Marine Sponge Endosymbionts: Structural and Functional Specificity of the Microbiome within *Euryspongia arenaria* Cells

Qi Yang,[a,b,c] Jackson K. B. Cahn,[c*] Jörn Piel,[c] Yue-Fan Song,[d] Wei Zhang,[b] Hou-Wen Lin[a]

[a]Center for Marine Drugs, State Key Laboratory of Oncogene and Related Genes, Department of Pharmacy, Renji Hospital, School of Medicine, Shanghai Jiao Tong University, Shanghai, China
[b]Centre for Marine Bioproducts Development, College of Medicine and Public Health, Flinders University, South Australia, Australia
[c]Institute of Microbiology, Eidgenössische Technische Hochschule (ETH) Zurich, Zurich, Switzerland
[d]College of Food Science and Engineering, Key Laboratory of Aquatic Product Processing and Utilization of Liaoning Province, Dalian Ocean University, Dalian, China

ABSTRACT Sponge microbiomes are typically profiled by analyzing the community DNA of whole tissues, which does not distinguish the taxa residing within sponge cells from extracellular microbes. To uncover the endosymbiotic microbiome, we separated the sponge cells to enrich the intracellular microbes. The intracellular bacterial community of sponge *Euryspongia arenaria* was initially assessed by amplicon sequencing, which indicated that it hosts three unique phyla not found in the extracellular and bulk tissue microbiomes. These three phyla account for 66% of the taxonomically known genera in the intracellular microbiome. The shotgun metagenomic analysis extended the taxonomic coverage to viruses and eukaryotes, revealing the most abundant signature taxa specific to the intracellular microbiome. Functional KEGG pathway annotation demonstrated that the endosymbiotic microbiome hosted the greatest number of unique gene orthologs. The pathway profiles distinguished the intra- and extracellular microbiomes from the tissue and seawater microbiomes. Carbohydrate-active enzyme analysis further discriminated each microbiome based on their representative and dominant enzyme families. One pathway involved in digestion system and family esterase had a consistently higher level in intracellular microbiome and could statistically differentiate the intracellular microbiome from the others, suggesting that triacylglycerol lipases could be the key functional component peculiar to the endosymbionts. The identified higher abundance of lipase-related eggNOG categories further supported the lipid-hydrolyzing metabolism of endosymbiotic microbiota. *Pseudomonas* members, reported as lipase-producing bacteria, were only in the endosymbiotic microbiome, meanwhile *Pseudomonas* also showed a greater abundance intracellularly. Our study aided a comprehensive sponge microbiome that demonstrated the taxonomic and functional specificity of endosymbiotic microbiota.

IMPORTANCE Sponges host abundant microbial symbionts that can produce an impressive number of novel bioactive metabolites. However, knowledge on intracellular (endosymbiotic) microbiota is scarce. We characterize the composition and function of the endosymbiotic microbiome by separation of sponge cells and enrichment of intracellular microbes. We uncover a noteworthy number of taxa exclusively in the endosymbiotic microbiome. We unlock the unique pathways and enzymes of endosymbiotic taxa. This study achieves a more comprehensive sponge microbial community profile, which demonstrates the structural and functional specificity of the endosymbiotic microbiome. Our findings not only open the possibility to reveal the low abundant and the likely missed microbiota when directly sequencing the sponge bulk tissues, but also warrant future in-depth exploration within single sponge cells.

KEYWORDS sponge *Euryspongia arenaria*, sponge cells, endosymbiotic microbiome, functional specificity, structural specificity

Marine sponges (phylum Porifera), which evolved ca. 750 million years ago, are the oldest extant metazoans (1). There are 9,490 valid species reported from marine and freshwater environments to date, which are found across tropical, temperate, and polar regions (2). Sponges are the most prolific sources of marine natural products, contributing almost 30% of all marine natural products reported during 2001 to 2010 (3). On average, about 200 to 350 new compounds were discovered from sponges each year since 2011 (4–13). The sponge species belonging to genus *Euryspongia* selected in this study is considered as a promising source of bioactive natural products since the initial report of euryfuran production (14). A new cyclized 9,11-secosterol enol-ether was first isolated from South Australian sponge *E. arenaria* (15). Various secondary metabolites with different bioactivities, such as cytotoxic, antifungal, antihistaminic, anti-HIV, and anti-cancer activities (16–24), have subsequently been reported from sponges belonging to genus *Euryspongia*.

Sponges form symbiotic relationships with often complex communities of microorganisms (25–30). Up to 40% to 60% of the tissue volume of certain sponge species can consist of microorganisms at a density exceeding 10^9 microbial cells per mL of sponge tissue, orders of magnitude greater than concentrations in surrounding seawater or sediment (31–33). The associated microbes produce diverse metabolites that have stimulated the interest of the pharmaceutical industry (34–38). Typically, the prokaryotic composition of sponge microbiomes is analyzed by high-throughput 16S rRNA gene-based sequencing approaches (27–30, 39–43). Recently, metagenomic sequencing provided new insights into the functional gene repertoires of sponges and the metabolic pathways carried by sponge symbionts (26, 44–46). However, metagenomic methods can be challenging in samples with low-microbial abundance or in those, like sponges, dominated by host DNA (47). For the members of a microbiome not yet characterized by culture-based methods, a significant fraction of reads in the metagenome may still remain unmatched after assembly (48). The quality of assembly can be impacted by the complexity of the community and the sequencing technology (49).

Sponge microbiomes are predominantly studied via directly sequencing the community DNA of whole tissue samples. They have shown host specificity and remarkable stability across a large range of environmental conditions (25, 27, 29, 41, 43), although for certain sponge species, shifts in community structure under different biogeographic locations and seasons have been observed (50–55). A pretreatment to enrich extracellular prokaryotes before DNA isolation has also been employed for the investigation of sponge microbiomes (44, 56–60). However, knowledge on intracellular sponge microbiota is scarce (61–64) and only two types of endosymbiotic bacteria have been identified—calcifying bacteria (Calcibacteria) (61–63) and renieramycin-producing bacteria (*Candidatus* Endohaliclona renieramycinifaciens) (64). Of particular interest are the endosymbiotic microbiomes of archaeocytes and choanocytes, which constitute the stem cell population in all four classes of sponges and generate all cell types for reconstituting dissociated tissues (65).

Our study aims to reveal the sponge cell endosymbiotic microbiome by sequencing the metagenomic DNA isolated from the sponge cells after enrichment and purification, as well as to profile its functional specificity compared to the extracellular, whole tissue, and seawater microbiomes. A South Australia locally growing *Euryspongia arenaria* was selected as the model sponge in our study because this sponge could be maintained in a controlled aquarium system under stable living conditions in pilot experiments of sponge aquaculture and sponge cell separation. An optimized sponge cell purification strategy was utilized to collect the sponge cell fractions. A newly developed multi-primer amplicon sequencing approach was applied to profile the microbial communities. Finally, shotgun metagenomic sequencing was further employed to investigate the functional gene diversity specific to the endosymbiotic microbiome. With this study, we provide a comprehensive profile of the *E. arenaria* microbiome to
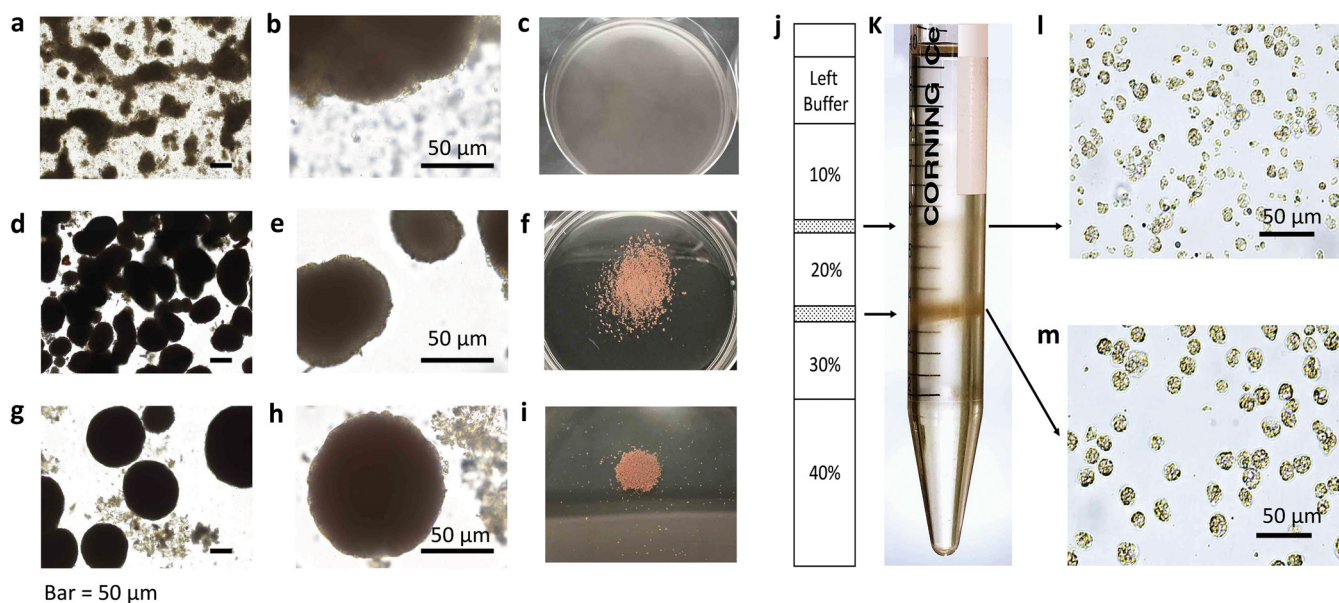
FIG 1 Sponge cells cultivation and purification. (a, b) Aggregated sponge cells by *in vitro* cultivation in artificial seawater for 2 h, vigorous cells started gathering. (c) Cells were inoculated in a Petri dish with artificial seawater after 2 h. (d, e) 6 h cultivation, vigorous cells formed solid aggregates. (f) Cells were inoculated in a Petri dish with artificial seawater after 6 h. (g, h) 12 h cultivation, aggregates became bigger in size. (i) Cells were cultivated in a Petri dish with artificial seawater after 12 h. (j) Illustration of Ficoll-CMFASW density gradient and cell fractions after centrifugation. (k) Ficoll-CMFASW density gradient and cell fractions after centrifugation in tube. (l) Optical microscopy image of sponge cell fraction between 10% and 20% Ficoll-CMFASW gradients. (m) Optical microscopy image of sponge cell fraction between 20% and 30% Ficoll-CMFASW gradients.

demonstrate the specificity of its intracellular microbial community regarding structural and functional features. Metagenome based annotations allowed for a correlation analysis between the functions and their potential microbial origins.

RESULTS

Enrichment of sponge cells and extracellular bacteria. To investigate intra- and extracellular microbiomes, sponge tissues were first dissociated for target fraction enrichment. Sponge archaeocytes (triangle, Fig. S1a and b) and choanocytes (circle, Fig. S1a and b) were enriched at 100 × g centrifugation (Fig. S1c) to separate out most of the extracellular bacteria. The microbial population was enriched gradually in the pellets as the centrifugation speed increased (Fig. S1d to f). The pellets from the centrifugations at 2,000 × g and 16,000 × g were combined to profile the extracellular bacteria community. The enriched sponge cells were further purified by *in vitro* cultivation in artificial seawater for 12 h (Fig. 1) until the aggregates formed "primmorphs" (66) with smooth surfaces (Fig. 1g and h). An additional gradient density centrifugation step separated two cell fractions: one was enriched in archaeocyte cells, comprising the band between 20% and 30% Ficoll (Fig. 1l), and the other was enriched in choanocyte cells with some epithelial cells present, comprising the 10% to 20% Ficoll layer (Fig. 1m). Each fraction was enriched by more than 70% for individual cell types. DNA sequence analysis was carried out on these two purified sponge cell fractions to reveal the microbial community of the archaeocyte- and choanocyte-dominant fraction, respectively.

Ultrastructural characterization of sponge cells. To characterize sponge cell types, transmission electron microscopy (TEM) was utilized to observe and measure cells. The typical archaeocytes in the purified fraction were usually 10 μm or above in diameter (Fig. 2a) with prominent and solid lysosomes and large nuclei, which were 3 μm or above in diameter. The choanocytes, usually less than 8 μm in diameter (Fig. 2b), were generally smaller than archaeocytes. In fixed tissues, the choanocytes were surrounded by cilia. Vacuoles were frequently found in archaeocytes with digested bacteria inside (Fig. 2c). In addition, bacteria-like cells were detected to be scattered in the cytoplasm of
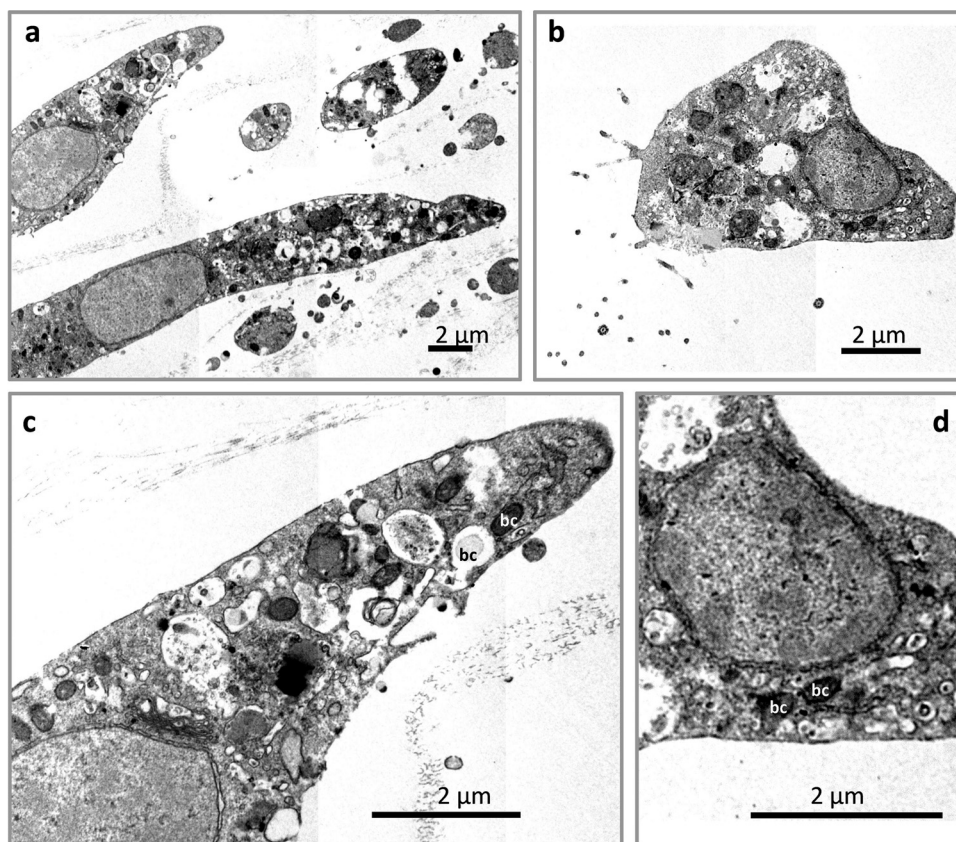
FIG 2 Morphological characterization of sponge cells and endosymbiotic microbe identification by transmission electron microscopy. (a) Archaeocytes. (b) Choanocytes. (c) Endosymbiotic bacteria-like cells within one archaeocyte cell shown in a. (d) Endosymbiotic bacteria-like cells within one choanocyte cell shown in b. The white "bc" highlights the non-digested bacteria like cells, the black "bc" highlights the digested bacteria like cells.

both archaeocytes (Fig. 2c) and choanocytes (Fig. 2d). The most observed bacteria-like organisms are 0.5 μm long and rod-shaped with condensed contents.

