1 **Diversity and structure of the deep-sea sponge microbiome in the equatorial**

2 **Atlantic Ocean**

3 Sam E. Williams*[1, 2], Gilda Varliero[3], Miguel Lurgi[4], Jem Stach[5], Paul R. Race[5] and Paul

4 Curnow*[1]

5 [1]School of Biochemistry, University of Bristol, University Walk, Bristol BS8 1TD, UK

6 [2]Novo Nordisk Foundation Center for Biosustainability, Technical University of Denmark, Søltofts Plads, Building 220, 2800

7 Kgs. Lyngby, Denmark.

8 [3]Rhizosphere Processes Group, Swiss Federal Institute for Forest, Snow and Landscape Research (WSL), Birmensdorf,

9 Switzerland

10 [4]Department of Biosciences, Swansea University, Singleton Park, Swansea, SA2 8PP, UK.

11 [5]School of Natural and Environmental Sciences, Newcastle University, Newcastle upon Tyne NE1 7RU, UK

12 *Corresponding authors. Email: saedwi@biosustain.dtu.dk, p.curnow@bristol.ac.uk

13

14 **Abstract**

15 Sponges (phylum Porifera) harbour specific microbial communities that drive the ecology

16 and evolution of the host. Understanding the structure and dynamics of these communities is

17 emerging as a primary focus in marine microbial ecology research. Much of the work to date

18 has focused on sponges from warm and shallow coastal waters, while sponges from the

19 deep ocean remain less well-studied. Here, we present a metataxonomic analysis of the

20 microbial consortia associated with 23 deep-sea sponges. We identify a high abundance of

21 archaea relative to bacteria across these communities, with certain sponge microbiomes

22 comprising more than 90% archaea. Specifically, the archaeal family *Nitrosopumilaceae* are

23 prolific, comprising over 99% of all archaeal reads. Our analysis revealed sponge microbial

24 communities mirror the host sponge phylogeny, indicating a key role for host taxonomy in

25 defining microbiome composition. Our work confirms the contribution of both evolutionary

26 and environmental processes to the composition of microbial communities in deep-sea

27 sponges.

28

29 **Keywords**: Sponge, Porifera, Deep-sea, Microbiome, Metataxonomics, 16S rRNA gene

30 analysis

31
32

33

34

Importance

The deep ocean is the largest biome on Earth, accounting for >90% of the planet's marine environment. Despite this it remains a largely unexplored ecosystem, with less than 0.01% of the deep seafloor having been quantitatively sampled. Deep-sea sponges are ancient metazoans which harbour complex microbial communities and much still remains to be learned about the composition and diversity of these unique microbiomes. In an effort to address this, here we report a metataxonomic analysis of the microbial consortia associated with 23 deep-sea sponges from the equatorial Atlantic Ocean. Our findings reveal intricate, species-specific microbial communities dominated by ammonia-oxidizing archaea. This study highlights the significant role sponges play in shaping microbial consortia, providing new insights into deep-sea ecosystem dynamics. Importantly, our findings provide a scientific basis for understanding the evolutionary relationships between sponges and their symbiotic microorganisms.

Introduction

Sea sponges (phylum Porifera) live in intimate symbiotic relationships with complex microbial communities (1-3) Microbial abundance within sponge tissue can be orders of magnitude greater than the surrounding seawater (4). Sponge microbiotas are generally dominated by the Pseudomonadota and a few other core phyla (5), but at least 30 other variable phyla are commonly identified in studies of metataxonomic communities. These communities exhibit unusual features such as the presence of the sponge-associated phyla known as Poribacteria, which are only rarely observed living outside the sponge hosts (6, 7). Large-scale comparative surveys have confirmed that microbial communities are relatively stable within, but vary greatly in richness and diversity across, sponge species (6-9). This species dependence suggests that particular microorganisms might be selected for by sponge-microbe or microbe-microbe interactions. However, these relationships might be modulated by other biotic and abiotic factors; for example, water temperature appears to have a major influence on microbial community structure inside hosts (7, 10).

Much of our current knowledge of sponge-associated microbes is derived from samples collected in relatively shallow coastal waters (5). There are fewer studies investigating sponges which inhabit the deep ocean (11), at least in part due to the difficulty in obtaining such samples. Sponge communities from mesopelagic and abyssal depths might be expected to differ from their shallow-water relatives given the drastically different ambient conditions - cold, dark, nutrient-poor - and the consequent changes to the microbial life in the seawater around the sponge. Current evidence suggests that deep-water sponge-associated bacterial communities remain species-specific, and that sponges with both High- and Low Microbial Abundance (HMA and LMA, respectively) are found at these depths (8, 12, 13). Perhaps the major differences in the community structure of deep-sea sponges are the general loss of Cyanobacteriota and a lower occurrence of Poribacteria, as well as the increased abundance of Archaea (12-14). Steinert and colleagues also showed that Hexactinellid glass sponges, which commonly inhabit deep and cold waters, had lower overall microbiome diversity than deep-water demosponges and that there were differences between the bacteria associated with the two classes (12). Archaea are found to be more abundant than bacteria in some deep-water sponges and are almost entirely dominated by the phylum Thermoproteota, specifically family *Nitrosopumilaceae* (14, 15). Only two other archaeal phyla appear to be present (Methanobacteriota, Nanoarchaeota), all at low abundance and without an obvious relationship to sponge species or class (12-14). As with shallow-water sponges, the precise composition of deep-water sponge communities reflects the complex relationship between evolutionary and environmental factors (8, 16).

85   The recent shift towards whole genome-based taxonomy has significantly revised the

86   microbial tree of life (17). Traditional taxonomic assignment via outdated 16S rRNA

87   databases fail to account for this revision and perform poorly in standardised tests (18). The

88   newly developed, Greengenes2, unifies whole genome taxonomy and 16S rRNA, offers

89   improved accurate taxonomic assignment and has not yet been applied to the sponge

90   microbiome (19). As well as their fundamental role in sponge ecology, the microbial

91   communities of deep-sea sponges are also of increasing interest as a resource for novel

92   therapeutics. Sponge symbionts are well-established as a fertile repository of bioactive

93   natural products (20) and the diverse composition of sponge microbiota has the potential for

94   high and novel functional diversity (21). The microbial consortia of deep-sea sponges remain

95   largely untapped in such biodiscovery programmes, where metataxonomic data collected for

96   community analysis could guide future drug discovery programmes (22, 23).

97

98   In this study, we investigate the microbiome of 23 diverse sponges from 5 underexplored

99   deep-water seamounts in the Atlantic Ocean. We focus on elucidating the impact of host-

100   sponge class on microbial composition and use updated taxonomic assignments to better

101   profile these communities. We investigate factors shaping these microbial communities,

102   including the biogeographical factors, and sponge host phylogeny on microbiome

103   composition. Specifically, we explore the extent to which sponge host phylogeny correlates

104   with microbiome structure. This study adds novel insight into host phylogeny's impact on the

105   sponge microbiome and explores previously unstudied sponge taxa.

