vibrio* coral pathogens. *mBio* **11**, 4 (2020).
75. Rubio-Portillo, E. et al. *Vibrio* communities in scleractinian corals differ according to health status and geographic location in the Mediterranean Sea. *Syst. Appl. Microbiol.* **41**, 131–138 (2018).
76. Corinaldesi, C. et al. Changes in coral forest microbiomes predict the impact of marine heatwaves on habitat-forming species down to mesophotic depths. *Sci. Total Environ.* **823**, 153701 (2022).
77. Maire, J. et al. Intracellular bacteria are common and taxonomically diverse in cultured and in hospite algal endosymbionts of coral reefs. *ISME J.* **15**, 2168–2170 (2021).
78. Bernasconi, R., Stat, M., Koenders, A. & Huggett, M. J. Global networks of *Symbiodinium*-Bacteria within the coral holobiont. *Micro. Ecol.* **77**, 794–807 (2019).
79. Zhang, L., Hirai, M. & Shoda, M. Removal characteristics of dime-thyl sulfide by a mixture of *Hyphomicrobium* Sp-I55 and *Pseudomonas-Acidovorans* Dmr-11. *J. Ferment Bioeng.* **74**, 174–178 (1992).
80. Lawson, C. A., Raina, J. B., Kahlke, T., Seymour, J. R. & Suggett, D. J. Defining the core microbiome of the symbiotic dinoflagellate, *Symbiodinium*. *Environ. Microbiol. Rep.* **10**, 7–11 (2018).
81. Camp, E. F. et al. Revealing changes in the microbiome of *Symbiodiniaceae* under thermal stress. *Environ. Microbiol.* **22**, 1294–1309 (2020).
82. Matthews, J. L. et al. *Symbiodiniaceae*-bacteria interactions: rethinking metabolite exchange in reef-building corals as multi-partner metabolic networks. *Environ. Microbiol.* **22**, 1675–1687 (2020).
83. Garrido, A. G., Machado, L. F. & Zilberberg, C. Leite DCdA. Insights into 'Symbiodiniaceae phycosphere' in a coral holobiont. *Symbiosis* **83**, 25–39 (2020).
84. Wu, Y. idemp: Barcode Demultiplex for Illumina I1, R1, R2 fastq.gz files. <https://github.com/yhwu/idemp>. (2017).
85. Andrews, S. FASTQC. A quality control tool for high throughput sequence data. <http://www.bioinformatics.babraham.ac.uk/projects/fastqc>. (2010).
86. Bolger, A. M., Lohse, M. & Usadel, B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**, 2114–2120 (2014).
87. Callahan, B. J. et al. DADA2: high-resolution sample inference from Illumina amplicon data. *Nat. Methods* **13**, 581–583 (2016).
88. Team R. RStudio: Integrated Development for R. RStudio, PBC, Boston, MA. (2020).
89. McMurdie, P. J. & Holmes, S. phyloseq: an R package for reproducible interactive analysis and graphics of microbiome census data. *PLoS ONE* **8**, e61217 (2013).
90. Quast, C. et al. The SILVA ribosomal RNA gene database project: Improved data processing and web-based tools. *Nucleic Acids Res.* **41**, D590–D596 (2013).
91. Sonett D., Brown T., Bengtsson-Palme J., Padilla-Gamiño J. L., Zaneveld J. R. The Organelle in the room: under-annotated mitochondrial reads bias coral microbiome analysis. Preprint available at bioRxiv (2021).
92. Janouskovec, J. et al. Apicomplexan-like parasites are polyphyletic and widely but selectively dependent on cryptic plastid organelles. *Elife* **8**, e49662 (2019).
93. Vohsen, S. A. et al. Deep-sea corals provide new insight into the ecology, evolution, and the role of plastids in widespread apicomplexan symbionts of anthozoans. *Microbiome* **8**, 34 (2020).
94. Camacho, C. et al. BLAST+: architecture and applications. *BMC Bioinform.* **10**, 421 (2009).
95. Wright, E. S. Using DECIPHER v2.0 to analyze big biological sequence data in R. *R. J.* **8**, 352–359 (2016).
96. Schliep, K. P. phangorn: phylogenetic analysis in R. *Bioinformatics* **27**, 592–593 (2011).