Purity validation by consistent Raman spectra of individual sponge cells. To evaluate and validate the homogeneity of the sponge cells within the same fraction, *in situ* non-destructive detection by Raman spectroscopy was applied to directly characterize and compare the metabolite profile of individual living sponge cells. The observed fingerprint bands pointed to the presence of certain chemical components and intensity in the spectra in each sponge cell and its symbionts. The observed spectra of sponge cells in archaeocyte fraction are composed of three major bands located at $1,010 \pm 5$ cm^{-1}, $1,151 \pm 8$ cm^{-1}, and $1,508 \pm 6$ cm^{-1} (Fig. 3a). The band pattern of sponge choanocyte fraction is $1,008 \pm 8$ cm^{-1}, $1,148 \pm 6$ cm^{-1}, and $1,541 \pm 8$ cm^{-1} (Fig. 3b). The trends of the intensity ratio are different between the two cell fractions (Fig. 3). The Raman spectra of four randomly selected cells in each of five investigation areas within the same subfraction showed a highly consistent band pattern and intensity ratios. The spectra obtained by investigating five representative areas showed a stable consistency. Only slight variation between the subfractions was observed regarding the actual intensity of the major peaks.

Microbial community compositions of different sponge cell types assessed by 16S rRNA gene amplicon sequencing. The 16S rRNA gene amplicon sequencing data covering the V1 to V8 regions revealed the sponge microbial communities of the prepared four types of samples. On average, sequence reads of 225,908 ($<$ 0.08% variation between biological replicates), 544,449 ($<$ 0.02% variation), 601,804 ($<$ 0.04% variation), and 543,060 ($<$ 0.06% variation) were obtained from the sponge tissue sample,
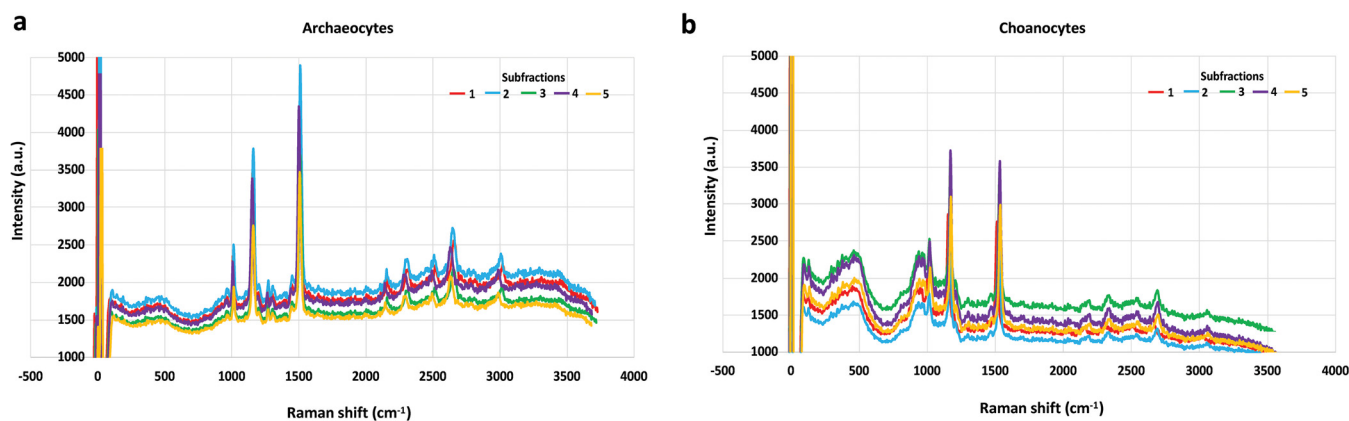
FIG 3 Raman spectra of single sponge cell for validation of homogenous cell sorting. (a) Spectra of each of 100 cells from five subfractions of purified sponge archaeocyte fraction. (b) Spectra of each of 100 cells from five subfractions of purified sponge choanocyte fraction. The combination is generated after normalization of zero peak.

archaeocyte fraction, choanocyte fraction), and the extracellular bacteria fraction, respectively (Table S1a). Rarefactions using three metrics (PD whole tree, Chao1, and observed species) demonstrated the archaeocyte fraction and choanocyte fraction had higher microbial richness compared to the other two communities (extracellular bacterial fraction and tissue sample) (Supplemental file 1). The curves generated using Shannon index of all samples already reached a plateau at this sequencing depth, suggesting that the sequencing was deep enough (Supplemental file 1). Combined, 21 phylum-level microbial OTUs, including 17 phyla, three candidate phyla, and one unassigned OTU, were revealed. The comparison between the communities based on the relative abundance of each OTU is provided as a hierarchical clustering heatmap in Fig. 3a. The dendrogram on top, obtained by using a "correlation" distance measurement and the "complete linkage" clustering method (67), indicated that the two purified sponge cell fractions displayed a similar community composition and structure (reflected by the relative abundance of each microbial taxon in the community) that were highly distinct from those of the sponge tissue and the enriched extracellular bacteria fraction (Fig. 4a). Moreover, the beta diversity using Bray-Curtis distance among the four microbial communities ($P = 0.0001$ by PERMANOVA) further supported the grouping pattern (Fig. 4b).

Both the whole-sponge tissues and the extracellular bacterial fraction showed higher microbial diversity of phylum-level OTUs than either cell fraction (Fig. 4c). However, members of three microbial phyla, Chlamydiae, Spirochaetes, and *Candidatus* Tectomicrobia, were only found in the intracellular cell fractions (Table S1b). Chlamydiae members were only obtained from the sponge archaeocytes. Sequencing of whole sponge tissue and bacteria-enriched fractions revealed the same microbial phyla, with 10 unique phyla (nine known phyla and one candidate phylum) compared with the sponge cell fractions. Eight phyla were commonly shared among the microbial communities of the whole sponge tissue and the three fractions.

A comparison at genus level (affiliated and candidate genera) of the microbial communities demonstrated that the sponge tissue contributed the highest number of unique known (candidate) genera, i.e., 31, compared with 24 found in the bacteria-enriched fraction and 19 from the two sponge cell fractions (Fig. 4d and Table S1c). Within the cell fractions, the archaeocyte fraction shared 57 known genera with the choanocyte fraction, while these two fractions had 13 and six unique genera, respectively. Only three genera were found present in all tested samples.

Unique 16S rRNA gene sequence distribution among different types of sponge cell. To go beyond taxonomic assignments, the distribution of unique sequences among the four sample types were investigated to compare their microbial communities. Based on a single-nucleotide difference cut-off for distinctive sequence identification, a total of
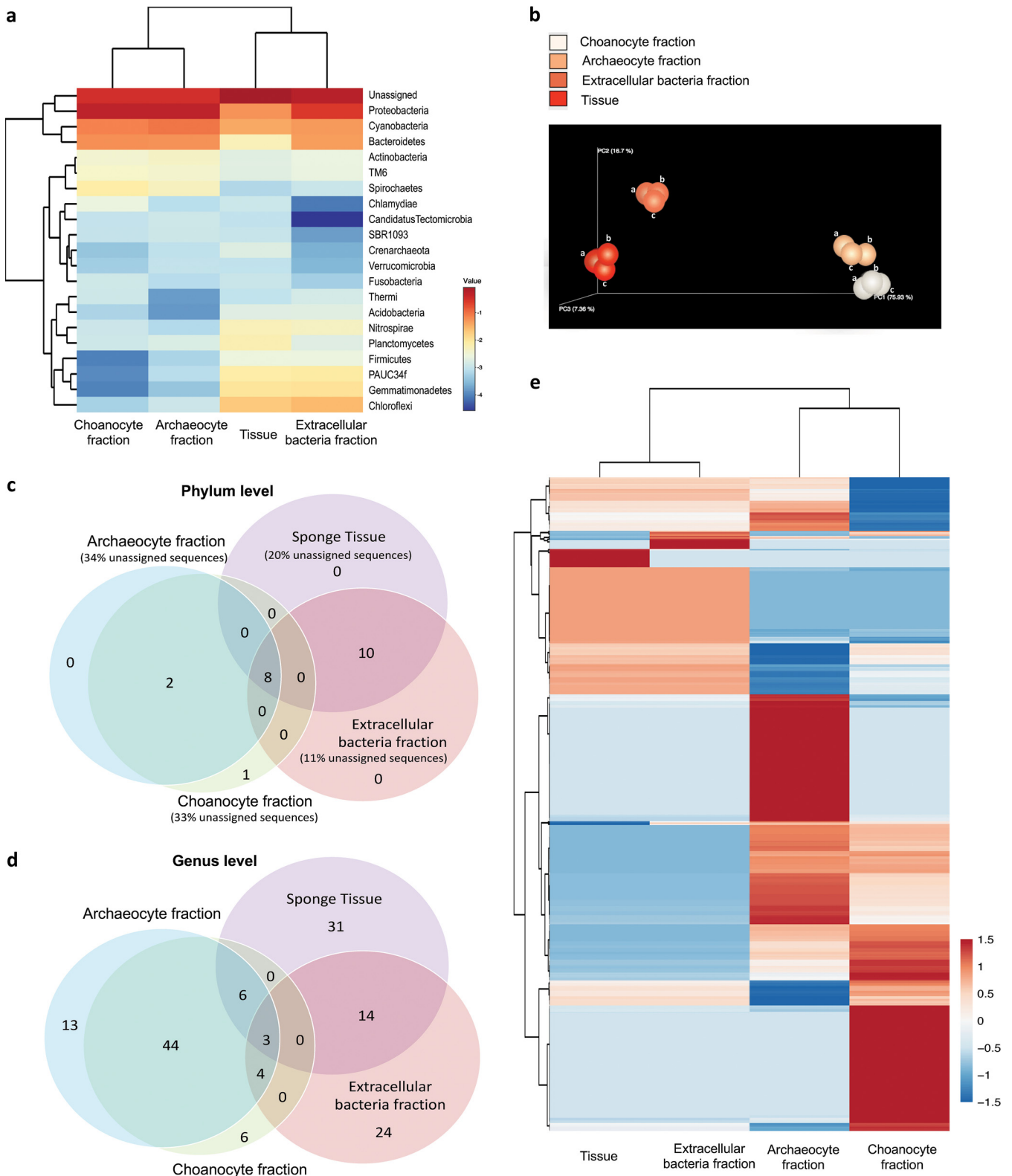
FIG 4 Comparisons among the four microbial communities revealed by amplicon analysis. (a) Taxonomic heatmap for phylum-level OTUs. (b) Principal coordinates analysis (PCoA) plot derived from Bray-Curtis distance among microbial communities ($P = 0.0001$ by PERMANOVA). Labels a to c distinguish the three biological replicates. (c) Venn diagram for microbial phylum level OTUs distribution. (d) Venn diagram for microbial genus level OTUs distribution. All the unclassified genus-level OTUs are counted as one. (e) Abundance of unique sequences in the four microbial communities. Data in heatmaps are centered and the unit scaling is applied. Both rows and columns are clustered using maximum distance method. The value of the color bar is calculated by $\log_{10}$ of the relative abundance of each OTU. The red and blue colors reflect the highest and lowest relative abundance in the community, respectively.

500 unique sequences were detected. Among them, 118 sequences could be assigned at the order level, but belong to an unidentified family; 382 sequences could be assigned to a known family; and 139 sequences could be assigned down to the genus level (Supplemental file 1). Notably, sequences assigned to the genus *Legionella*, a group of Gram-negative bacteria that includes human pathogens, were only detected in the extracellular bacteria fraction and the tissue sample (Supplemental file 1). A comparison between the communities is shown in a heatmap using hierarchical clustering analysis based on the relative abundance of the unique sequences inferred from raw data (Fig. 4e). The distribution of the unique sequences was found to be specific to each of the four microbial communities (Fig. 4e and Supplemental file 1).

In-depth microbiota identification from different types of sponge cell. Using the unique sequences analysis approach, an in-depth characterization of the identified microbiotas was conducted to further distinguish the microbiomes at the sequence level. The in-depth comparison of multiple 16S rRNA gene sequences assigned to the same genus or family allows differentiation between organisms at the species/subspecies level or inference of potential new species. Twelve families were detected that occurred in all the microbial communities of sponge tissue, purified sponge cell fractions, and the enriched extracellular bacteria fraction (Fig. 5 and Supplemental file 2), although distributions and abundances within these families varied across the four microbial communities. Additionally, the unique sequences specific to the microbial community of each of the cell fractions were identified. For the families of Alteromonadaceae, Campylobacteraceae, Coxiellaceae, and Pseudoalteromonadaceae, both sponge cell fractions had distinguishable unique sequences that were not present in the extracellular bacterial fraction and tissue sample (Fig. 5a to d). For Flavobacteriaceae, Rhodobacteraceae, and Vibrionaceae, the cell fractions shared unique sequence entries (Fig. 5e to g). For Legionellaceae (Fig. 5h), the unique sequences came from archeocyte fraction and the extracellular bacteria fraction. Moreover, for Endozoicomonadaceae, Phyllobacteriaceae, Oceanospirillaceae, and Synechococcaceae, the archeocyte fraction (Fig. 5i and j), extracellular bacteria fraction (Fig. 5k), or the tissue sample (Fig. 5l) were indicated as the single contributor for the unique sequence(s).