106

107

108

109

110

Materials and Methods

Sample Selection

A subset of 23 deep-sea sponges were chosen from a larger collection which was described previously (24). Sponges were chosen to represent a variety of depths and sampling sites. All the chosen samples were collected between depths of 569-2618 m (Average depth: 1455 m) from five different sampling sites in the Atlantic Ocean by the NERC research vessel RRS James Cook during research cruise JC094 (25) (Figure 1). Samples were provided to us directly by the cruise Chief Scientist Laura Robinson, University of Bristol. Sample handling and preparation was performed according to the protocols adopted for the Earth Microbiome Project (EMP) (26) and subsequently the Sponge Microbiome Project (5). Local sea water controls were not available for sequencing and so are absent from the dataset.
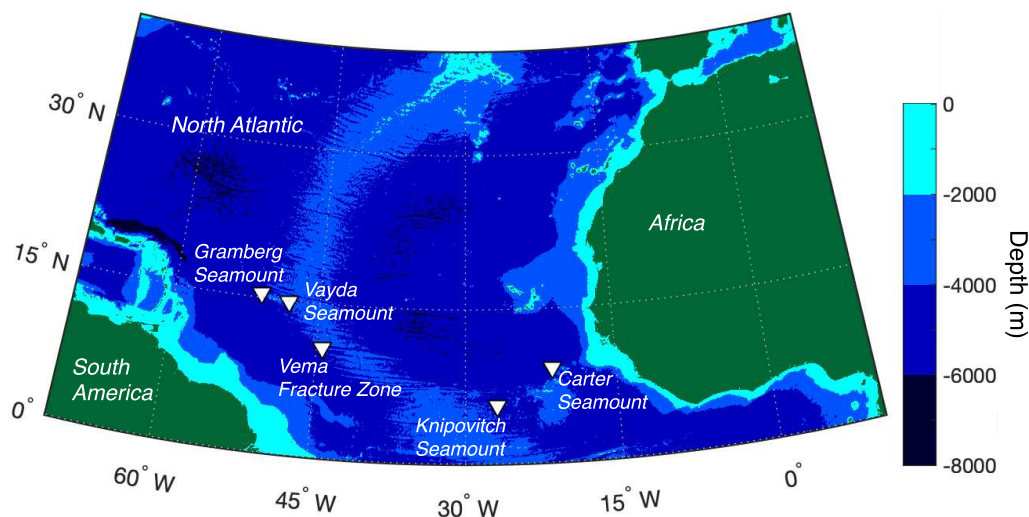

Figure 1. Map of the five deep-sea sampling areas (triangles) from the JC094 research cruise in the mid-Atlantic Ocean. Graphic created using ETOPO1 bathymetry.

DNA extraction

DNA was extracted from 0.25 g of sponge tissue using the DNeasy PowerSoil Kit (Qiagen, Hilden, Germany) using the optimized procedure of Marotz et al. (27) and including a blank extraction to account for intrinsic kit or reagent contamination (28). Prior to extraction, sponge samples were washed three times in sterile artificial sea water (ASW; Crystal Sea Marine Mix, Marine Enterprise International). All extractions were performed in a laminar flow hood.

Identification of sponge species

Sponge taxonomy was assigned based upon the mitochondrial cytochrome oxidase subunit I (COI) or the 28S rRNA gene. The COI gene was amplified through PCR using the universal primers LCO1490 and HCO2198 (Supplementary Table 1) (29). The reaction comprised 20 µL Platinum™ Hot Start PCR Master Mix (2X) (Invitrogen, Thermo Fisher Scientific, Waltham, MA, USA), 1.6 µL of each primer at 10 pmol/µl, 14.8 µL deionised water and 2 µL DNA template. Thermal cycling was performed as described by Yang et al. (30): 1 minute denaturation at 94 °C; 5 cycles of 94 °C for 30 sec, annealing at 45 °C for 90 sec and extension at 72 °C for 1 min; 35 cycles of 94 °C for 30 sec, 51 °C for 40 sec and 72 °C for 1 min; and a final extension step at 72 °C for 5 min. 5 µL of the amplified PCR fragment was analysed on a 1% TAE agarose gel to confirm the expected COI gene fragment size of approximately 680bp. Samples unable to be identified by the COI locus were reanalysed with 28S primers NL4F and NL4R (Supplementary Table 1) (31), using the same PCR reagents but with thermal cycling as follows: 10-min initial denaturation at 95 °C; 35 cycles of: 95 °C, 56 °C and 72 °C for 1 minute each; and a final extension step of 72 °C for 7 minutes (30). 5 µL of these reactions were analysed on a 1% TAE agarose gel to confirm 28S gene fragment size 858 to 1010 bp. PCR products were then either purified directly using Monarch PCR and DNA Cleanup Kit (New England Biolabs, Hitchin, UK), or if more than one band was present extracted from 1% TAE gel using the QIAquick Gel Extraction Kit (Qiagen, Hilden, Germany). Nucleotide BLAST (BLASTn) was used for taxonomic assignment (32). Sponge taxonomy was assigned at the species level if BLASTn % identity >97%, we recognise this threshold is arbitrary but sufficient for the purposes of this study (33). For identities below this, assignment was made to the nearest genus, recognizing that lower % identities yield approximate, not precise, taxonomic classifications. To distinguish between sponges of the same species, or unassigned sponges the original sample identifier (BXXXX) was added as a suffix.

Sponge Phylogenetic Analysis

COI gene sequences were aligned with MAFFT (v7.490) using the L-INS-I algorithm (34). A phylogenetic tree was then constructed using the IQ-TREE2 (v2.2.2.7) with model finder, the best-fitting model was TIM2+F+I+G4 and the analysis included 1,000 ultra-fast bootstrap replicates (35). The trimmed COI gene from the slime sponge *Oscarella carmela* (GenBank: NC_009090.1) was used as an outgroup on which the tree was rooted. The tree was visualised and annotated with iTOL v6 (36).

Sponge microbiome sequencing

170 The V4 and V5 regions of the microbial 16S rRNA gene were amplified via PCR with the

171 universal primers 515F-926R (26) combined with Illumina sequencing adaptors

172 (Supplementary Table 1). Thermal cycling followed the standardised protocol of the Earth

173 Microbiome Project (26). This entailed an initial denaturation at 94 °C for 3 minutes, followed

174 by 35 cycles of 94 °C for 45 seconds, 50 °C for 60 seconds and 72 °C for 90 seconds, with a

175 final extension of 10 minutes at 72 °C. The presence of the expected PCR products was

176 confirmed by visualising 5 µL of the reaction on a 1% agarose gel. For most sponge samples

177 this produced a single intense band at approximately 478 bp, and the amplicon pool was

178 cleaned with the QIAquick PCR Purification Kit (Qiagen 28104). If spurious bands were

179 identified the 478 bp fragment was extracted using the QIAquick Gel Extraction Kit (Qiagen

180 28704) alongside a blank gel lane extraction as a further negative control for environmental

181 contamination.

182

183 Illumina sequencing was performed by the genomics facility at the University of Bristol. After

184 initial quality assessment and normalisation of amplicon libraries, samples were indexed with

185 the XT Index Kit (Nextera®) and sequenced using the MiSeq platform (Illumina). Base

186 calling and quality assessment was performed with Real Time Analysis version 1.18.54.0. At

187 least 200,000 reads were recovered for each sample.

188

189 *Sponge microbiome sequencing and bioinformatic analyses*

190 Microbiome analysis was performed with QIIME2 2023.7 (37) and the data processing

191 pipeline is available at https://github.com/Sam-Will/Deep_Sponge_Micro. Adaptors and

192 primers were removed from the demultiplexed sequence data using the cutadapt plugin (38)

193 and the subsequent data was denoised with DADA2 (v. 1.26) (39) via the denoised-pair

194 command. Truncation parameters were chosen based on the quality plots to ensure a

195 minimum of 30 bp overlap between reads. Amplicon sequence variants (ASVs) were

196 assigned taxonomy using classify-sklearn feature-classifier (40) trained against the

197 Greengenes2 2022.10 full length sequences classifier (19). Data was imported to R version

198 4.0.2 for further analysis, and the following packages were used (41). Inherent laboratory

199 contamination during sample handling and eDNA extraction was assessed by analysing

200 blank samples from the DNA extraction kit (350,173 sequences) and from the gel extraction

201 kit (398 sequences). Twenty-six of the ASVs in these negative controls were identified as

202 likely true contaminants by the decontam 1.8.0 package in combined mode (42). These

203 sequences were removed from the sponge dataset (Supplementary Table 2). Identified

204 contaminants were removed from all samples in the data set prior to further analysis.