97. Paradis, E. & Schliep, K. ape 5.0: an environment for modern phylogenetics and evolutionary analyses in R. *Bioinformatics* **35**, 526–528 (2019).
98. Yu, G. Using ggtree to visualize data on tree-like structures. *Curr. Protoc. Bioinforma.* **69**, e96 (2020).
99. Cameron, E. S., Schmidt, P. J., Tremblay, B. J., Emelko, M. B. & Muller, K. M. Enhancing diversity analysis by repeatedly rarefying next generation sequencing data describing microbial communities. *Sci. Rep.* **11**, 22302 (2021).
100. McMurdie, P. J. & Holmes, S. Waste not, want not: why rarefying microbiome data is inadmissible. *PLoS Comput Biol.* **10**, e1003531 (2014).
101. Bates, D., Machler, M., Bolker, B. & Walker, S. Fitting linear mixed-effects models using lme4. *J. Stat. Softw.* **67**, 1–48 (2015).
102. Lenth, R., Singmann, H., Love, J., Buerkner, P. & Herve, M. R. Package ‘emmeans’. *Am Stat.* **34**, 216–221 (2019).
103. Bartón, K. R. Package ‘MuMIn’. Model selection and model averaging based on information criteria. <https://cran.r-project.org/web/packages/MuMIn/MuMIn.pdf>. (2018).
104. Dinno, A. & Dinno, M. A. Package ‘dunn.test’. CRAN. <https://cran.r-project.org/web/packages/dunn.test/dunn.test.pdf>. (2017).
105. Kandlikar, G. S. et al. ranacapa: an R package and Shiny web app to explore environmental DNA data with exploratory statistics and interactive visualizations. *F1000Res* **7**, 1734 (2018).
106. Oksanen, J., Blanchet, F. G., Kindt, R. & Legendre, P. R Package ‘vegan’: Community Ecology Package. <https://cran.r-project.org/web/packages/vegan/vegan.pdf>. (2014).
107. Wickham H. ggplot2: Elegant graphics for data analysis. Springer-Verlag New York. (2016).
108. Hsieh, T. C., Ma, K. H. & Chao, A. iNEXT: an R package for rarefaction and extrapolation of species diversity (Hill numbers). *Methods Ecol. Evol.* **7**, 1451–1456 (2016).
109. Taudière, A. R package ‘MiscMetabar’: miscellaneous functions for metabarcoding analysis. <https://github.com/adrietaudiere/MiscMetabar>. (2021).
110. Gu, Z., Gu, L., Eils, R., Schlesner, M. & Brors, B. Circlize implements and enhances circular visualization in R. *Bioinformatics* **30**, 2811–2812 (2014).
111. Lahti L., Shetty S., Ernst F. R. Package ‘microbiome’. (2022).
112. 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* **6**, 1–14 (2018).
113. McCauley, M. V1-V2 Microbiome Library. <https://doi.org/10.6084/m9.figshare.23290076.v1> (2023).
114. McCauley, M. V3 Microbiome Library. <https://doi.org/10.6084/m9.figshare.23290094.v1> (2023).
115. McCauley, M. V4 Microbiome Library (phyloseq). <https://doi.org/10.6084/m9.figshare.23290097.v1> (2023).
116. McCauley, M. V5-V6 Microbiome Library (phyloseq). <https://doi.org/10.6084/m9.figshare.23290106.v1> (2023).
117. McCauley, M. V7-V8 Microbiome Library (phyloseq). <https://doi.org/10.6084/m9.figshare.23290109.v1> (2023).
118. McCauley, M. V4\_p Microbiome Library (phyloseq). <https://doi.org/10.6084/m9.figshare.23290100.v1> (2023).
119. Kayal, E. et al. Phylogenomics provides a robust topology of the major cnidarian lineages and insights on the origins of key organismal traits. *BMC Evol. Biol.* **18**, 68 (2018).
120. Kayal, E., Roure, B., Philippe, H., Collins, A. G. & Lavrov, D. V. Cnidarian phylogenetic relationships as revealed by mitogenomics. *BMC Evol. Biol.* **13**, 5 (2013).

## Acknowledgements

We greatly thank all contributing scientists for their published sequences and correspondence. Especially, we would like to thank Allison Kerwin, Aki Ohdera, and Monica Medina for the inclusion of unpublished sequences in our dataset. We would also like to thank James Strother for providing assistance and access to his computational servers, and Marlène Chiarello for her contributions to one of our figures. Additional thanks to Joe Ryan, Virginia Weis, and Mark Martindale for their insights. This work was supported by the National Science Foundation under grant no. NSF EF-2025476 and NSF IOS-2124120 awarded to SL, and NSF DEB-1831531 awarded to CRJ. Any opinions, findings, conclusions, and/or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

## Author contributions

M.M., C.R.J.: conceptualized the study. M.M.: conducted the review, collected and processed sequences, created sequence libraries, analyzed data, created figures, and drafted the paper. C.R.J., S.L.: contributed to, and assisted in data analysis and presentation, supported the study financially. M.M., T.L.G., C.R.J., S.L.: edited manuscript. All authors have approved the final version of the manuscript.

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s41467-023-39876-6>.

**Correspondence** and requests for materials should be addressed to M. McCauley.

**Peer review information** *Nature Communications* thanks Sebastian Fraune and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. A peer review file is available.

**Reprints and permissions information** is available at <http://www.nature.com/reprints>

**Publisher’s note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

This is a U.S. Government work and not under copyright protection in the US; foreign copyright protection may apply 2023