To explore potential novel microbiota from the unique sequences specific to individual sponge cell fractions, these multiple sequences belonging to the same families/ genera were further distinguished at the genus/species level by conducting individual BLAST searches in the NCBI gene database. Moreover, the base variations between the sequences assigned as the same family/genus were observed. For example, the sequences (seq3, seq13, and seq21) assigned to genus *Synechococcus* in family Synechococcaceae were further inferred to match the species *Synechococcus rubescens* with similarity of 98%, 96%, and 96%, respectively. These in-depth classifications and the phylogenetic trees (Supplemental file 2) could identify potentially novel microbiota.

Community structure specificity of the sponge intracellular microbiome uncovered by metagenome sequencing. 16S rRNA gene amplicon sequencing is useful for the identification of microbial community members, but shotgun metagenomic sequencing is required to identify the functional potential of a community by a marker gene approach. Considering the challenge of metagenomics in samples dominated by host DNA (Fig. S2), we removed the sequences belonging to sponges (Phylum Porifera) using the Burrows-Wheeler Aligner (BWA) before further processing. To better present the microbial community, the sequences of other eukaryotes except fungi were also filtered out for the rest of the analyses. As a result, an average of 70.2 million paired-end reads (150 bp) were collected for each sample (Supplemental file 3). After quality check, more than 80% reads were retained for *de novo* assembly (contigs or scaffolds). Approximately 4.1 million metagenes were then clustered into 2,442,690 non-redundant genes (unigenes).

Annotations using reference sequences in the nr database assigned the unigenes to protein types and taxonomic affiliations. About 85% unigenes were assigned to prokaryotes (bacteria and archaea), 14% were viral genes, and only 1% were annotated as fungi. Metagenome sequencing significantly enlarged the microbiome coverage (141 phyla including eight fungal and one viral phyla) compared with amplicon analysis (21 phyla)
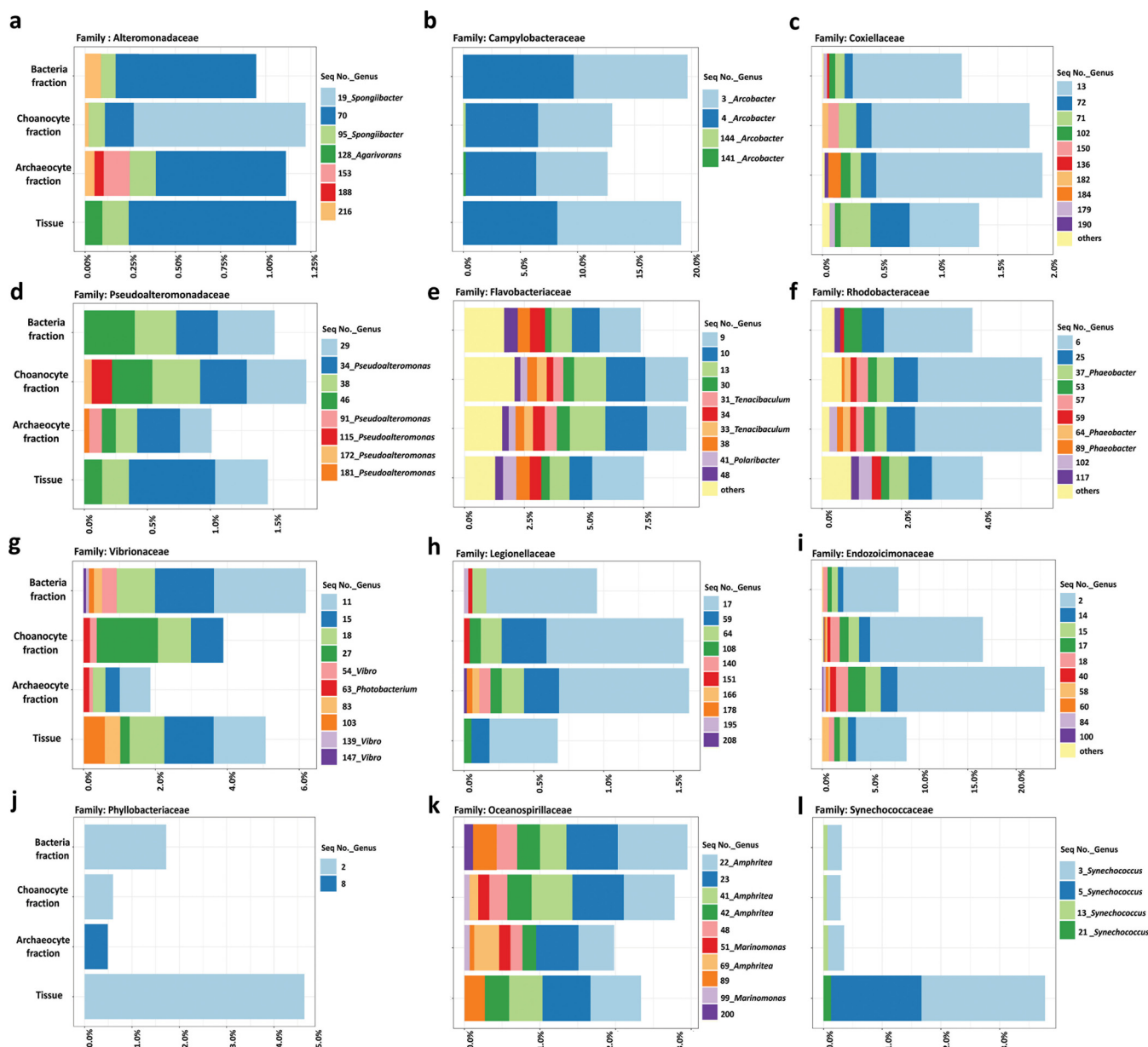
FIG 5 Distribution of multiple sequences within the same family or genus revealed by amplicon analysis. The title of the bars is labeled in one of the three subfigures sorted in the same row (i.e., subfigures a, d, g, and j). The other subfigures (b, c, e, f, h, i, k, and l) follow the same vertical axis format. The titles of each subfigure are the family names. Bacteria fraction, extracellular bacteria enriched; Tissue, whole sponge sample.

(Fig. 6a), but the dominant phyla remained consistent between the two approaches, as did the relative-abundance patterns for each type of samples (Fig. 6a). Given the consistency, the dominant core microbiota in each sample type were identified based on both data sets: Acidobacteria, Actinobacteria, Cyanobacteria, Proteobacteria, and Unclassified for the tissue microbiome (89.6% and 60.2% relative sequence abundance in the amplicon and metagenomic data, respectively), as well as for the extracellular microbiome (87.7% and 92.6%). In comparison, Bacteroidetes, Cyanobacteria, Proteobacteria, Spirochaetes, and Unclassified were the core taxa for the intracellular microbiome (99.1% and 82.9%). The seawater microbiome contained members of *Ca.* Tectomicrobia, Firmicutes, Proteobacteria, Unclassified virus phylum, and Unclassified bacterial and archaeal phyla as the most dominant phyla (84.3% relative sequence abundance in the metagenomic data).

Considering all the unigenes, the profiled microbiomes of the four sample types can be clearly distinguished based on their specific composition and structure of the
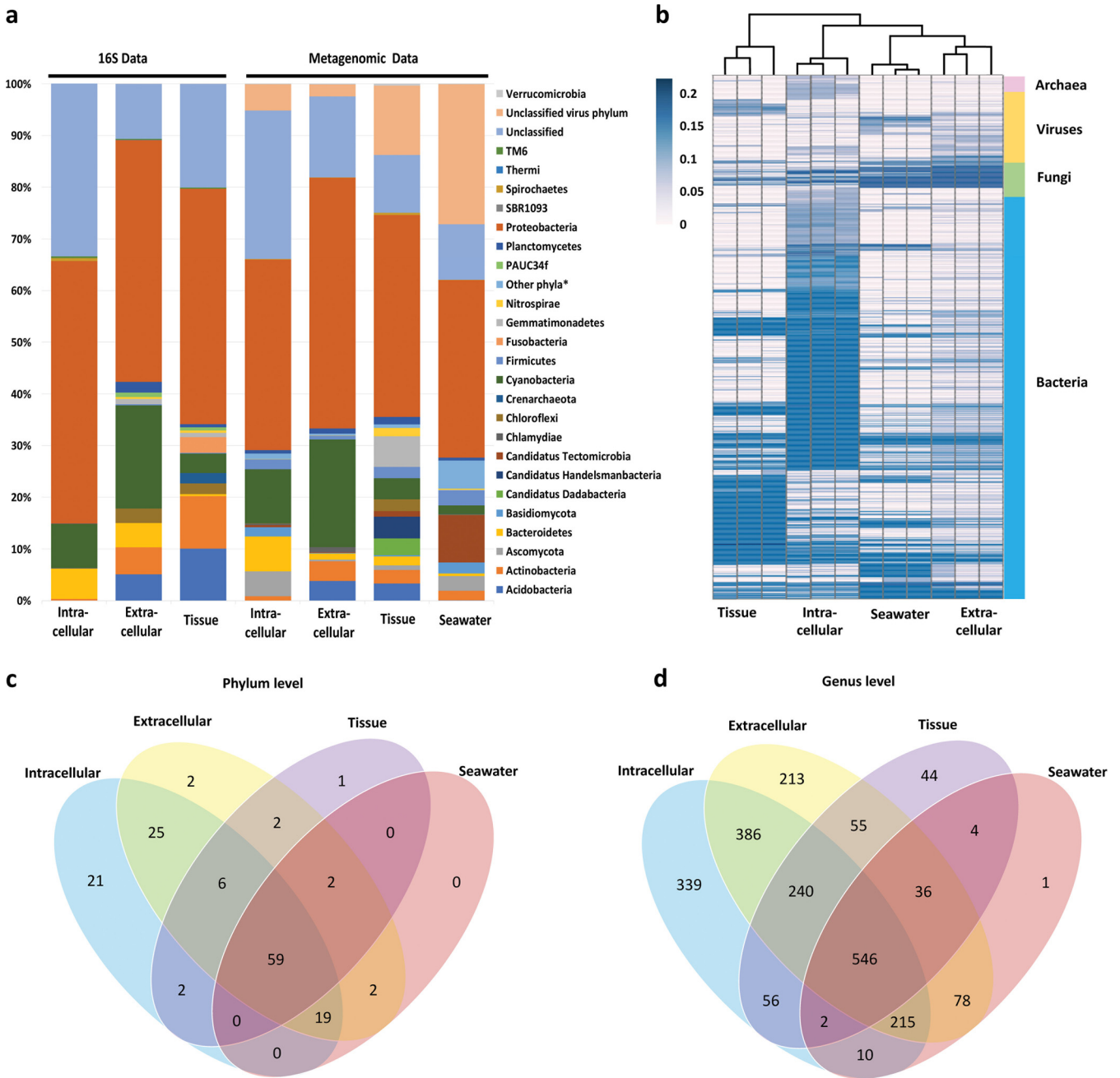
FIG 6 Taxonomic distribution and diversity comparisons among the four microbiomes revealed by metagenomic analysis. (a) Phylum-level taxa comparisons between 16S rRNA gene amplicon and metagenomic analysis. The metagenomic data excluded the Porifera sequences. Other phyla* includes the sum of all the rest phyla (relative abundance < 0.01% each) generated by metagenomic data. The intracellular profile by amplicon sequencing included two data sets of cell archaeocyte fraction and choanocyte fraction samples. The intracellular profile by metagenomics was based on the sequencing data of cell archaeocyte fraction and choanocyte fraction mixture. (b) Similarity among the microbiomes of the four sample types (three biological replicates) based on the relative abundance of each genus-level microbial unigenes. (c) Venn diagram for phylum-level taxa distribution among four microbiomes. (d) Venn diagram for genus-level taxa distribution among four microbiomes.