205 Phylogenies were constructed with phyloseq 1.42.0 and ASV alpha diversity values were

206 calculated using the 'estimate_richness()' function and assessed with aov() (41). Rarefaction

207   curves were calculated with iNEXT (v 2.0.20) (43). Permutational multivariate analyses of

208   variance (PERMANOVA), non-metric multidimensional scaling (NMDS) and the distance

209   decay analyses were calculated using Vegan (v 2.5.7) (44). Before performing

210   PERMANOVA, the dataset was tested for homoscedasticity using the function betadisper()

211   and then anova(). As homoscedasticity was observed (p > 0.05 in ANOVA test),

212   PERMANOVA was then performed on the Bray-Curtis dissimilarity matrix calculated on the

213   Hellinger transformed ASV dataset (1000 permutations). PERMANOVA was performed

214   using all samples except from sample Hexactinellid B01175 which was the only

215   representative of >2500 m depth. To calculate the community correlation with sampling

216   depth (i.e., distance decay) and geographical distances, mantel statistical tests ("method =

217   Spearman") were performed on the Bray-Curtis matrix computed on Hellinger transformed

218   ASV dataset and Euclidean matrix calculated on depth data, and geographical distances

219   (1000 permutations). *p*-values obtained from the Mantel tests were corrected using the

220   Holm-Bonferroni method (45). Geographical distances were calculated using the R library

221   geosphere (v 1.5-1.8) (46). NMDS was performed on the Hellinger transformed ASV

222   dataset. Plots and data manipulation were performed using gplots (v 3.1.1), RColorBrewer (v

223   1.1.2), tidyr (v 1.1.2), ggplot2 (v 3.3.3) and svglite (v 2.1.2). A dendrogram of the sponge

224   microbiome community structure was obtained using the hclust function ("method =

225   ward.D2") and was formatted using ape (v 5.7-1) and TreeTools (v 1.10.0) packages (47,

226   48). All of the above steps were done in the R environment (49). Dendroscope (v 3.8.10)

227   (50) was to create a tanglegram depicting similarities between the microbiome community

228   structure and sponge phylogeny.

229

230   *Data Availability*

231   The COI sequences of sponges in this study have been deposited in GenBank assigned the

232   accession numbers PP092504-PP092519 and OP036683.1. All raw 16S rRNA gene

233   amplicon sequencing data is available in the NCBI Sequence Read Archive, associated with

234   BioProject number PRJNA702029. Additionally, the data processing pipeline, including

235   Qiime2 taxonomy, representative sequences, phylogenetic trees, and metadata, can be

236   found at https://github.com/Sam-Will/Deep_Sponge_Micro.

237

Results

Classification of sponge samples

Sponge taxonomy was initially based upon the gene for mitochondrial cytochrome oxidase (29). This approach was used to successfully classify 17 of the 23 sponges considered here at the genus or species level (Figure 2). Assignments for certain samples to the nearest genus, including *Euplectella, Desmacella*, and *Aspidoscopulia,* are tentative since they have a low BLAST identity to their nearest relative and thus are likely novel. Of the remaining 6 sponges, 2 were subsequently identified to species level by the 28S rRNA gene (*Hertwigia* sp. MD-2008 97.09% and *Trachycladus* sp. NCl325 97.51%). The four remaining sponges were identified only to the class level as Hexactinellid, based on a visual inspection of spicules. Overall, the 23 deep-sea sponges studied here comprise 13 demosponges with representatives from 6 orders, 8 families and 10 species; and 10 hexactinellids from 2 orders, 3 families and 5 species (Figure 2, Supplementary Table 3). To our knowledge this dataset comprises several sponge genera whose microbiota has not previously been characterised, including *Rhabderemia, Trachycladus, Euplectella* and *Aspidoscopulia* (11, 12, 51).

9

Tree scale: 0.1

Homoscleromorpha - Oscarella carmela

Euplectella sp B1101 [96.11%]

Aspidoscopulia sp. B1963 [86.91%]

Farrea sp USNM 1122183 [99.11%]

Aphrocallistes beatrix B01657 [99.69%]

Aphrocallistes beatrix B0790 [99.54%]

Polymastia corticata [99.85%]

Phorbas bihamiger [97.50%]

Desmacella hyalina [98.37%]

Desmacella sp. B01623 [92.15%]

Desmacella sp. B1951 [92.86%]

Rhabderemia sp. 1 RAD-2021 [98.47%]

Stelletta tuberosa B01165 [98.76%]

Stelletta tuberosa B1955 [100%]

Geodia megastrella [100%]

Geodia barretti [99.69%]

Vulcanella aberrans B01671 [99.36%]

Vulcanella aberrans B01116 [99.38%]
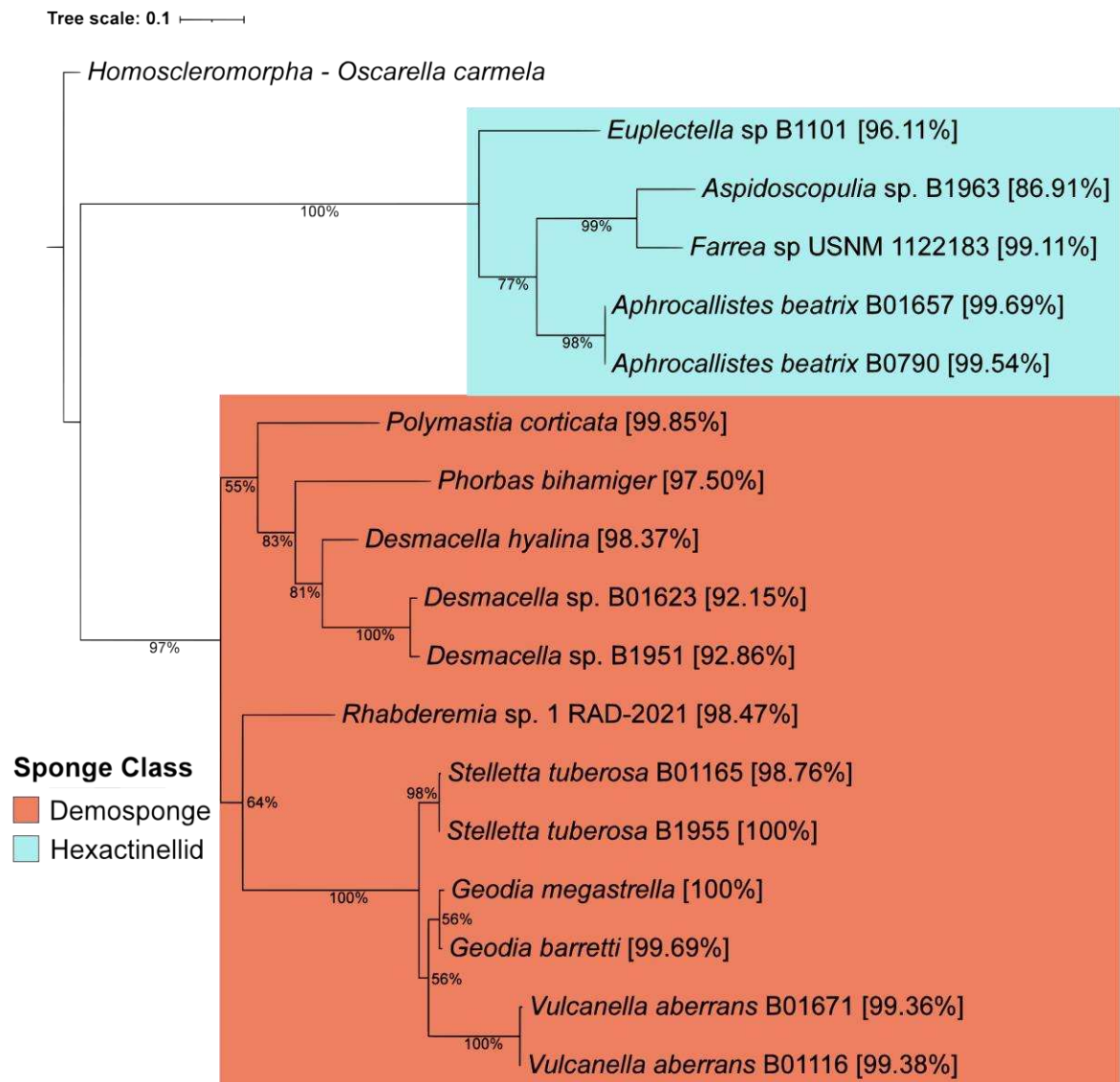
Sponge Class
- Demosponge
- Hexactinellid

Figure 2. Maximum likelihood tree for 17 deep-sea sponge samples based on the mitochondrial cytochrome oxidase (COI) gene barcode. Bootstrap values are shown as a percentage of 1000 replications. Taxon branches are labelled with the closest relative from BLASTn (% ID) and the sampling site. Demospongia are highlighted in orange and Hexactinellida in blue. The slime sponge *Oscarella carmela* COI sequence is used as an outgroup.

Microbial amplicon sequence variants (ASVs) and richness

Sequence abundance from the sponge microbiota for each sponge sample ranged from 200,223 to 437,147 with an average of 292,684 features per sponge. While individual sponges varied, across the whole dataset there was higher amount of archaeal (3,514,471) than bacterial abundance (3,185,734) and for *Hertwigia sp.* MD-2008 and Hexactinellid B01175 archaea represented over 90% of the reads (Supplementary Figure 1). Sequencing depth was sufficiently deep to be confident in recovering an accurate estimate of diversity, with all samples reaching at least 200,000 reads after DADA2 denoising (Supplementary

10

269   Table 4) and approaching saturation in rarefaction curves (Figure 3a, Supplementary Figure

270   2).

271

272   The total number of ASVs (*i.e.* richness) present across the 23 sponges was 11,465. The

273   microbiota of sponge *Desmacella sp.* B1951 had the greatest number of ASVs with 1864,

274   and this was over 20 times the number of ASVs in the least diverse sponge (*Aphrocallistes*

275   *beatrix* B01657, 80 ASVs; Figure 3a, Supplementary Table 5). There was considerably

276   higher bacterial richness at the Kingdom level with 10,844 ASVs assigned to Bacteria, 595

277   ASVs assigned to Archaea, and 26 ASVs that could not be assigned. Hexactinellid samples

278   (n = 10) had 4698 bacterial and 315 archaeal ASVs, while Demosponge samples (n = 13)

279   collectively had 6753 bacterial and 378 archaeal ASVs. Number of observed ASVs was

280   slightly higher in demosponges compared to hexactinellids, but this difference was not

281   statistically significant (p value = 0.298). Demosponges did, however, exhibit a significantly

282   higher Shannon diversity index (p value = 0.0124) (Figure 3b). The range of ASVs within the

283   dataset was large however, and certain hexactinellid samples from the *Euplectellidae* family

284   (*Euplectella* sp. B1101 and *Hertwigia* sp. MD-2008) had amongst the highest number of

285   ASVs (Supplementary Table 5). Examining the impact of depth on richness and diversity, no

286   correlation was observed between increasing depth on observed ASVs (p value = 0.189414)

287   or Shannon diversity (p value = 0.7624). However, a relationship emerged between depth

288   and archaeal ASVs in demosponges, with a decrease in the number of observed archaea as

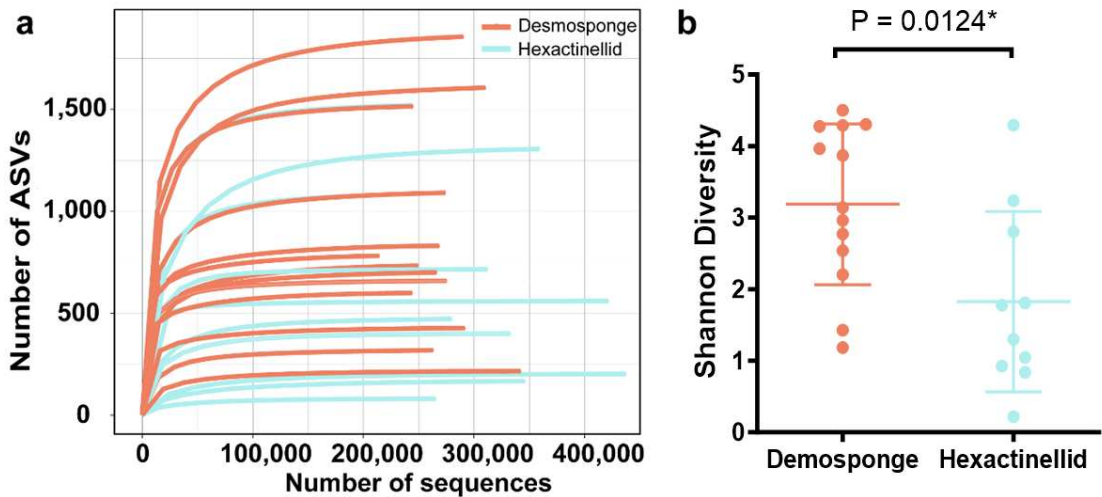289   depth increased, this pattern was exclusive to demosponges (p value = 0.002209).

290

291

292   Figure 3. Microbial amplicon sequence variants (ASVs) and Shannon diversity derived from 23 deep-

293   sea sponges. (**a**) Rarefaction curves of 16S rRNA gene diversity for deep-sea Atlantic sponge

294   samples displayed by sponge class. Results from the hexactinellid *Euplectella* sp B1101 (1534 ASVs)

295   and the demosponge *Vulcanella aberrans* B0116 (1530 ASVs) superimpose and are difficult to

11

296   distinguish on this plot. (**b**) Shannon diversity associated with each sponge, presented by sponge

297   class.

298

299   *Taxonomic distribution and relative abundance of the sponge microbiota*

300   We identified a total of 41 bacterial phyla and 4 archaeal phyla across the 23 sponge

301   species analysed (Figure 4). There were 8 phyla found across all sponges in the dataset:

302   Thermoproteota, Pseudomonadota, Chloroflexota, Desulfobacterota, Bacteroidota,

303   Actinobacteriota, Gemmatimonadota and Myxococcota. Bdellovibrionota was present in low

304   abundance in all samples except *Desmacella hyalina*. Acidobacteria, a dominant phylum in

305   certain sponges, was present in all sponges except for *Aphrocallistes beatrix* B01657. There

306   was a significant amount of reads unclassified even at the phylum level, this was particularly

307   prominent in the two novel *Desmacella* sponge species, which had 56.11% (sp. B01623)

308   and 32.16% (sp. B1951) of the relative abundance from unclassified reads. The diversity of

309   phyla hosted varied across species, with *Aphrocallistes beatrix* B01657 hosting as few as 13

310   phyla, while *Vulcanella aberrans* B01671 exhibited a rich diversity of 38 phyla.

311

312   Archaea were more abundant than bacteria, with most archaeal abundance represented by

313   phylum Thermoproteota (99.78%). This came almost completely from the *Nitrosopumilaceae*

314   family (99.73%) which were by far the most abundant single taxa in our samples, where the

315   second and third most abundant were family *CG1-02-33-12* (Nanoarchaeota phylum) and

316   *Thalassarchaeaceae* (Thermoplasmatota phylum) with 0.11% and 0.04%. The most

317   abundant bacteria were Pseudomonadota, particularly the classes Alpha- and

318   Gammaproteobacteria, although Alphaproteobacteria organisms were lacking from both

319   samples of *Aphrocallistes beatrix* obtained from the Vayda sea mount. The sponge

320   associated Poribacteria phylum, thought to be primarily associated with sponges, were

321   notably scarce among our deep-sea samples, but was identified in 5 demosponge samples

322   and 2 hexactinellids. When present, this phylum comprised a tiny fraction of the overall