genus-level microbial members (Fig. 6b) that covered archaea, bacteria, fungi, and viruses. Principal coordinates analysis (PCoA) using Bray-Curtis distance also demonstrated the clearly distinguished four samples (*P* = 0.0001 by PERMANOVA) (Supplemental file 3). The extracellular and seawater samples were considerably more similar to one another than to either the intracellular or tissue samples. At the phylum level, the intracellular microbiome hosted a considerable higher number of the unique taxa than the other three communities (Fig. 6c). There were 120 sponge-derived phyla, 65 of which were shared
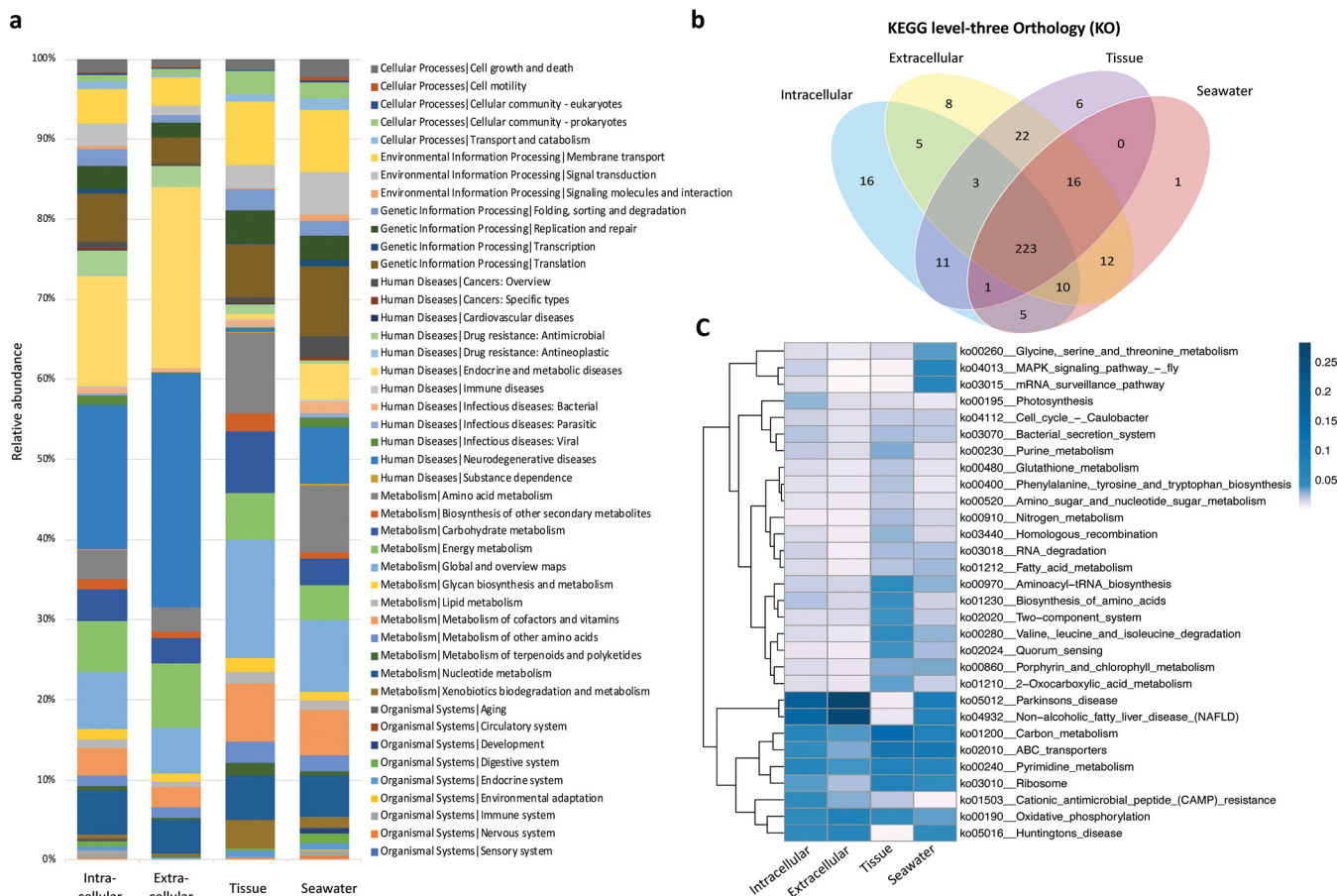
FIG 7 KEGG pathways comparison among the four microbiomes revealed by metagenomic sequencing. (a) Functional KEGG level-two pathway distribution. (b) Venn diagram depicting the numbers of KEGG level-three orthology (KO) in each sample. The unique 16 pathways include ko00290 (Valine, leucine and isoleucine biosynthesis), ko00250 (Alanine, aspartate and glutamate metabolism), ko00450 (Selenocompound metabolism), ko00626 (Naphthalene degradation), ko00624 (Polycyclic aromatic hydrocarbon degradation), ko00514 (Other types of O-glycan biosynthesis), ko00591 (Linoleic acid metabolism), ko00943 (Isoflavonoid biosynthesis), ko00640 (Propanoate metabolism), ko00710 (Carbon fixation in photosynthetic organisms), ko00130 (Ubiquinone and other terpenoid-quinone biosynthesis), ko05131 (Shigellosis), ko05231 (Choline metabolism in cancer), ko05110 (*Vibrio cholerae* infection), ko04122 (Sulfur relay system), and ko04520 (Adherens junction). (c) Heatmap depicting the distribution of the top 30 dominate level-three pathways. Clustering is based on the relative abundance of each KO.

between all the sample types. Of the 82 seawater-derived phyla, 73% were also found among the sponge samples. Consistently, down to the genus level, the greatest number of unique taxa were found in the intracellular microbiome, which included 17 archaeal genera, 267 bacterial genera, 47 fungal genera, and eight viruses genera, followed by extracellular microbiome (eight archaeal genera, 174 bacterial genera, 126 fungal genera, and five viruses genera), tissue microbiome (three archaeal genera, 37 bacterial genera, two fungal genera, and two viruses genera), and then seawater microbiome with only one bacterial genus (Fig. 6d).

Functional specificity of the sponge intracellular microbiome uncovered by metagenome sequencing. Functional annotations were firstly obtained by a BLAST search of the clustered unigenes against the KEGG orthology (KO) database; for all microbiomes, KOs were primarily assigned to 45 KEGG level-two pathways (Fig. 7a). Within the KEGG level-three pathways, we obtained 338 sponge-derived pathways (226 shared between the groups); 99% of seawater-derived pathways were also included in this grouping. We further identified 16 unique pathways from the intracellular microbiome (Fig. 7b), a greater number than that found to be unique to the extracellular (eight), tissue (six), and seawater (one) samples. The 16 pathways belong to categories metabolism (11 pathways), human diseases (three pathways), genetic information processing (one pathway), and cellular processes (one pathway). Considering the relative abundance of

each level-three pathway, it is found that the sponge tissue and seawater microbiomes contained a much more even distribution among the dominant pathways (Fig. 7c, indicating top 30 dominant KOs); in contrast, the intra- and extracellular microbiomes were highly enriched in a few KOs.

To examine differences in carbohydrate metabolism between these microbial communities, functional annotations were generated by querying the carbohydrate-active enzyme (CAZy) database, and the hierarchical annotation was analyzed based on the distribution of six enzyme classes in the different microbiomes (Fig. 8). Overall, the enzymes annotated from the metagenomes showed a high level of glycoside hydrolases (GH), glycosyl transferases (GT), and carbohydrate-binding modules (CBM) in the studied microbiomes. Pairwise comparison of the samples demonstrated the largest difference between the intracellular fraction and the tissue microbiome, with statistically significant differences in auxiliary activities (AA, P-value < 0.05), carbohydrate esterases (CE, P-value < 0.005), GH (P-value < 0.05), and polysaccharide lyases (PL, P-value < 0.005) (Fig. 8a). The most common variances between the samples were the abundances of CE and GT as they both had the greatest number of comparisons, between each other among the four microbiomes, showing statistically significant difference. The tissue sample had the largest number of significant-difference pairs against each of the other three microbiomes for different CAZy classes. Correspondingly, the family-level CAZys were also dominated by GH (family 16), GT (families 2, 4, 47, and 51), and CBM (families 2, 6, 50, and 53) (Fig. 8b). Moreover, different microbiomes can be distinguished by the enrichment of specific enzymes. There are seven, 16, and 11 unique CAZy families specific to intra-, extracellular, and tissue samples, respectively (Fig. 8c). All of the seawater related CAZys were covered by sponge-derived ones. Among the intracellular specific enzymes, we found enzymes involved in degradation (i.e., carboxyl esterase in CE10 esterases). Other typical serine hydrolases families (CE1 lipase and CE5 cutinase) were further analyzed. The results showed that they were dominant in the intracellular microbiome and had the highest abundance (55% and 43% for CE1 and CE5, respectively) compared with other three microbiomes, followed by the extracellular microbiome (33% and 34% for CE1 and CE5, respectively) (Fig. 8d and e).

Unique functional features of the sponge intracellular microbiome. An analysis of the functional features and taxonomic structure of the microbiomes allowed us to better understand the specific functional roles of the sponge microbiota (Fig. 9). The significant-difference analysis indicated the three most differentially abundant KEGG pathways in each of the six pairwise comparisons (Fig. 9a). The pathway ko04972 (involved in the digestive system) was the only one consistently identified with higher abundance in intracellular microbiome compared with extracellular, tissue, and seawater microbiomes. Notably, for the analysis of CAZys, only the family CE10 (carboxyl esterase) kept a consistently higher abundance in intracellular microbiome (Fig. 9b). For the eggNOG annotated clusters of orthologous genes (COGs), the COGs belonging to groups I (lipid transport and metabolism), L (replication, recombination and repair), Q (secondary metabolites biosynthesis, transport and catabolism), and T (signal transduction mechanisms) showed higher abundances in the intracellular microbiome when compared against each of the other three microbiomes (Fig. 9c). In addition, the composition and relative abundance of each COG for the four microbial communities were distinguishable between each other (Fig. S3). To link the functions with their microbial origins, the taxa reported as lipase-producing bacteria were obtained by searching the eggNOGs database (Supplemental file 4). Among these identified bacterial genera, *Pseudomonas* was found to be unique to the sponge intracellular microbiome (Supplemental file 4) and showed consistently a higher abundance compared to other three microbiomes in the pairwise significant-different analysis (Fig. 9d).

DISCUSSION

Our study validated that the strategy of directly sequencing different cell fractions of sponge tissue, in particular the sponge stem cell population, could offer a more comprehensive profile of sponge microbiomes. Using the sponge *Euryspongia arenaria*
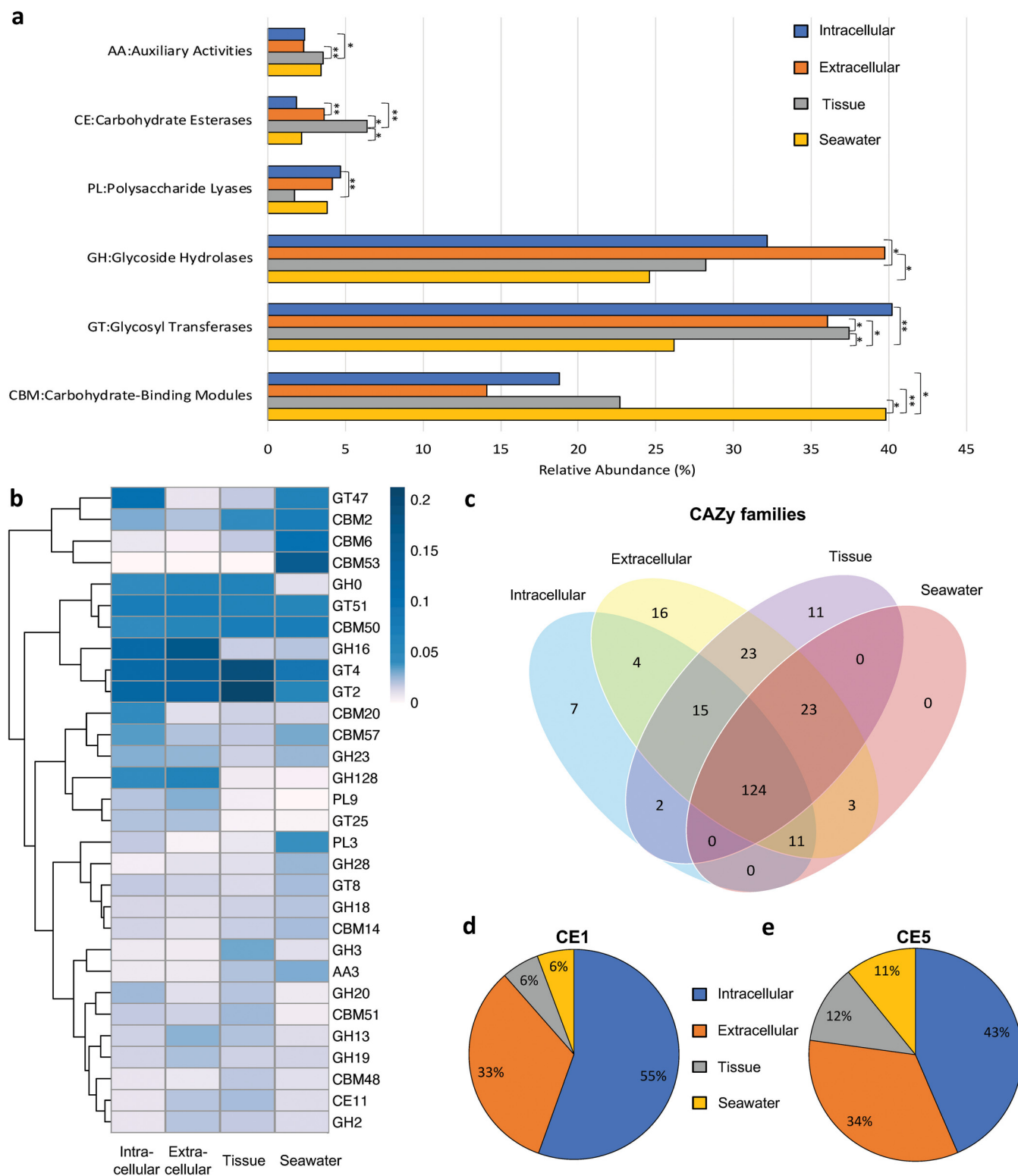
FIG 8 Carbohydrate-degrading enzymes (CAZy) comparison among the four microbiomes revealed by metagenomic sequencing. (a) Relative abundance of the unigenes assigned to six carbohydrate active enzyme classes. (b) Heatmap depicting the distribution of the top dominate 30 CAZy families belonging to the classes. Clustering is based on the relative abundance of each family. (c) Venn diagram depicting the numbers of CAZy families in each sample. (d) Distribution of CAZy family CE1. (e) Distribution of CAZy family CE5.

FIG 9 Significantly differential analysis of the functional and taxonomic features among the four microbiomes. (a) KEGG level three pathways comparison; ko 03018, RNA degradation; ko 04972, pancreatic secretion; ko 03070, bacterial secretion system; Ko 04013, MAPK signaling pathway; Ko 00920, sulfur metabolism; ko 01503, cationic antimicrobial peptide resistance; ko 05012, Parkinson's disease; ko 00280, Valine, leucine, and isoleucine degradation; ko 00270, cysteine and methionine metabolism; ko 00330, arginine and proline metabolism; ko 01503, cationic antimicrobial peptide (CAMP) resistance; ko 04932, non-alcoholic fatty liver disease (NAFLD); ko 03440, homologous recombination. (b) Carbohydrate-degrading enzymes family level comparison. (c) EggNOG categories (COGs) comparison; C, energy production and conversion; L, replication, recombination, and repair; P, inorganic ion transport and metabolism; I, lipid transport and metabolism; E, amino acid transport and metabolism; H, coenzyme transport and metabolism; T, signal transduction mechanisms; O, posttranslational modification, protein turnover, chaperones; Q, secondary metabolites biosynthesis, transport, and catabolism. (d) Genus-level taxa comparison. The comparison is conducted between each other among the four samples. The top three results with the smallest *P*-value (< 0.05) are presented.

as a model, we found the endosymbiotic microbiome hosted unique taxa from the phylum to genus level compared to the extracellular, tissue, and surrounding seawater microbiomes. The functional specificity of endosymbiotic microbiota was demonstrated by their unique pathways and enzyme families.