323   diversity, falling beneath the >0.5% relative abundance threshold for plotting on Figure 4.
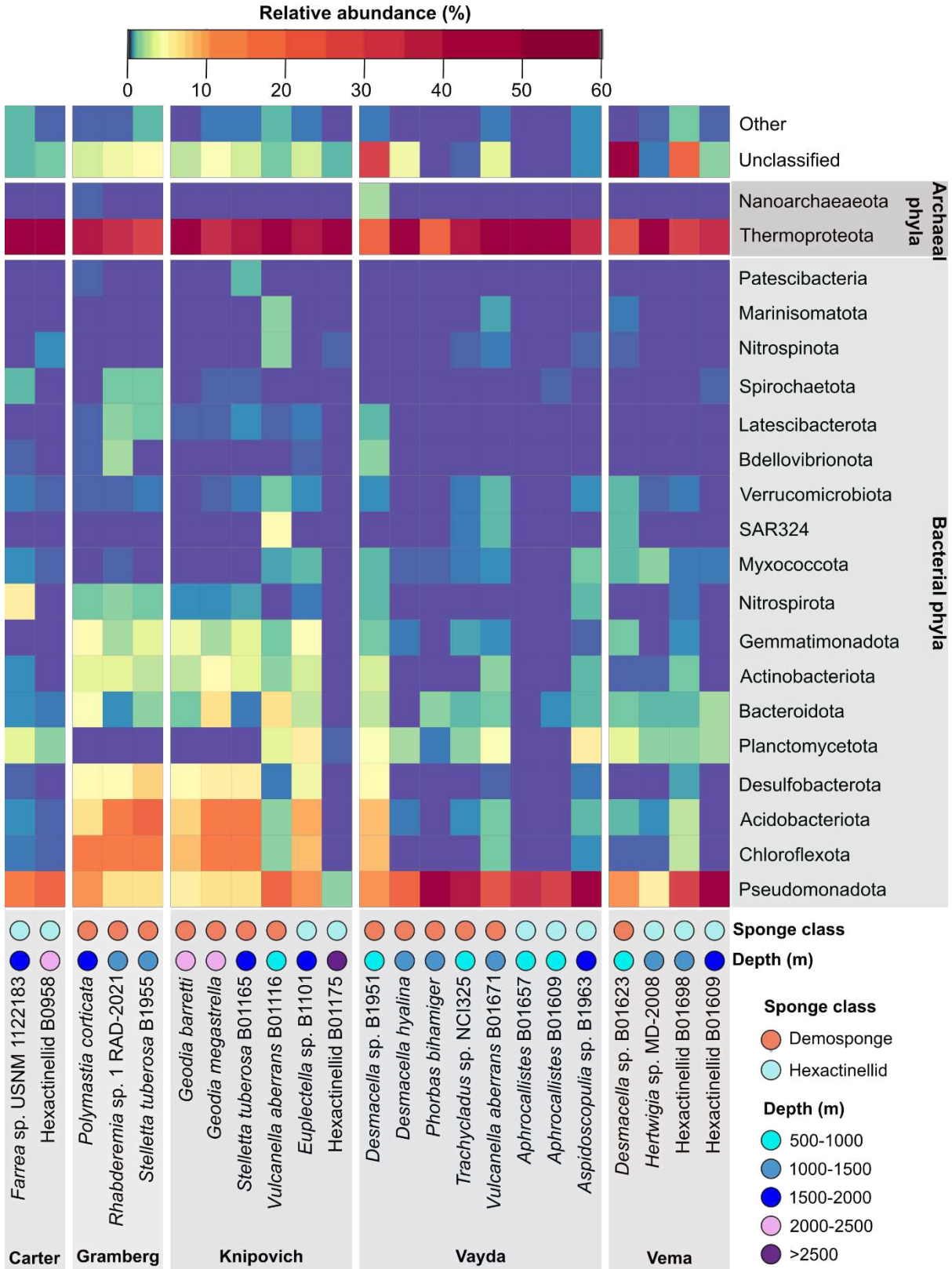
Figure 4. Abundant phyla in the deep-sea sponge microbiome, grouped by sampling location. Only phyla present with a relative abundance > 0.5% of total taxa in at least one sample are reported.

Sponge species and geographic influence on the deep-sponge microbiome

329 To determine the influence of biotic and abiotic factors on the sponge microbiome, we

330 performed Hellinger transformation on the ASV dataset and conducted a PERMANOVA

331 analysis against three factors: sponge taxonomic class, sampling site, and sampling depth

332 category. The results suggest that sponge class only explained 9% of variation in the

333 microbiome, whereas sampling site and the depth category were responsible for 32% and

334 16% of the total variation, respectively (Table 1). Mantel tests showed significant correlation

335 (p-value = 0.013) between the ASV dataset and sponge depth (r = 0.223) but not with

336 geographic distances (r = 0.031, p-value = 0.302). These results show that while sampling

337 site and depth impact microbiome and illustrate the important influence of locality on

338 community structure (52).

339

340 Table 1. PERMANOVA (A) and Mantel (B) tests performed on Hellinger-transformed ASV dataset,

341 using Bray-Curtis dissimilarity matrix with 1000 permutations. *Depth categories as in Figure 4.

342 **Adjusted p-values using the Holm-Bonferroni method.

343

	Factor	R^2	p
A PERMANOVA	Sampling Site	0.319	0.001
	Depth Category*	0.164	0.003
	Sponge Class	0.085	0.002
	Sampling Site: Depth Category*	0.211	0.009

	Factor	r	p**
B Mantel test	Geographical distance	0.031	0.302
	Depth	0.223	0.013

344

345 Next, we conducted a Non-Metric Multidimensional Scaling (NMDS) analysis of the

346 community composition against the same factors of sponge class, site and depth (our

347 dataset was insufficient for within-species comparisons). The NMDS groupings by sponge

348 class were overlapping to some degree, although could be distinguished (Figure S5a).

349 Sampling site and depth did not seem to be associated with clearly distinct clusters (Figure

350 S5b-c). The most obvious individual cluster for sampling site was formed by the three

351 demosponge samples from Gramberg Seamount (Figure S5b). Collectively, while

352 PERMANOVA reveals that site and depth are generally important factors in explaining

353 variation in microbial communities, our limited dataset does not show clear and obvious

354   distinctions in community composition when comparing at the finer-grained level of sites or

355   specific depths.

356

357   Finally, previous studies have suggested that sponge-associated microbial communities

358   exhibit a high degree of host specificity at lower taxonomic ranks (11). We sought to

359   investigate this phenomenon at the species level. To accomplish this, we performed a co-

360   phylogenetic analysis relating the phylogenetic structure of sponge microbial communities

361   compared to the COI gene-based phylogenetic tree from the host sponges. This

362   comparative analysis showed similar, but not identical clading between the microbial

363   assemblage and the host phylogeny (Figure 5). Our findings highlight the predominance of

364   host species in shaping the structure of the deep-sea sponge microbiome, mostly

365   irrespective of variations in depth or geographical distribution of the same sponge species.

366   One exception here is the two samples of *Stelletta tuberosa*, where it's clear variation in

367   depth and site play a larger role. Additionally, an anomaly was noted in the sponge

368   *Euplectella* sp. B1101, which demonstrated a phylogenetic affinity more closely aligned with

369   *Geodia* sponges than with other members of the Hexactinellida.
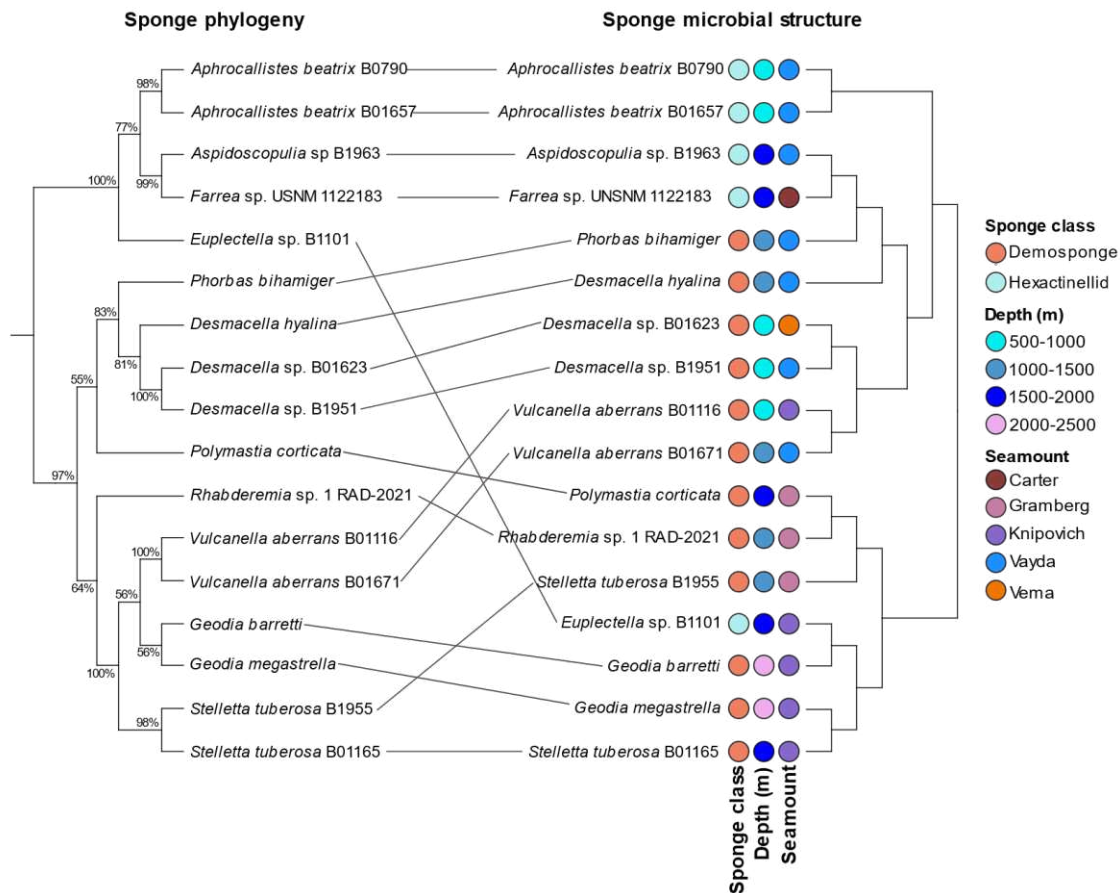

Figure 5. The association between sponge host phylogeny based on COI gene sequences and associated microbial community structure. The left tree represents the COI gene-based phylogenetic

15

relationships among sampled sponge species, while the right tree illustrates the phylogenetic structure of their respective microbial communities. Lines connecting the two trees highlight the observed congruence, demonstrating the influence of host phylogeny on microbial assemblage.

370

371

372

16

Discussion

373
374 The microbial ecology of the deep-sea remains underexplored. This study employs

375 metataxonomic methods to investigate the microbial community structure of 23 sponges

376 from 5 Atlantic seamounts, including many sponges that have not previously been

377 characterised and are likely to be unique to the deep ocean. We reveal a microbiome

378 predominantly consisting of ammonia-oxidizing archaea from the family *Nitrosopumilaceae*.

379 Importantly, our data provides further evidence that host-sponge phylogeny is associated

380 with microbiome composition, contributing to a further understanding of deep-sea sponge

381 microbial ecology.

382

383 **Thermoproteota dominate the sponge microbiome**

384 Our findings further contribute to the understanding of deep-sea sponge microbiota,

385 corroborating prior observations of high abundance of archaea and particularly dominance of

386 Thermoproteota (6, 12-14). Metagenome-assembled genomes of sponge associated

387 *Nitrosopumilaceae* show a high degree of symbiotic adaption (53). In our study an

388 overwhelming 99.75% of archaeal reads were identified as Thermoproteota, underscoring

389 their key role to oxidise ammonia, and so contribute to nitrogen fixation (54, 55).

390

391 **Methodological Limitations**

392 The absence of seawater controls limits the extent to which this study can be compared to

393 others. Though we implemented stringent rinsing protocols, seawater contamination remains

394 a concern as do reads present in the blank extraction. The sponges studied here are a

395 subset of a larger collection collected to study biomineralization (56) and subsequently

396 bioprospecting (24, 57, 58). The samples thus represent an unbiased general survey, with a

397 modest sample size, rather than a systematic effort to understand a particular sponge

398 species or geographical location, for this analysis we direct readers to large systematic

399 studies of the deep-sea sponge microbiome (11, 12).

400

401 **Comparison with large scale studies**

402 Our findings align with Busch et al.'s extensive study of 1077 deep-sea sponges (11), but

403 add novel, previously unstudied genera of deep-sea sponges. The use of GreenGenes2

404 reference tree improves microbial taxonomic assignment and facilitating comparability with

405 the Genome Taxonomy Database (19, 59). Our study also confirms the presence of both

406 LMA and HMA sponges in deep-sea environments (60). Interestingly, while hexactinellids

407 generally appear to be LMA, we do identify a potential HMA Hexactinellid class sponge from

408 the genus *Euplectella.* a genus that has not been characterised previously and has a

409 microbiome similar to phylogenetically distance sponges. Busch et al. noted that

17

410 *Amphidiscella caledonica* (family *Euplectellidae*) diverged from typical LMA glass sponges,

411 showing high Chloroflexota abundance and several other members of *Euplectellidae* family

412 clustered with more with demosponges. We identify two further *Euplectellidae* genera

413 (*Euplectella, Hertwigia*) which deviate from the classical composition of LMA glass sponges.

414

415 **Genus-Specific Comparisons**

416 Our study also includes two representatives of the demosponge genus *Geodia* collected

417 from the Knipovich seamount. Both sponges have similar community profiles to samples of

418 *Geodia hentscheli* and *Geodia barretti* from other studies (6, 16, 61)*,* being dominated by

419 Pseudomonadota, Chloroflexota and Acidobacteriota, as well as Desulfobacterota and

420 Gemmatimonadota. Although these other studies did report a higher abundance of

421 Poribacteria that was not observed here. Otherwise, samples of *Stelletta tuberosa* from

422 Gramberg and Knipovich seamounts had relatively dissimilar profiles based on the

423 tanglegram analysis, whereas the two *Vulcanella aberrans* samples from Knipovich and

424 Vayda sites clustered together strongly. The relative dominance of Chloroflexota and

425 Acidobacteriota in *Stelletta tuberosa* was previously observed in another member of the

426 genus, *Stelletta normani* obtained from canyons of the North Atlantic and is consistent with

427 the assignment of this species as a deep-sea HMA sponge (13) .

428

429 **Influence of sponge host and biogeographical factors**

430 Deep-sea sponge microbial communities are influenced by the sponge host species and

431 biogeographical factors such as nutrient profile and the surrounding seawater (16). Our

432 study emphasises the role of depth in shaping sponge associated communities, a finding

433 which is supported by Busch et al. (11), the *Tara* Oceans project (62) but contrasting with

434 other systemic studies (12). A study of seamounts impact on the microbial composition of

435 sponges did highlight the impact of depth on phyla distribution (16), suggesting depth could

436 impact communities in an environment specific manner. The deep-sea sponge microbiome

437 project highlighted a weak but significant biogeographical impact on community composition,

438 albeit over a much larger distance of 10,000 km (11). In our study, significant differences

439 between sponges at various sites suggest that isolation by seamounts, rather than

440 geographical distance, influences community structure. Influence of sampling site over

441 geographical distance has been previously observed in deep-sea microbiomes both in

442 sponge hosts and deep-sea sediments (52, 63).

443

444 We also identified accordance between sponge host phylogeny and sponge microbiome

445 structure (Figure 5), indicating host-driven microbiome composition. For example, the

446 microbiomes of two *Aphrocallistes beatrix* sponges*,* both collected from the Vayda

18

447 seamount, displayed a consistent microbial community, also supporting previous findings

448 from the same species (12). The identification of a co-phylogenetic relationship, known as

449 phylosymbiosis, has been previously identified in coral (64), animal (65) and shallow reef

450 sponge microbiomes (66). While our findings indicate an influence of host phylogeny on

451 microbiome structure, this relationship is not absolute, reflecting that deep-sea sponges may

452 prioritize functional redundancy over species-specificity in their microbial associations. This

453 is supported by a comparative study of 39 HMA sponge species, revealing a more

454 pronounced species-specificity in microbial composition among shallow water sponges than

455 in their deep-sea counterparts (67).

456

457 **Conclusion**

458 Overall, our study explores the microbiomes of previously unstudied deep-sea sponges and

459 identify depth and sponge phylogeny to be drivers of microbial composition. Our results

460 particularly highlight distinct microbiomes in the *Euplectellidae* family, differing from other

461 glass sponges which warrant further investigation.

462

463 **Author Contributions:**

464 S.E.W. set out the methodology of the study, performed data collection, conducted
465 bioinformatic data curation and initial investigations of the data. S.E.W. and G.V. conducted
466 data analysis and produced visualizations for the manuscript. S.E.W. wrote the original draft
467 of the manuscript with help from G.V., P.C., M.L., J.E.M and P.R.R. All authors reviewed and
468 edited the manuscript. P.C. and P.R.R. conceived the project, provided supervision and
469 acquired funding. All authors have read and agreed to the published version of the
470 manuscript.

479

480

481

482

483

484

485

486

References:

1. Taylor MW, Radax R, Steger D, Wagner M. Sponge-associated microorganisms: evolution, ecology, and biotechnological potential. Microbiol Mol Biol Rev. 2007;71(2):295-347.