To profile the sponge endosymbiotic microbiome using community DNA extraction and next generation sequencing, it is critical to obtain a pure sponge cell fraction. Therefore, for different sponge species the most challenging step is to develop a purification strategy specific to the composition of that sponge. Typically, differential centrifugation is a well-developed method to separate the cells based on their different relative densities (68–70). Density gradient centrifugation, usually with Ficoll as the medium, has also been commonly used in the fine-scale sponge cell separation (71–80). It was noted in our study that the cells of *E. arenaria* were highly sensitive to the dissociation conditions and centrifugation (Fig. S1). A cell purification protocol (Fig. S4a) was developed and validated after optimization tailored to the properties of different *E. arenaria* cell types. Cell aggregation (81–84), a unique characteristic of sponge cells, was additionally employed as an efficient approach to further enrich cells by excluding damaged sponge cells. The formation of *E. arenaria* primmorphs thereby selected for healthy sponge cells and removed the contaminants or heterogenous cells, such as dead and broken sponge cells (Fig. S4b). This combined *in vitro* cultivation-differential centrifugation strategy greatly improved the cell purification efficiency (Fig. 1a to i) and produced two target cell fractions (Fig. 1j to m). The purification efficiency was further validated by the statistically consistent Raman spectra of individual cells within each fraction (Fig. 3), demonstrating their chemical homogeneity. TEM examination eventually distinguished and identified the cell types in our study (Fig. 2a and b) and revealed the presence of endosymbiotic microbes in both archaeocytes (Fig. 2c) and choanocytes (Fig. 2d). Matching the reported morphological features of archaeocytes (85, 86), we also found that the cells enriched in fraction I are motile and relatively large and contain a large nucleus with a single large nucleolus. For the cells in fraction II, we identified choanocytes due to their most critical morphology feature, i.e., flagella (65, 86).

Unlike the traditional sponge microbiome investigations, in which DNA isolated directly from sponge tissues is sequenced, the strategy here provided a new insight into the completeness of the sponge microbiome. The comprehensive data obtained by investigating preseparated fractions of sponges greatly impacted the observed coverage of the microbial profile (Fig. 4e and Fig. 6b). The profile generated by bulk sequencing of the sponge tissue sample missed out 14% (three out of 21) of microbial phyla and 62.8% (91 out of 145) of the identifiable genera (Fig. 4c and d; Table S1b and c). This is qualitatively consistent with similar studies in which the enrichment of the microbial fraction in a metagenome has also been reported to improve the resolution of microbiome characterization (44, 56–58, 60, 87). However, these investigations did not focus on the comparison between the revealed microbiomes pre- and post-enrichment of the microbial fraction. Only one study on two sponges (*Cymbastela concentrica* and *Scopalina* sp.) (59) found that the microbial OTUs obtained by a sequencing of whole sponge tissue had only a small overlap with those revealed from the bacterial pellet, which is consistent with the finding in our study that parts of the sponge microbiome could be hidden depending on the sample processing strategy (Fig. 4d and Fig. 6d). Based on these findings, it seems likely that the substantial differences were due to a sample preparation approach that allows for more efficient recovery of DNA from the microbial community. When the generated microbiomes highly depend on DNA quality, a more focused sequencing from bacteria-enriched fractions (intra- and extracellular) can therefore offer more comprehensive sequencing results. Our study paves the way toward approaching a complete microbiome for any other marine sponges.

Most sponge symbiotic microorganisms inhabit the mesohyl tissue, which constitutes most of the sponge body and is formed of an extracellular matrix populated by sponge cells (88). Only through TEM have researchers gradually become aware of the

presence of intracellular bacterial symbionts in these tissues (86, 89–97). Consumption of symbiotic microorganisms has been proposed as a possible food source for sponges (98, 99), and indeed phagocytosis and subsequent intracellular digestion of bacteria are the presumed mechanisms of nutrient transfer between a carnivorous sponge *Cladorhiza* and its methanotrophic symbionts (100, 101). Calcification was reported to be mediated by endosymbiotic bacteria (calcibacteria) within archeocyte-like cells of sponges in genus *Hemimycale* (61–63). A recent study sorted sponge homogenate based on particle size and localized the production of defense chemicals to intracellular symbionts of *Haliclona* sponges (64).

In our study, we focused on two sponge cell types and conducted a tailored cell sorting and purification so as to produce a detailed profile of the microbiome both intra- and extracellularly using our recently developed efficient multi-primer amplicon sequencing (102) and shotgun metagenomics. The different types of enriched sponge cells hosted specific microbial taxa, extending from genus to phylum, (Fig. 4) that were distinct from the ones of the extracellular and tissue samples. Notably, the community comparison indicated a high similarity between two cell fractions (Fig. 4a and e), which also guided us to pool them together for the shotgun metagenomic analysis. The microbiome structural specificity was further explained by additionally comparing it with the seawater sample (Fig. 6). The unique taxa specific to the intracellular microbiome were identified with significantly increased taxa number by using metagenomes. Interestingly, Archaea were found almost exclusively in the intracellular fractions. In contrast, their virus abundances were much less than the other three microbiomes (Fig. 6b).

As is common with such studies, a substantial percentage of the sequences we obtained by amplicon or metagenome sequencing could not be classified taxonomically. In our study, only 34% of the 16S rRNA gene amplicons could be identified in the archaeocytes-enriched fraction, 33% in the choanocytes-enriched fraction, 20% in tissue, and 11% in extracellular bacteria fraction (Fig. 4c). Although amplicon sequencing provides less capacity to explore untapped microbial taxa from a complex microbiome, the advantage of the primer based 16S rRNA gene sequences is that they can be effectively BLASTed against comprehensive databases such as Greengenes to identify the microbial taxa. Each of the multiple representative sequences assigned to the same genus or family (Fig. 5) were re-classified to a deeper taxonomic level for a better taxonomic resolution (Data sets S2). The inferred taxa include the species belonging to Alpha-, Gamma-, and Epsilon-proteobacteria, as well as Bacteroidetes and Cyanobacteria, which could represent microbial resources for gene and metabolite discovery.

Sponge symbionts have been shown to be the actual producers of the diverse natural products with a range of important bioactivities (64, 103–106). Among these natural products producers, cultivated bacteria represent only a minute fraction and the uncultured majority are generally perceived as a large, untapped resource of new drug candidates. Sponges belonging to the genus *Euryspongia* have been shown to contain abundant secondary metabolites with various activities (15–24). However, there are no comprehensive metagenomic annotations for the microbiomes of the sponges *Euryspongia* to allow for a correlation analysis between the functions and their potential microbial origins. The presence of functional features specific to the intra- or extracellular microbiomes was demonstrated by KEGG pathway distribution (Fig. 7); in particular, the highest number of unique level-three pathways was identified in the intracellular microbiome (Fig. 7b). The functional specificity of the intracellular endosymbionts can be further distinguished using their carbohydrate-active enzymes (CAZys) profile (Fig. 8a to c). Moreover, the statistically differential comparison guided us to identify an KEGG pathway (ko 04972) and a CAZy family (CE10) specific to the endosymbiotic microbiome (Fig. 9a and b). The enrichment of the pathway ko04972, designated by KEGG as the pancreatic secretion pathway involved in digestive system, may be related to a high abundance of muscarinic acetylcholine receptor and cholecystokinin receptors (107–109), likely acquired by horizontal gene transfer and possibly required for the maintenance of an endosymbiotic lifestyle (110–112). Comparing the involved enzymes in the pathway with the ones belonging to family CE10, we found one

shared entry-triacylglycerol lipase (EC 3.1.1.3) that was applied as polyethylene terephthalate hydrolase (113). In addition, the eggNOGs v5.0.0 database (114) supports the correlation analysis to link the annotated functional COGs and enzymes with their microbial origins, which allowed us to understand which microbial taxon carries which specific function. By doing so, we identified 40 microbial species containing homologs of triacylglycerol lipase (EC 3.1.1.3) enzymes that matched with the unique taxa in intracellular microbiome (Supplemental file 4). Consistently, the predicted COGs (i.e., functional categories L, T, and I) for this lipase family (EC 3.1.1.3) were also found to be specific to the intracellular microbiome with relative high abundance (Fig. 9c). Moreover, among the matched 40 microbial taxa, *Pseudomonas* members were identified to be specific to the intracellular microbiome by statistical differential analysis at the microbial genus-level comparison (Fig. 9d). These findings provide a valuable guidance to conduct studies on the functional gene repertoire of intracellular microbiomes in diverse sponges.

In conclusion, sequencing pure sponge cell fractions by combining both amplicon and metagenomic data provided a comprehensive analysis of the *E. arenaria* microbiome and revealed for the first time the structural and functional specificity of the intracellular endosymbiotic microbiome compared to its extracellular, tissue, and seawater microbiomes. The sponge cell purification and extracellular bacteria enrichment could be considered as an approach for profiling a complete sponge microbiome, in particular for uncovering the low abundant and likely missed microbiota when sequencing the sponge bulk tissues. Metagenome-based annotations of various features revealed potential functions correlated to the unique intracellular microbiota. One promising finding on metagenome-annotated serine hydrolases families enriched within intracellular microbiome highlighted the sponge endosymbiotic taxa as an understudied source for novel functional bacteria resources mining. Additional downstream analyses, such as conducting a whole-genome assembly from metagenomes to identify gene clusters for natural products or applying single-cell sequencing to directly reveal the sponge endosymbionts, are still necessary to further our understanding and utilization of sponge intracellular endosymbionts.

MATERIALS AND METHODS

Sponge sample collection and identification. Sponge specimens ($n \geq 3$ in each of the two collections) were collected via scuba diving at depths of 4 to 15 meters at Rapid Bay, Adelaide, South Australia (35°31'16.6"S, 138°11'07.5"E) in January of 2018. The healthy sponge specimens were soaked in flowing natural seawater to clean the contaminations. Preparation for histological sections and spicule preparations followed the methods in "Sponguide" (115). The classification refers to the study for Demosponge classification (116). The specimens were identified as *Euryspongia arenaria* (Queensland Museum Registration No. G301355) (Class: Demospongiae; Order: Dictyoceratida; Family: Dysideidae) using morphological features (117). Seawater samples were collected in parallel with sponge sampling using 50 mL tubes ($n = 3$) and kept in $-80°C$ freezer for further analysis.

Sponge cells and extracellular bacteria enrichment. Three sponge specimens were soaked in sterile calcium/magnesium-free artificial seawater (CMFASW) for to remove natural seawater. The sponge tissues were then fixed using 2.5% glutaraldehyde in CMFASW. The cell dissociation protocol (Fig. S4a) was developed based on the study (79). The cell purification was conducted following the protocol shown in Figure S4b (79). Continuous Ficoll-CMFASW gradients with densities of 10%, 20%, 30%, and 40% (wt/vol %) were applied to separate the sponge cells. One cell fraction was collected between 20% and 30% Ficoll-CMFASW gradients; the other fraction was collected between 10% and 20% gradients. The pellets from high-speed centrifugation were kept for extracellular bacterial community analysis.

Sponge cell type characterization and endosymbiont identification. The morphology and structure of sponge cells in the two purified fractions were characterized by TEM based on a protocol (118) with optimizations. Briefly, ultrathin sections were obtained with the Ultramicrotome Leica EM UC7 and stained first with 2% uranyl acetate for 30 min and then with lead citrate for 15 min at room temperature. Observations were conducted using a FEI Tecnai G2 Spirit TEM with Olympus-SIS Veleta CCD camera. Analysis was conducted via an EDS system comprising an Apollo XLT SDD running EDAX's TEAM software.

Raman microspectroscopy. *In situ* measurement of chemical composition within each single sponge cell was recorded with a WITec Alpha 300R Raman microscope (WITec Instrument Corp., Germany) equipped with CCD detector (119, 120). The confocal Raman spectra were generated following the instrument operating procedure. 10 μL suspension of purified sponge cell fraction was dropped onto middle of a glass slide and then covered by a cover slip. Five slides (five subfractions) were prepared for each cell fraction. Microscope was adjusted manually for focusing before turning on the Raman laser. Five investigation areas were selected for each slide for data collection: (i) top left corner; (ii) top right corner; (iii) middle; (iv) bottom left corner; and (v) bottom right corner. Four sponge cells in

each area were selected randomly to profile single Raman spectra using 532 nm laser after manual validation of the laser intensity.

DNA extraction. The DNA extraction protocol utilized for sponge tissue is the Cetyltrimethylammonium bromide (CTAB)-based method (39, 121) with the modifications described in Yang et al. (122). For purified sponge cell fractions and enriched bacteria fraction, the extraction was conducted by an extraction kit Quick-DNA/RNA Microprep Plus Kit (ZYMO Research, USA, Cat. No. D7005). Three biological replicates for each of the five types of the samples (sponge tissue, two cell fractions, extracellular bacteria sample, and seawater sample) were applied. Purity and quantity of DNA were determined with a NanoDrop ND-1000 Spectrophotometer (Thermo Fisher Scientific, Wilmington, DE, USA). Each sample was extracted in triplicate. The qualified DNA samples (A260/280: 1.8-2.0; Conc. > 100 ng/μL) extracted from three replicates for each sample type were selected and kept at $-20°C$ for subsequent PCRs and sequencing.

Illumina MiSeq amplicon library and sequencing. A multiprimer based amplicon sequencing approach, which was developed and validated in our previous studies (102, 123), was utilized. Three primer sets (28F-519R, 518F-926R, and 803F-1392R) were employed for 16S rRNA gene region V1 to V8. For each primer set, negative controls and multiple amplifications ($n > 3$) were applied for each of the three qualified DNA samples extracted from different sample types. Each pooled amplicon was sequenced multiple times ($n \geq 3$) from both ends of paired-end library preparations (2 $\times$ 300 bp) using sequencing kit version 3.0. The raw data was processed by base-calling using Illumina Genome Analyzer Pipeline software (GAPipeline version 1.4.0.). The DNA standard provided by ZymoBIOMIC was employed as the quality control (both cellular standard and DNA standard) for each sequencing run (Fig. S5).

16S rRNA gene sequencing data processing and analysis. The demultiplexing and quality filter (at Phred $\geq$ Q20) was processed by script split_libraries.py in QIIME pipeline (version 1.9.1) (124). The primer sequences and the poor quality sequences were trimmed by the software Trimmomatic (version 0.35) (125). Chimeric sequences were removed by the USEARCH tool (126). A minimum of 99,864 filtered sequence reads per primer set was collected for each biological replicate of same sample type. The sequencing reads generated by three primer sets were jointly applied based on the previous methodology validation (102).