2. Kiran GS, Sekar S, Ramasamy P, Thinesh T, Hassan S, Lipton AN, et al. Marine sponge microbial association: towards disclosing unique symbiotic interactions. Mar Environ Res. 2018;140:169-79.

3. Hentschel U, Piel J, Degnan Sm, Taylor MW. Genomic insights into the marine sponge microbiome Nat Rev Microbiol. 2012;10:641-54.

4. Hentschel U, Usher KM, Taylor MW. Marine sponges as microbial fermenters. FEMS Microbiology Ecology. 2006;55:167-77.

5. Moitinho-Silva L, Nielsen S, Amir A, Gonzalez A, Ackermann GL, Cerrano C, et al. The sponge microbiome project. Gigascience. 2017;6(10).

6. Thomas T, Moitinho-Silva L, Lurgi M, Björk JR, Easson C, Astudillo-García C, et al. Diversity, structure and convergent evolution of the global sponge microbiome. Nat Commun. 2016.

7. Schmitt S, Tsai P, Bell J, Fromont J, Ilan M, Lindquist N, et al. Assessing the complex sponge microbiota: core, variable and species-specific bacterial communities in marine sponges. Isme J. 2012;6(3):564-76.

8. Reveillaud J, Maignien L, Murat Elen A, Huber JA, Apprill A, Sogin ML, et al. Host-specificity among abundant and rare taxa in the sponge microbiome. The ISME Journal. 2014;8:1198-209.

9. Diez-Vives C, Taboada S, Leiva C, Busch K, Hentschel U, Riesgo A. On the way to specificity - microbiome reflects sponge genetic cluster primarily in highly structured populations. Molecular Ecology. 2020;29(22):4412-27.

10. Lurgi M, Thomas T, Wemheuer B, Webster NS, Montoya JM. Modularity and predicted functions of the global sponge microbiome network. Nat Commun. 2019;10:992.

11. Busch K, Slaby BM, Bach W, Boetius A, Clefsen I, Colaço A, et al. Biodiversity, environmental drivers, and sustainability of the global deep-sea sponge microbiome. Nat Commun. 2022;13(1):5160.

12. Steinert G, Busch K, Bayer K, Kodami S, Arbizu PM, Kelly M, et al. Compositional and Quantitative Insights Into Bacterial and Archaeal Communities of South Pacific Deep-Sea Sponges (Demospongiae and Hexactinellida). Frontiers in Microbiology. 2020;11(716).

13. Kennedy J, Flemer B, Jackson SA, Morrissey JP, O'Gara F, Dobson ADW. Evidence of a Putative Deep Sea Specific Microbiome in Marine Sponges. Plos One. 2014;9(3).

14. Jackson SA, Flemer B, McCann A, Kennedy J, Morrissey JP, O'Gara F, et al. Archaea Appear to Dominate the Microbiome of Inflatella pellicula Deep Sea Sponges. Plos One. 2013;8(12):8.

15. Zhang S, Song W, Wemheuer B, Reveillaud J, Webster N, Thomas T. Comparative Genomics Reveals Ecological and Evolutionary Insights into Sponge-Associated Thaumarchaeota. mSystems. 2019;4(4):e00288-19.

16. Busch K, Hanz U, Mienis F, Mueller B, Franke A, Martyn Roberts E, et al. On giant shoulders: how a seamount affects the microbial community composition of seawater and sponges. Biogeosciences. 2020;17:3471-86.

17. Parks DH, Chuvochina M, Waite DW, Rinke C, Skarshewski A, Chaumeil PA, et al. A standardized bacterial taxonomy based on genome phylogeny substantially revises the tree of life. Nat Biotechnol. 2018;36(10):996-+.

18. Odom AR, Faits T, Castro-Nallar E, Crandall KA, Johnson WE. Metagenomic profiling pipelines improve taxonomic classification for 16S amplicon sequencing data. Sci Rep. 2023;13(1):13957.

19. McDonald D, Jiang Y, Balaban M, Cantrell K, Zhu Q, Gonzalez A, et al. Greengenes2 unifies microbial data in a single reference tree. Nat Biotechnol. 2023.

540   20.   Taylor MW, Tsai P, Simister RL, Deines P, Botte E, Ericson G, et al. 'Sponge-
541   specific' bacteria are widespread (but rare) in diverse marine environments. The ISME
542   Journal. 2013;7(2):438-43.
543   21.   Rust M, Helfrich EJN, Freeman CJ, Nanudorn P, Field CM, Rückert C, et al. A
544   multiproducer microbiome generates chemical diversity in the marine sponge *Mycale*
545   *hentscheli*. PNAS. 2020;117(17):9508-18.
546   22.   Wemheuer F, Taylor JA, Daniel R, Johnston E, Meinicke P, Thomas T, et al.
547   Tax4Fun2: prediction of habitat-specific functional profiles and functional redundancy based
548   on 16S rRNA gene sequences. Environmental Microbiome. 2020;15(1):11.
549   23.   Loureiro C, Galani A, Gavriilidou A, Chaib de Mares M, van der Oost J, Medema MH,
550   et al. Comparative Metagenomic Analysis of Biosynthetic Diversity across Sponge
551   Microbiomes Highlights Metabolic Novelty, Conservation, and Diversification. mSystems.
552   2022;7(4):e0035722.
553   24.   Williams SE, Stennett HL, Back CR, Tiwari K, Ojeda Gomez J, Challand MR, et al.
554   The Bristol Sponge Microbiome Collection: A Unique Repository of Deep-Sea
555   Microorganisms and Associated Natural Products. Antibiotics. 2020;9(8):509.
556   25.   Robinson LF. RRS James Cook Cruise JC094, October 13–November 30 2013,
557   Tenerife-Trinidad. TROPICS, Tracing Oceanic Processes using Corals and Sediments.
558   Reconstructing abrupt Changes in Chemistry and Circulation of the Equatorial Atlantic
559   Ocean: Implications for global Climate and deep-water Habitats. Pangaea. 2014.
560   26.   Thompson LR, Sanders JG, McDonald D, Amir A, Ladau J, Locey KJ, et al. A
561   communal catalogue reveals Earth's multiscale microbial diversity. Nature. 2017.
562   27.   Marotz C, Amir A, Humphrey G, Gaffney J, Gogul G, Knight R. DNA extraction for
563   streamlined metagenomics of diverse environmental samples. BioTechniques.
564   2017;62(6):290-3.
565   28.   Salter SJ, Cox MJ, Turek EM, Calus ST, Cookson WO, Moffatt MF, et al. Reagent
566   and laboratory contamination can critically impact sequence-based microbiome analyses.
567   BMC Biology. 2014;12(1):87.
568   29.   Folmer O, Black M, Hoeh W, Lutz R, Vrijenhoek R. DNA primers for amplification of
569   mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates. Mol Mar
570   Biol Biotechnol. 1994;3(5):294-9.
571   30.   Yang Q, Franco CMM, Sorokin SJ, Zhang W. Development of a multilocus-based
572   approach for sponge (phylum Porifera) identification: refinement and limitations. Sci Rep-Uk.
573   2017;7(1):41422.
574   31.   Nichols SA. An evaluation of support for order-level monophyly and interrelationships
575   within the class Demospongiae using partial data from the large subunit rDNA and
576   cytochrome oxidase subunit I. Molecular Phylogenetics and Evolution. 2005;34(1):81-96.
577   32.   Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search
578   tool. J Mol Biol. 1990;215(3):403-10.
579   33.   Timmers MA, Vicente J, Webb M, Jury CP, Toonen RJ. Sponging up diversity:
580   Evaluating metabarcoding performance for a taxonomically challenging phylum within a
581   complex cryptobenthic community. Environmental DNA. 2022;4(1):239-53.
582   34.   Katoh K, Standley DM. MAFFT multiple sequence alignment software version 7:
583   improvements in performance and usability. Mol Biol Evol. 2013;30(4):772-80.
584   35.   Minh BQ, Schmidt HA, Chernomor O, Schrempf D, Woodhams MD, von Haeseler A,
585   et al. IQ-TREE 2: New Models and Efficient Methods for Phylogenetic Inference in the
586   Genomic Era. Molecular Biology and Evolution. 2020;37(5):1530-4.
587   36.   Letunic I, Bork P. Interactive Tree Of Life (iTOL) v5: an online tool for phylogenetic
588   tree display and annotation. Nucleic Acids Res. 2021;49(W1):W293-W6.
589   37.   Bolyen E, Rideout JR, Dillon MR, Bokulich NA, Abnet C, Al-Ghalith GA, et al. QIIME
590   2: Reproducible, interactive, scalable, and extensible microbiome data science. PeerJ
591   Preprints. 2018;6:e27295v2.
592   38.   Martin M. Cutadapt removes adapter sequences from high-throughput sequencing
593   reads. EMBnetjournal. 2011;17(1):10.