Closed-reference picking was selected in our study (124). Reads were assigned to OTUs based on their best hit to Greengenes database (version: gg_13_8_otus/rep_set/97_otus.fasta) (127) at greater than or equal to 97% sequence identity (128). Greengenes was selected as the reference gene database due to its intermediate number of the sequences and sufficient taxa coverage at each rank from phylum to genus compared to SILVA and RDP (129, 130). While SILVA is a comprehensive database, the risk of overestimating the diversity is high due to the highest number of false-positives (130). In addition, the multi-primer approach applied in this study has demonstrated powerful performance on increasing the revealed microbial taxa coverage and the classification accuracy (102, 123). Hence, our analysis was based on Greengenes database. The uclust (126) consensus taxonomy classifier was applied to assign taxonomy to each OTU. The taxonomy metadata was further conducted by ClustVis to visualize the microbial community heatmap (67) with the implementation of R package of "pheatmap." In the heatmap, both rows and columns are clustered using "correlation" distance measurement and "complete linkage" clustering method. The rows are centered, and the unit scaling is applied. Unique-sequence abundance heatmap was built directly from the abundance of unique amplicon sequences inferred from raw sequencing data. To select sequence variants, DADA2 plugin (131) was implemented in QIIME2 pipeline (https://qiime2.org) to generate the exact "sequence variants" based on statistical error correction model. The involved unique sequences were exported as FASTA format data to conduct BLAST search for subspecies identification.

Illumina HiSeq metagenomic library preparation and sequencing. Two purified sponge cell fractions I and II were pooled for metagenomic sequencing. Library preparations were constructed following the manufacturer's protocol (VAHTS Universal DNA Library Prep Kit for Illumina). Replicates were applied. The PCR products were cleaned up using VAHTSTM DNA Clean Beads, validated using an Agilent 2100 Bioanalyzer (Agilent Technologies, Palo Alto, CA, USA), and quantified by Qubit 3.0 Fluorometer (Invitrogen, Carlsbad, CA, USA). Then libraries with different indices were multiplexed and loaded on an Illumina HiSeq instrument according to manufacturer's instructions (Illumina, San Diego, CA, USA). Sequencing was carried out using a 2 $\times$ 150 paired-end configuration on Illumina HiSeq instrument.

Metagenomic data processing and analysis. The raw reads from metagenome sequencing were processed by cutadapt (v2.0) (132) to remove adaptor sequences and low-quality reads. The reads were then aligned to the host genome using BWA (v0.7.13) (133, 134) to remove host contamination. The quality optimized clean reads were then processed by MEGAHIT (v1.1.3) (135) for *de novo* assembly. Different K-mers (39, 59, 79, 119) were tested, and the value was selected based on the largest N50 of the assembled scaffolds. Contigs (continuous sequences within scaffolds) with more than 200 bp were retained in the final assembly for gene prediction. The metagenes were predicted using Prodigal (v3.02) (136). CD-HIT (v2.21) (137) was applied to reduce sequence redundancy and cluster the metagenes with the parameters of minimum identity 0.95 and coverage over 0.9. Approximate 2.4 million non-redundant metagenes (unigenes) were obtained.

To generate the taxonomic information of the unigenes, the sequences were blasted against the NCBI microbial nr (non-redundant protein) database using DIAMOND (v0.8.15.77) (138) with an E value $< 1e^{-5}$. Based on the lowest common ancestor algorithm in MEGAN (v6.4.4) (139), the taxonomic identities of each unigene were assigned. The sequences were further blasted against the KEGG database (for KOs and pathways), CAZy database (for carbohydrate-active enzymes), and eggNOG database (for COGs and their host organisms) using DIAMOND with E value $< 1e^{-5}$. For the taxonomic and functional abundance

profiling, the reads belonging to different samples were mapped separately against all the unigenes using SOAPAligner (140) with the default setting.

To determine the similarity or difference of taxonomic components between different samples, clustered heatmap was constructed based on the relative abundance of each unigene at the genus level. The dendrogram was generated by R (pheatmap package) using *complete* method based on *bray* distance. T-test was employed to indicate the significant-difference pairs for CAZy abundance annotation. The statistical difference comparison among the samples were conducted by Metastats (141) to identify the taxonomic and functional features.

Data availability. Raw 16S rRNA gene sequences, the metagenomic sequencing reads and assembly have been uploaded to the NCBI Sequence Read Archive (SRA) data depository, with the BioProject accession number PRJNA699379.

SUPPLEMENTAL MATERIAL

Supplemental material is available online only.
SUPPLEMENTAL FILE 1, XLSX file, 1.4 MB.
SUPPLEMENTAL FILE 2, XLSX file, 19 MB.
SUPPLEMENTAL FILE 3, XLSX file, 0.1 MB.
SUPPLEMENTAL FILE 4, XLSX file, 0.01 MB.
SUPPLEMENTAL FILE 5, PDF file, 2.7 MB.

ACKNOWLEDGMENTS

REFERENCES

1. Feuda R, Dohrmann M, Pett W, Philippe H, Rota-Stabelli O, Lartillot N, Wörheide G, Pisani D. 2017. Improved modeling of compositional heterogeneity supports sponges as sister to all other animals. Curr Biol 27:1–7.

2. de Voogd NJ, Alvarez B, Boury-Esnault N, Carballo JL, Cárdenas P, Díaz M-C, Dohrmann M, Downey R, Hajdu E, Hooper JNA, Kelly M, Klautau M, Manconi R, Morrow CC, Pisera AB, Ríos P, Rützler K, Schönberg C, Vacelet J, van Soest RWM. 2022. World Porifera database. https://www.marinespecies.org/porifera.

3. Mehbub MF, Lei J, Franco CMM, Zhang W. 2014. Marine sponge derived natural products between 2001 and 2010: trends and opportunities for discovery of bioactives. Mar Drugs 12:4539–4577. https://doi.org/10.3390/md12084539.

4. Blunt JW, Copp BR, Keyzers RA, Munro MHG, Prinsep MR. 2013. Marine natural products. Nat Prod Rep 30:237–323. https://doi.org/10.1039/c2np20112g.

5. Blunt JW, Copp BR, Keyzers RA, Munro MHG, Prinsep MR. 2014. Marine natural products. Nat Prod Rep 31:160–258. https://doi.org/10.1039/c3np70117d.

6. Blunt JW, Copp BR, Keyzers RA, Munro MHG, Prinsep MR. 2015. Marine natural products. Nat Prod Rep 32:116–211. https://doi.org/10.1039/c4np00144c.

7. Blunt JW, Copp BR, Keyzers RA, Munro MHG, Prinsep MR. 2016. Marine natural products. Nat Prod Rep 33:382–431. https://doi.org/10.1039/c5np00156k.

8. Blunt JW, Copp BR, Keyzers RA, Munro MHG, Prinsep MR. 2017. Marine natural products. Nat Prod Rep 34:235–294. https://doi.org/10.1039/c6np00124f.

9. Blunt JW, Carroll AR, Copp BR, Davis RA, Keyzers RA, Prinsep MR. 2018. Marine natural products. Nat Prod Rep 35:8–53. https://doi.org/10.1039/C7NP00052A.

10. Carroll AR, Copp BR, Davis RA, Keyzers RA, Prinsep MR. 2019. Marine natural products. Nat Prod Rep 36:122–173. https://doi.org/10.1039/c8np00092a.

11. Carroll AR, Copp BR, Davis RA, Keyzers RA, Prinsep MR. 2020. Marine natural products. Nat Prod Rep 37:175–223. https://doi.org/10.1039/c9np00069k.

12. Carroll AR, Copp BR, Davis RA, Keyzers RA, Prinsep MR. 2021. Marine natural products. Nat Prod Rep 38:362–413. https://doi.org/10.1039/d0np00089b.

13. Carroll AR, Copp BR, Davis RA, Keyzers RA, Prinsep MR. 2022. Marine natural products. Nat Prod Rep. https://doi.org/10.1039/D1NP00076D.

14. Hall LD, Sanders JKM. 1980. Complete analysis of ^{1}H NMR spectra of complex natural products using a combination of one- and two-dimensional techniques. 1-Dehydrotestosterone. J Am Chem Soc 102:5703–5711. https://doi.org/10.1021/ja00538a001.

15. van Altena IA, Butler AJ, Dunne SJ. 1999. A new cyclized 9,11-secosterol enol−ether from the Australian sponge *Euryspongia arenaria*. J Nat Prod 62:1154–1157. https://doi.org/10.1021/np9805591.

16. Boonlarppradab C, Faulkner DJ. 2007. Eurysterols A and B, cytotoxic and antifungal steroidal sulfates from a marine sponge of the genus *Euryspongia*. J Nat Prod 70:846–848. https://doi.org/10.1021/np060472c.

17. Dopeso J, Quiñoá E, Riguera R, Debitus C, Bergquist PR. 1994. Euryspongiols: ten new highly hydroxylated 9,11-secosteroids with antihistaminic activity from the sponge *Euryspongia* sp. stereochemistry and reduction. Tetrahedron 50:3813–3828. https://doi.org/10.1016/S0040-4020(01)90401-6.

18. Hallock YF, Cardellina JH, Boyd MR. 1998. (-)-Frondosins A and D, HIV-inhibitory sesquiterpene hydroquinone derivatives from *Euryspongia* sp. Nat Prod Lett 11:153–160. https://doi.org/10.1080/10575639808041212.

19. Urban S, Capon R. 1996. Deoxyspongiaquinones: new sesquiterpene quinones and hydroquinones from a Southern Australian marine sponge *Euryspongia* sp. Aust J Chem 49:611–615. https://doi.org/10.1071/CH9960611.

20. Hochlowski JE, Walker RP, Ireland C, Faulkner DJ. 1982. Metabolites of four nudibranchs of the genus *Hypselodoris*. J Org Chem 47:88–91. https://doi.org/10.1021/jo00340a018.

21. Vanaltena I, Miller D. 1989. A study of the terpenoids of the sponge *Euryspongia* sp. Aust J Chem 42:2181–2190. https://doi.org/10.1071/CH9892181.

22. Clark RJ, Garson MJ, Brereton IM, Kennedy JA. 1999. Vinylfurans revisited: a new sesquiterpene from *Euryspongia deliculata*. J Nat Prod 62:915–916. https://doi.org/10.1021/np980557g.

23. Yamazaki H, Takahashi O, Kanno S-i, Nakazawa T, Takahashi S, Ukai K, Sumilat DA, Ishikawa M, Namikoshi M. 2015. Absolute structures and bioactivities of euryspongins and eurydiene obtained from the marine sponge *Euryspongia* sp. collected at Iriomote Island. Bioorg Med Chem 23:797–802. https://doi.org/10.1016/j.bmc.2014.12.049.

24. Guzmán EA, Pitts TP, Diaz MC, Wright AE. 2019. The marine natural product Scalarin inhibits the receptor for advanced glycation end products (RAGE) and autophagy in the PANC-1 and MIA PaCa-2 pancreatic cancer cell lines. Invest New Drugs 37:262–270. https://doi.org/10.1007/s10637-018-0635-4.

25. Thomas T, Moitinho-Silva L, Lurgi M, Björk JR, Easson C, Astudillo-García C, Olson JB, Erwin PM, López-Legentil S, Luter H, Chaves-Fonnegra A, Costa R, Schupp PJ, Steindler L, Erpenbeck D, Gilbert J, Knight R, Ackermann G, Lopez JV, Taylor MW, Thacker RW, Montoya JM, Hentschel U, Webster NS. 2016. Diversity, structure and convergent evolution of the global sponge microbiome. Nat Commun 7:11870. https://doi.org/10.1038/ncomms11870.

26. Webster NS, Thomas T. 2016. The sponge hologenome. mBio 7: e00135-16–e00116. https://doi.org/10.1128/mBio.00135-16.

27. Moitinho-Silva L, Nielsen S, Amir A, Gonzalez A, Ackermann GL, Cerrano C, Astudillo-Garcia C, Easson C, Sipkema D, Liu F, Steinert G, Kotoulas G, McCormack GP, Feng G, Bell JJ, Vicente J, Björk JR, Montoya JM, Olson JB, Reveillaud J, Steindler L, Pineda MC, Marra MV, Ilan M, Taylor MM, Polymenakou P, Erwin PM, Schupp PJ, Simister RL, Knight R, Thacker RW, Costa R, Hill RT, Lopez-Legentil S, Dailianis T, Ravasi T, Hentschel U, Li Z, Webster NS, Thomas T. 2017. The sponge microbiome project. GigaScience 6:1–7. https://doi.org/10.1093/gigascience/gix077.

28. Pita L, Rix L, Slaby BM, Franke A, Hentschel U. 2018. The sponge holobiont in a changing ocean: from microbes to ecosystems. Microbiome 6: 18. https://doi.org/10.1186/s40168-018-0428-1.

29. Hardoim C, Ramaglia A, Lôbo-Hajdu G, Custódio M. 2021. Community composition and functional prediction of prokaryotes associated with sympatric sponge species of southwestern Atlantic coast. Sci Rep 11:16. https://doi.org/10.1038/s41598-021-88288-3.

30. Ribeiro B, Padua A, Oliveira BFRd, Puccinelli G, da Costa Fernandes F, Laport MS, Klautau M. 2022. Uncovering the microbial diversity of two exotic calcareous sponges. Microb Ecol. https://doi.org/10.1007/s00248-022-01980-3.

31. Hentschel U, Hopke J, Horn M, Friedrich AB, Wagner M, Hacker J, Moore BS. 2002. Molecular evidence for a uniform microbial community in sponges from different oceans. Appl Environ Microbiol 68:4431–4440. https://doi.org/10.1128/AEM.68.9.4431-4440.2002.

32. Taylor MW, Tsai P, Simister RL, Deines P, Botte E, Ericson G, Schmitt S, Webster NS. 2013. 'Sponge-specific' bacteria are widespread (but rare) in diverse marine environments. ISME J 7:438–443. https://doi.org/10.1038/ismej.2012.111.

33. de Voogd NJ, Cleary DFR, Polónia ARM, Gomes NCM. 2015. Bacterial community composition and predicted functional ecology of sponges, sediment and seawater from the thousand islands reef complex, West Java, Indonesia. FEMS Microbiol Ecol 91:fiv019.

34. Indraningrat AAG, Hauke S, Detmer S. 2016. Bioprospecting sponge-associated microbes for antimicrobial compounds. Mar Drugs 14:87. https://doi.org/10.3390/md14050087.

35. Brinkmann CM, Marker A, Kurtböke DI. 2017. An overview on marine sponge-symbiotic bacteria as unexhausted sources for natural product discovery. Diversity 9:40–51. https://doi.org/10.3390/d9040040.

36. Bhushan A, Peters EE, Piel J. 2017. Entotheonella bacteria as source of sponge-derived natural products: Opportunities for biotechnological production. Prog Mol Subcell Biol 55:291–314. https://doi.org/10.1007/978-3-319-51284-6_9.

37. Agarwal V, Blanton JM, Podell S, Taton A, Schorn MA, Busch J, Lin Z, Schmidt EW, Jensen PR, Paul VJ, Biggs JS, Golden JW, Allen EE, Moore BS. 2017. Metagenomic discovery of polybrominated diphenyl ether biosynthesis by marine sponges. Nat Chem Biol 13:537–543. https://doi.org/10.1038/nchembio.2330.

38. Geahchan S, Ehrlich H, Rahman MA. 2021. The anti-viral applications of marine resources for COVID-19 treatment: an overview. Drugs 19:409. https://doi.org/10.3390/md19080409.

39. Schmitt S, Tsai P, Bell J, Fromont J, Ilan M, Lindquist N, Perez T, Rodrigo A, Schupp PJ, Vacelet J, Webster N, Hentschel U, Taylor MW. 2012. Assessing the complex sponge microbiota: core, variable and species-specific bacterial communities in marine sponges. ISME J 6:564–576. https://doi.org/10.1038/ismej.2011.116.

40. Schmitt S, Hentschel U, Taylor M. 2012. Deep sequencing reveals diversity and community structure of complex microbiota in five Mediterranean sponges. Hydrobiologia 687:341–351. https://doi.org/10.1007/s10750-011-0799-9.

41. Reveillaud J, Maignien L, Eren AM, Huber JA, Apprill A, Sogin ML, Vanreusel A. 2014. Host-specificity among abundant and rare taxa in the sponge microbiome. ISME J 8:1198–1209. https://doi.org/10.1038/ismej.2013.227.

42. Rodríguez-Marconi S, De la Iglesia R, Díez B, Fonseca CA, Hajdu E, Trefault N. 2015. Characterization of bacterial, archaeal and eukaryote symbionts from Antarctic sponges reveals a high diversity at a three-domain level and a particular signature for this ecosystem. PLoS One 10: e0138837. https://doi.org/10.1371/journal.pone.0138837.

43. Cristi A, Parada-Pozo G, Morales-Vicencio F, Cárdenas CA, Trefault N. 2021. Variability in host specificity and functional potential of Antarctic sponge-associated bacterial communities. Front Microbiol 12:771589. https://doi.org/10.3389/fmicb.2021.771589.

44. Botté ES, Nielsen S, Wahab MAA, Webster J, Robbins S, Thomas T, Webster NS. 2019. Changes in the metabolic potential of the sponge microbiome under ocean acidification. Nat Commun 10:4134. https://doi.org/10.1038/s41467-019-12156-y.

45. Moreno-Pino M, Cristi A, Gillooly JF, Trefault N. 2020. Characterizing the microbiomes of Antarctic sponges: a functional metagenomic approach. Sci Rep 10:645. https://doi.org/10.1038/s41598-020-57464-2.

46. Morganti TM, Slaby BM, de Kluijver A, Busch K, Hentschel U, Middelburg JJ, Grotheer H, Mollenhauer G, Dannheim J, Rapp HT, Purser A, Boetius A. 2022. Giant sponge grounds of Central Arctic seamounts are associated with extinct seep life. Nat Commun 13:15. https://doi.org/10.1038/s41467-022-28129-7.

47. Knight R, Vrbanac A, Taylor BC, Aksenov A, Callewaert C, Debelius J, Gonzalez A, Kosciolek T, McCall L, McDonald D, Melnik AV, Morton JT, Navas J, Quinn RA, Sanders JG, Swafford AD, Thompson LR, Tripathi A, Xu ZZ, Zaneveld JR, Zhu Q, Caporaso JG, Dorrestein PC. 2018. Best practices for analysing microbiomes. Nat Rev Microbiol 16:410–422. https://doi.org/10.1038/s41579-018-0029-9.

48. Quince C, Walker AW, Simpson JT, Loman NJ, Segata N. 2017. Shotgun metagenomics, from sampling to analysis. Nat Biotechnol 35:833–844. https://doi.org/10.1038/nbt.3935.

49. Whelan FJ, Waddell B, Syed SA, Shekarriz S, Rabin HR, Parkins MD, Surette MG. 2020. Culture-enriched metagenomic sequencing enables in-depth profiling of the cystic fibrosis lung microbiota. Nat Microbiol 5: 379–390. https://doi.org/10.1038/s41564-019-0643-y.

50. White JR, Patel J, Ottesen A, Arce G, Blackwelder P, Lopez JV. 2012. Pyrosequencing of bacterial symbionts within *Axinella corrugata* sponges: diversity and seasonal variability. PLoS One 7:e38204. https://doi.org/10.1371/journal.pone.0038204.

51. Burgsdorf I, Erwin PM, López-Legentil S, Cerrano C, Haber M, Frenk S, Steindler L. 2014. Biogeography rather than association with cyanobacteria structures symbiotic microbial communities in the marine sponge *Petrosia ficiformis*. Front Microbiol 10:1–11.

52. Luter HM, Widder S, Botté ES, Wahab MA, Whalan S, Moitinho-Silva L, Thomas T, Webster NS. 2015. Biogeographic variation in the microbiome of the ecologically important sponge, *Carteriospongia foliascens*. PeerJ 3: e1435. https://doi.org/10.7717/peerj.1435.

53. Weigel BL, Erwin PM. 2016. Intraspecific variation in microbial symbiont communities of the sun sponge, *Hymeniacidon heliophila*, from intertidal and subtidal habitats. Appl Environ Microbiol 82:650–658. https://doi .org/10.1128/AEM.02980-15.

54. Yang Q, Zhang W, Franco CMM. 2019. Response of sponge microbiomes to environmental variations, p 181–247. *In* Li Z (ed), Symbiotic microbiomes of coral reefs sponges and corals. Springer, Dordrecht, the Netherlands.

55. Steffen K, Indraningrat AAG, Erngren I, Haglöf J, Becking LE, Smidt H, Yashayaev I, Kenchington E, Pettersson C, Cárdenas P, Sipkema D. 2022. Oceanographic setting influences the prokaryotic community and metab-olome in deep-sea sponges. Sci Rep 12:16. https://doi.org/10.1038/s41598 -022-07292-3.

56. Engelberts JP, Robbins SJ, de Goeij JM, Aranda M, Bell SC, Webster NS. 2020. Characterization of a sponge microbiome using an integrative ge-nome-centric approach. ISME J 14:1100–1110. https://doi.org/10.1038/ s41396-020-0591-9.

57. Thomas T, Rusch D, DeMaere MZ, Yung PY, Lewis M, Halpern A, Heidelberg KB, Egan S, Steinberg PD, Kjelleberg S. 2010. Functional genomic signatures of sponge bacteria reveal unique and shared fea-tures of symbiosis. ISME J 4:1557–1567. https://doi.org/10.1038/ismej .2010.74.

58. Bayer K, Jahn MT, Slaby BM, Moitinho-Silva L, Hentschel U. 2018. Marine sponges as chloroflexi hot spots: genomic insights and high-resolution visualization of an abundant and diverse symbiotic clade. mSystems 3. https://doi.org/10.1128/mSystems.00150-18.

59. Esteves AIS, Amer N, Nguyen M, Thomas T. 2016. Sample processing impacts the viability and cultivability of the sponge microbiome. Font Microbiol 7.

60. Karimi E, Ramos M, Gonçalves JMS, Xavier JR, Reis MP, Costa R. 2017. Comparative metagenomics reveals the distinctive adaptive features of the *Spongia officinalis* endosymbiotic consortium. Front Microbiol 8. https://doi.org/10.3389/fmicb.2017.02499.

61. Uriz MJ, Agell G, Blanquer A, Turon X, Casamayor EO. 2012. Endosymbi-otic calcifying bacteria: a new cue to the origin of calcification in Meta-zoa? Evolution 66:2993–2999. https://doi.org/10.1111/j.1558-5646.2012 .01676.x.

62. Garate L, Blanquer A, Uriz M-J. 2015. Calcareous spherules produced by intracellular symbiotic bacteria protect the sponge *Hemimycale colu-mella* from predation better than secondary metabolites. Mar Ecol Prog Ser 523:81–92. https://doi.org/10.3354/meps11196.

63. Garate L, Sureda J, Agell G, Uriz MJ. 2017. Endosymbiotic calcifying bac-teria across sponge species and oceans. Sci Rep 7:43674. https://doi.org/ 10.1038/srep43674.

64. Tianero MD, Balaich JN, Donia MS. 2019. Localized production of defence chemicals by intracellular symbionts of Haliclona sponges. Nat Microbiol 4:1149–1159. https://doi.org/10.1038/s41564-019-0415-8.

65. Funayama N. 2018. The cellular and molecular bases of the sponge stem cell systems underlying reproduction, homeostasis and regeneration. Int J Dev Biol 62:513–525. https://doi.org/10.1387/ijdb.180016nf.

66. Custodio MR, Prokic I, Steffen R, Koziol C, Borojevic R, Brümmer F, Nickel M, Müller WEG. 1998. Primmorphs generated from dissociated cells of the sponge Suberites domuncula: a model system for studies of cell pro-liferation and cell death. Mach Ageing Dev 105:45–59. https://doi.org/10 .1016/S0047-6374(98)00078-5.

67. Metsalu T, Vilo J. 2015. Clustvis: a web tool for visualizing clustering of multivariate data using principal component analysis and heatmap. Nucleic Acids Res 43:W566–570. https://doi.org/10.1093/nar/gkv468.

68. Bewley CA, Holland ND, Faulkner DJ. 1996. Two classes of the metabolites from *Theonella swinhoei* are localized in distinct populations of bacterial symbionts. Experientia 52:716–722. https://doi.org/10.1007/BF01925581.

69. Richelle-Maurer E, Braekman JC, De Kluijver MJ, Gomez R, van de Vyver G, van Soest RWM, Devijver C. 2001. Cellular location of (2R,3R,7Z)-2-ami-notetradec-7-ene-1,3-diol, a potent antimicrobial metabolite produced by the Caribbean sponge *Haliclona vansoesti*. Cell Tissue Res 306: 157–165. https://doi.org/10.1007/s004410100437.

70. Laroche M, Imperatore C, Grozdanov L, Costantino V, Mangoni A, Hentschel U, Fattorusso E. 2007. Cellular localization of secondary metab-olites isolated from the Caribbean sponge *Plakortis simplex*. Mar Biol 151: 1365–1373. https://doi.org/10.1007/s00227-006-0572-1.

71. John HA, Campo MS, Mackenzie AM, Kemp RB. 1971. Role of different sponge cell types in species specific cell aggregation. Nat New Biol 230: 126–128. https://doi.org/10.1038/newbio230126b0.

72. Pomponi SA, Willoughby R. 1993. Sponge cell culture for production of bioactive metabolites, p 396–340. *In* van Soest RWM, van Kempen TMG,

73. Braekman JC (ed), Sponges in time and space. Balkema, Rotterdam, Amsterdam.

73. Uriz MJ, Turon X, Galera J, Tur JM. 1996. New light on the cell location of avarol within the sponge *Dysidea avara* (Dendroceratida). Cell Tissue Res 285:519–527. https://doi.org/10.1007/s004410050668.

74. Flowers AE, Garson MJ, Webb RI, Dumdei EJ, Charan RD. 1998. Cellular origin of chlorinated diketopiperazines in the dictyoceratid sponge *Dysi-dea herbacea* (Keller). Cell Tissue Res 292:597–607. https://doi.org/10 .1007/s004410051089.

75. Garson MJ, Flowers AE, Webb RI, Charan RD, McCaffrey EJ. 1998. A sponge/ dinoflagellate association in the haplosclerid sponge *Haliclona* sp.: cellular origin of cytotoxic alkaloids by Percoll density gradient fractionation. Cell Tissue Res 293:365–373. https://doi.org/10.1007/s004410051128.

76. Salomon CE, Deerinck T, Ellisman MH, Faulkner DJ. 2001. The cellular localization of dercitamide in the Palauan sponge *Oceanapia sagittaria*. Mar Biol 139:313–319.

77. Zhang XY, Cao XP, Zhang W, Yu XJ, Jin MF. 2003. Primmorphs from archaeocytes-dominant cell population of the sponge *Hymeniacidon per-leve*: improved cell proliferation and spiculogenesis. Biotechnol Bioeng 84:583–590. https://doi.org/10.1002/bit.10811.

78. Sun L-M, Song Y-F, Qu Y, Yu X-J, Zhang W. 2007. Purification and *in vitro* cultivation of archaeocytes (stem cells) of the marine sponge *Hymeniaci-don perleve* (Demospongiae). Cell Tissue Res 328:223–237. https://doi .org/10.1007/s00441-006-0342-x.

79. Song Y-F, Qu Y, Cao X-P, Zhang W. 2011. Cellular localization of debro-mohymenialdisine and hymenialdisine in the marine sponge *Axinella* sp. using a newly developed cell purification protocol. Mar Biotechnol (NY) 13:868–882. https://doi.org/10.1007/s10126-010-9347-2.

80. Mukherjee S, Ray M, Ray S. 2015. Phagocytic efficiency and cytotoxic responses of Indian freshwatersponge (*Eunapius carteri*) cells isolated by density gradientcentrifugation and flow cytometry: a morphofunctional analysis. Zoology (Jena) 118:8–18. https://doi.org/10.1016/j.zool.2014.07 .002.

81. Wilson HV. 1907. A new method by which sponges may be artificially reared. Science 25:912–915. https://doi.org/10.1126/science.25.649.912.

82. Eerkes-Medrano D, Feehan CJ, Leys SP. 2015. Sponge cell aggregation: checkpoints in development indicate a high level of organismal com-plexity. Invertebr Biol 134:1–18. https://doi.org/10.1111/ivb.12072.

83. Lavrov AI, Kosevich IA. 2014. Sponge cell reaggregation: mechanisms and dynamics of the process. Russ J Dev Biol 45:205–223. https://doi .org/10.1134/S1062360414040067.

84. Lavrov AI, Kosevich IA. 2016. Sponge cell reaggregation: cellular structure and morphogenetic potencies of multicellular aggregates. J Exp Zool A Ecol Genet Physiol 325:158–177. https://doi.org/10.1002/jez.2006.

85. Funayama N. 2010. The stem cell system in demosponges: insights into the origin of somatic stem cells. Dev Growth Differ 52:1–14. https://doi .org/10.1111/j.1440-169X.2009.01162.x.

86. Simpson TL. 1984. The cell biology of sponges. Springer, New York, NY.

87. Robbins SJ, Song W, Engelberts JP, Glasl B, Slaby BM, Boyd J, Marangon E, Botté ES, Laffy P, Thomas T, Webster NS. 2021. A genomic view of the microbiome of coral reef demosponges. ISME J 15:1641–1654. https:// doi.org/10.1038/s41396-020-00876-9.

88. Hentschel U, Piel J, Degnan SM, Taylor MW. 2012. Genomic insights into the marine sponge microbiome. Nat Rev Microbiol 10:641–654. https:// doi.org/10.1038/nrmicro2839.

89. Vacelet J. 1970. Description de cellules a bactéries intranucléaires chez des éponges Verongia. J Microsc 9:333–346.

90. Vacelet J. 1975. Ètude en microscopie électronique de l'association entre bactéries et spongiaires du genre *Verongia* (Dictyoceratida.). J Microsc Biol Cell 23:271–288.

91. Vacelet J, Donadey C. 1977. Electron microscope study of the association between some sponges and bacteria. J Exp Mar Biol Ecol 30:301–314. https://doi.org/10.1016/0022-0981(77)90038-7.

92. Wilkinson CR. 1978. Microbial associations in sponges. I. Ecology, physi-ology, and microbial populations of coral reef sponges. Mar Biol 49: 161–167. https://doi.org/10.1007/BF00387115.

93. Wilkinson CR. 1978. Microbial associations in sponges. II. Numerical anal-ysis of sponge and water bacterial populations. Mar Biol 49:169–176. https://doi.org/10.1007/BF00387116.

94. Wilkinson CR. 1978. Microbial associations in sponges. III. Ultrastructure of the *in situ* associations in coral reef sponges. Mar Biol 49:177–185. https://doi.org/10.1007/BF00387117.

95. Rützler K. 1990. Associations between Caribbean sponges and photosynthetic organisms, p 455–466. *In* Rützler K (ed), New perspectives in sponge biology. Smithsonian Institution Press, Washington, DC.

96. Maldonado M, Young CM. 1998. Limits on the bathymetric distribution of keratose sponges: a field test in deep water. Mar Ecol Prog Ser 174:123–139. https://doi.org/10.3354/meps174123.

97. Maldonado M, Cortadellas N, Trillas MI, Rützler K. 2005. Endosymbiotic yeast maternally transmitted in a marine sponge. Biol Bull 209:94–106. https://doi.org/10.2307/3593127.

98. Sara M. 1971. Ultrastructural aspects of the symbiosis between two species of the genus *Aphanocapsa* (Cyanophyceae) and *Ircinia variabilis* (Demospongiae). Mar Biol 11:214–221. https://doi.org/10.1007/BF00401270.

99. Sara M, Bavestrello G, Cattaneo-Vietti R, Cerrano C. 1998. Endosymbiosis in sponges: relevance for epigenesis and evolution. Symbiosis 25:57–70.

100. Vacelet J, Boury-Esnault N, Fiala-Medioni A, Fisher CR. 1995. A methanotrophic carnivorous sponge. Nature 377:296–296. https://doi.org/10.1038/377296a0.

101. Vacelet J, Fiala-Medioni A, Fisher CR, Boury-Esnault N. 1996. Symbiosis between methane-oxidizing bacteria and a deep-sea carnivorous cladorhizid sponge. Mar Ecol Prog Ser 145:77–85. https://doi.org/10.3354/meps145077.

102. Yang Q, Franco CMM, Zhang W. 2019. Uncovering the hidden marine sponge microbiome by applying a multi-primer approach. Sci Rep 9:6214. https://doi.org/10.1038/s41598-019-42694-w.

103. Wilson MC, Mori T, Rückert C, Uria AR, Helf MJ, Takada K, Gernert C, Steffens UAE, Heycke N, Schmitt S, Rinke C, Helfrich EJN, Brachmann AO, Gurgui C, Wakimoto T, Kracht M, Crüsemann M, Hentschel U, Abe I, Matsunaga S, Kalinowski J, Takeyama H, Piel J. 2014. An environmental bacterial taxon with a large and distinct metabolic repertoire. Nature 506:58–62. https://doi.org/10.1038/nature12959.

104. Mori T, Cahn JKB, Wilson MC, Meoded RA, Wiebach V, Martinez AFC, Helfrich EJN, Albersmeier A, Wibberg D, Dätwyler S, Keren R, Lavy A, Rückert C, Ilan M, Kalinowski J, Matsunaga S, Takeyama H, Piel J. 2018. Single-bacterial genomics validates rich and varied specialized metabolism of uncultivated *Entotheonella* sponge symbionts. Proc Natl Acad Sci U S A 115:1718–1723. https://doi.org/10.1073/pnas.1715496115.

105. Rust M, Helfrich EJN, Freeman MF, Nanudorn P, Field CM, Rückert C, Kündig T, Page MJ, Webb VL, Kalinowski J, Sunagawa S, Piel J. 2020. A multiproducer microbiome generates chemical diversity in the marine sponge *Mycale hentscheli*. Proc Natl Acad Sci U S A 117:9508–9518. https://doi.org/10.1073/pnas.1919245117.

106. Storey MA, Andreassend SK, Bracegirdle J, Brown A, Keyzers RA, Ackerley DF, Northcote PT, Owen JG. 2020. Metagenomic exploration of the marine sponge *Mycale hentscheli* uncovers multiple polyketide-producing bacterial symbionts. mBio 11:e02997-19. https://doi.org/10.1128/mBio.02997-19.

107. Wessler I, Kirkpatrick CJ. 2008. Acetylcholine beyond neurons: the non-neuronal cholinergic system in humans. Br J Pharmacol 154:1558–1571. https://doi.org/10.1038/bjp.2008.185.

108. Horiuchi Y, Kimura R, Kato N, Fujii T, Seki M, Endo T, Kato T, Kawashima K. 2003. Evolutional study on acetylcholine expression. Life Sci 72:1745–1756. https://doi.org/10.1016/S0024-3205(02)02478-5.

109. Wessler I, Kirkpatrick CJ, Racke K. 1999. The cholinergic 'pitfall': acetylcholine, a universal cell molecule in biological systems, including humans. Clin Exp Pharmacol Physiol 26:198–205. https://doi.org/10.1046/j.1440-1681.1999.03016.x.

110. Conaco C, Tsoulfas P, Sakarya O, Dolan A, Werren J, Kosik KS. 2016. Detection of prokaryotic genes in the *Amphimedon queenslandica* genome. PLoS One 11:e0151092. https://doi.org/10.1371/journal.pone.0151092.

111. Degnan SM. 2014. Think laterally: horizontal gene transfer from symbiotic microbes may extend the phenotype of marine sessile hosts. Front Microbiol 5:638. https://doi.org/10.3389/fmicb.2014.00638.

112. Jackson DJ, Macis L, Reitner J, Wörheide G. 2011. A horizontal gene transfer supported the evolution of an early metazoan biomineralization strategy. BMC Evol Biol 11:238. https://doi.org/10.1186/1471-2148-11-238.

113. Danso D, Schmeisser C, Chow J, Zimmermann W, Wei R, Leggewie C, Li X, Hazen T, Streita WR. 2018. New insights into the function and global distribution of Polyethylene Terephthalate (PET)-degrading bacteria and enzymes in marine and terrestrial metagenomes. Appl Enrivon Micorbiol 84:e02773-17.

114. Huerta-Cepas J, Szklarczyk D, Heller D, Hernández-Plaza A, Forslund SK, Cook H, Mende DR, Letunic I, Rattei T, Jensen LJ, von Mering C, Bork P. 2019. eggNOG 5.0: a hierarchical, functionally and phylogenetically annotated orthology resource based on 5090 organisms and 2502 viruses. Nucleic Acids Res 47:D309–D314. https://doi.org/10.1093/nar/gky1085.

115. Hooper JNA. 2003. Sponguide: guide to sponge collection and identification. Queensland Museum. Available at http://www.qm.qld.gov.au/.

116. Morrow C, Cárdenas P. 2015. Proposal for a revised classification of the Demospongiae (Porifera). Front Zool 12:1:7. https://doi.org/10.1186/s12983-015-0099-8.

117. Hooper JNA, van Soest RWM. 2002. Systema Porifera. A guide to the classification of sponges. Kluwer Academic/Plenum Publishers, New York, NY.

118. Maldonado M. 2007. Intergenerational transmission of symbiotic bacteria in oviparous and viviparous demosponges, with emphasis on intracytoplasmically-compartmented bacterial types. J Mar Biol Ass 87:1701–1713. https://doi.org/10.1017/S0025315407058080.

119. Maia LF, Fernandes RF, Lobo-Hajdu G, de Oliveira LFC. 2014. Conjugated polyenes as chemical probes of life signature: use of Raman spectroscopy to differentiate polyenic pigments. Philos Trans R Soc A 372:20140200. https://doi.org/10.1098/rsta.2014.0200.

120. Nehrke G, Nouet J. 2011. Confocal Raman microscope mapping as a tool to describe different mineral and organic phases at high spatial resolution within marine biogenic carbonates: case study on *Nerita undata* (Gastropoda, Neritopsina). Biogeosciences 8:3761–3769. https://doi.org/10.5194/bg-8-3761-2011.

121. Taylor MW, Schupp PJ, Dahllöf I, Kjelleberg S, Steinberg PD. 2004. Host specificity in marine sponge-associated bacteria, and potential implications for marine microbial diversity. Environ Microbiol 6:121–130. https://doi.org/10.1046/j.1462-2920.2003.00545.x.

122. Yang Q, Franco CMM, Zhang W. 2015. Sponge-associated actinobacterial diversity: validation of the methods of actinobacterial DNA extraction and optimization of 16S rRNA gene amplification. Appl Microbiol Biotechnol 99:8731–8740. https://doi.org/10.1007/s00253-015-6875-8.

123. Yang Q, Franco CMM, Lin H-W, Zhang W. 2019. Untapped sponge microbiomes: structure specificity at host order and family levels. FEMS Microbiol Ecol 95:fiz136. https://doi.org/10.1093/femsec/fiz136.

124. Caporaso JG, Kuczynski J, Stombaugh J, Bittinger k, Bushman FD, Costello EK, Fierer N, Peña AG, G JK, Gordon JI, Huttley GA, Kelley ST, Knights D, Koenig JE, Ley RE, Lozupone CA, McDonald D, Muegge BD, Pirrung M, Reeder J, Sevinsky JR, Turnbaugh PJ, Walters WA, Widmann J, Yatsunenko T, Zaneveld J, Knight R. 2010. QIIME allows analysis of high-throughput community sequencing data. Nat Methods 7:335–336. https://doi.org/10.1038/nmeth.f.303.

125. Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 30:2114–2120. https://doi.org/10.1093/bioinformatics/btu170.

126. Edgar RC. 2010. Search and clustering orders of magnitude faster than BLAST. Bioinformatics 26:2460–2461. https://doi.org/10.1093/bioinformatics/btq461.

127. DeSantis TZ, Hugenholtz P, Larsen N, Rojas M, Brodie EL, Keller K, Huber T, Dalevi D, Hu P, A GL. 2006. Greengenes, a chimera-checked 16S rRNA gene database and workbench compatible with ARB. Appl Environ Microbiol 72:5069–5072. https://doi.org/10.1128/AEM.03006-05.

128. Chen W, Zhang CK, Cheng Y, Zhang S, Zhao H. 2013. A comparison of methods for clustering 16S rRNA sequences into OTUs. PLoS One 8:e70837. https://doi.org/10.1371/journal.pone.0070837.

129. Balvočiūtė M, Huson DH. 2017. SILVA, RDP, Greengenes, NCBI and OTT–how do these taxonomies compare? BMC Genom 18:114. https://doi.org/10.1186/s12864-017-3501-4.

130. Park S-C, Won S. 2018. Evaluation of 16S rRNA databases for taxonomic assignments using a mock community. Genomics Inform 16:e24. https://doi.org/10.5808/GI.2018.16.4.e24.

131. Callahan BJ, McMurdie PJ, Rosen MJ, Han AW, Johnson AJA, Holmes SP. 2016. DADA2: high-resolution sample inference from Illumina amplicon data. Nat Methods 13:581–583. https://doi.org/10.1038/nmeth.3869.

132. Martin M. 2011. Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet j 17:10–12. https://doi.org/10.14806/ej.17.1.200.

133. Li H, Durbin R. 2009. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics 25:1754–1760. https://doi.org/10.1093/bioinformatics/btp324.

134. Li H, Durbin R. 2010. Fast and accurate long-read alignment with Burrows-Wheeler transform. Bioinformatics 26:589–595. https://doi.org/10.1093/bioinformatics/btp698.

135. Li D, Luo R, Liu CM, Leung CM, Ting HF, Sadakane K, Yamashita H, Lam TW. 2016. MEGAHIT v1.0: a fast and scalable metagenome assembler driven by advanced methodologies and community practices. Methods 102:3–11. https://doi.org/10.1016/j.ymeth.2016.02.020.

136. Hyatt D, Chen G, LoCascio PF, Land ML, Larimer FW, Hauser LJ. 2010. Prodigal: prokaryotic gene recognition and translation initiation site identification. BMC Bioinformatics 11:119. https://doi.org/10.1186/1471 -2105-11-119.

137. Fu L, Niu B, Zhu Z, Wu S, Li W. 2012. CD-HIT: accelerated for clustering the next-generation sequencing data. Bioinformatics 28:3150–3152. https://doi.org/10.1093/bioinformatics/bts565.

138. Buchfink B, Xie C, Huson DH. 2015. Fast and sensitive protein alignment using DIAMOND. Nat Methods 12:59–60. https://doi.org/10.1038/nmeth .3176.

139. Huson DH, Auch AF, Qi J, Schuster SC. 2007. MEGAN analysis of meta-genomic data. Genome Res 17:377–386. https://doi.org/10.1101/gr .5969107.

140. Li R, Yu C, Li Y, Lam TW, Yiu SM, Kristiansen K, Wang J. 2009. SOAP2: an improved ultrafast tool for short read alignment. Bioinformatics 25: 1966–1967. https://doi.org/10.1093/bioinformatics/btp336.

141. White JR, Nagarajan N, Pop M. 2009. statistical methods for detecting differentially abundant features in clinical metagenomic samples. PLoS Comput Biol 5:e1000352. https://doi.org/10.1371/journal.pcbi .1000352.