594 39.   Callahan BJ, McMurdie PJ, Rosen MJ, Han AW, Johnson AJA, Holmes SP. DADA2:
595 High-resolution sample inference from Illumina amplicon data. Nature Methods.
596 2016;13(7):581-3.
597 40.   Bokulich NA, Kaehler BD, Rideout JR, Dillon M, Bolyen E, Knight R, et al. Optimizing
598 taxonomic classification of marker-gene amplicon sequences with QIIME 2's q2-feature-
599 classifier plugin. Microbiome. 2018;6(1).
600 41.   McMurdie PJ, Holmes S. phyloseq: An R Package for Reproducible Interactive
601 Analysis and Graphics of Microbiome Census Data. PLOS ONE. 2013;8(4):e61217.
602 42.   Davis NM, Proctor DM, Holmes SP, Relman DA, Callahan BJ. Simple statistical
603 identification and removal of contaminant sequences in marker-gene and metagenomics
604 data. Microbiome. 2018;6(1):226.
605 43.   Chao A, Gotelli NJ, Hsieh TC, Sander EL, Ma KH, Colwell RK, et al. Rarefaction and
606 extrapolation with Hill numbers: a framework for sampling and estimation in species diversity
607 studies. Ecol Monogr. 2014;84(1):45-67.
608 44.   Oksanen J, Kindt R, Legendre P, O'Hara B, Simpson GL, Solymos P, et al. vegan:
609 Community Ecology Package. 2008.
610 45.   Holm S. A Simple Sequentially Rejective Multiple Test Procedure. Scand J Stat.
611 1979;6(2):65-70.
612 46.   Hijmans RJ. geosphere: Spherical Trigonometry. R package version 1.5-18 ed2022.
613 47.   Paradis E, Schliep K. ape 5.0: an environment for modern phylogenetics and
614   evolutionary analyses in R. Bioinformatics. 2019;35:526-8.
615 48.   Smith MR. TreeTools: create, modify and analyse phylogenetic trees. 1.10.0 ed:
616 Comprehensive R Archive Network; 2019.
617 49.   R CT. R: A language and environment for statistical computing. In: R Foundation for
618 Statistical Computing V, Austria, editor. 2021.
619 50.   Huson DH, Scornavacca C. Dendroscope 3: An Interactive Tool for Rooted
620 Phylogenetic Trees and Networks. Syst Biol. 2012;61(6):1061-7.
621 51.   Busch K, Hanz U, Mienis F, Mueller B, Franke A, Roberts EM, et al. On giant
622 shoulders: how a seamount affects the microbial community composition of seawater and
623 sponges. Biogeosciences. 2020;17(13):3471-86.
624 52.   Burgsdorf I, Erwin PM, López-Legentil S, Cerrano C, Haber M, Frenk S, et al.
625 Biogeography rather than association with cyanobacteria structures symbiotic microbial
626 communities in the marine sponge *Petrosia ficiformis*. Frontiers in Microbiology. 2014;5.
627 53.   Bayer K, Busch K, Kenchington E, Beazley L, Franzenburg S, Michels J, et al.
628 Microbial Strategies for Survival in the Glass Sponge *Vazella pourtalesii*. mSystems.
629 2020;5(4):10.1128/msystems.00473-20.
630 54.   Tian R-M, Sun J, Cai L, Zhang W-P, Zhou G-W, Qiu J-W, et al. The deep-sea glass
631 sponge *Lophophysema eversa* harbours potential symbionts responsible for the nutrient
632 conversions of carbon, nitrogen and sulfur. Environmental Microbiology. 2016;18(8):2481-
633 94.
634 55.   Engelberts JP, Robbins SJ, de Goeij JM, Aranda M, Bell SC, Webster NS.
635 Characterization of a sponge microbiome using an integrative genome-centric approach.
636 The ISME Journal. 2020;14(5):1100-10.
637 56.   Cassarino L, Coath CD, Xavier JR, Hendry KR. Silicon isotopes of deep sea
638 sponges: new insights into biomineralisation and skeletal structure. Biogeosciences.
639 2018;15(22):6959-77.
640 57.   Back CR, Stennett HL, Williams SE, Wang L, Ojeda Gomez J, Abdulle OM, et al. A
641 new *Micromonospora* strain with antibiotic activity isolated from the microbiome of a mid-
642 Atlantic deep-dea dponge. Marine Drugs. 2021;19(2):105.
643 58.   Williams SE, Back CR, Best E, Mantell J, Stach JEM, Williams TA, et al. Discovery
644 and biosynthetic assessment of *Streptomyces ortus* sp. nov. isolated from a deep-sea
645 sponge. Microbial Genomics. 2023;9(5).
646 59.   Parks DH, Chuvochina M, Rinke C, Mussig AJ, Chaumeil PA, Hugenholtz P. GTDB:
647 an ongoing census of bacterial and archaeal diversity through a phylogenetically consistent,

648   rank normalized and complete genome-based taxonomy. Nucleic Acids Res.
649   2022;50(D1):D785-D94.
650   60.   Moitinho-Silva L, Steinert G, Nielsen S, Hardoim CCP, Wu YC, McCormack GP, et
651   al. Predicting the HMA-LMA Status in Marine Sponges by Machine Learning. Frontiers in
652   Microbiology. 2017;8.
653   61.   Luter HM, Bannister RJ, Whalan S, Kutti T, Pineda M-C, Webster N. Microbiome
654   analysis of a disease affecting the deep-sea sponge *Geodia barretti*. Fems Microbiology
655   Ecology. 2017;93(6):fix074.
656   62.   Sunagawa S, Coelho LP, Chaffron S, Kultima JR, Labadie K, Salazar G, et al.
657   Structure and function of the global ocean microbiome. Science. 2015;348(6237).
658   63.   Varliero G, Bienhold C, Schmid F, Boetius A, Molari M. Microbial Diversity and
659   Connectivity in Deep-Sea Sediments of the South Atlantic Polar Front. Frontiers in
660   Microbiology. 2019;10.
661   64.   Pollock FJ, McMinds R, Smith S, Bourne DG, Willis BL, Medina M, et al. Coral-
662   associated bacteria demonstrate phylosymbiosis and cophylogeny. Nat Commun.
663   2018;9(1):4921.
664   65.   Brooks AW, Kohl KD, Brucker RM, van Opstal EJ, Bordenstein SR. Phylosymbiosis:
665   Relationships and Functional Effects of Microbial Communities across Host Evolutionary
666   History. PLoS Biol. 2016;14(11):e2000225.
667   66.   O'Brien PA, Tan SJ, Yang CT, Frade PR, Andreakis N, Smith HA, et al. Diverse coral
668   reef invertebrates exhibit patterns of phylosymbiosis. Isme J. 2020;14(9):2211-22.
669   67.   Steinert G, Rohde S, Janussen D, Blaurock C, Schupp PJ. Host-specific assembly of
670   sponge-associated prokaryotes at high taxonomic ranks. Sci Rep-Uk. 2017;7.

671

672 # Supplementary materials

673 # Diversity and structure of the deep-sea sponge
674 # microbiome in the equatorial Atlantic Ocean

675 Sam E. Williams, Gilda Varliero, Miguel Lurgi, Jem Stach, Paul R. Race and Paul Curnow
676

677 **Table of Contents**

678

690

Figure S1. Relative abundance of Archaea and Bacteria in the deep-sea sponge microbiome. Overall relative abundance is taken from the entire dataset.
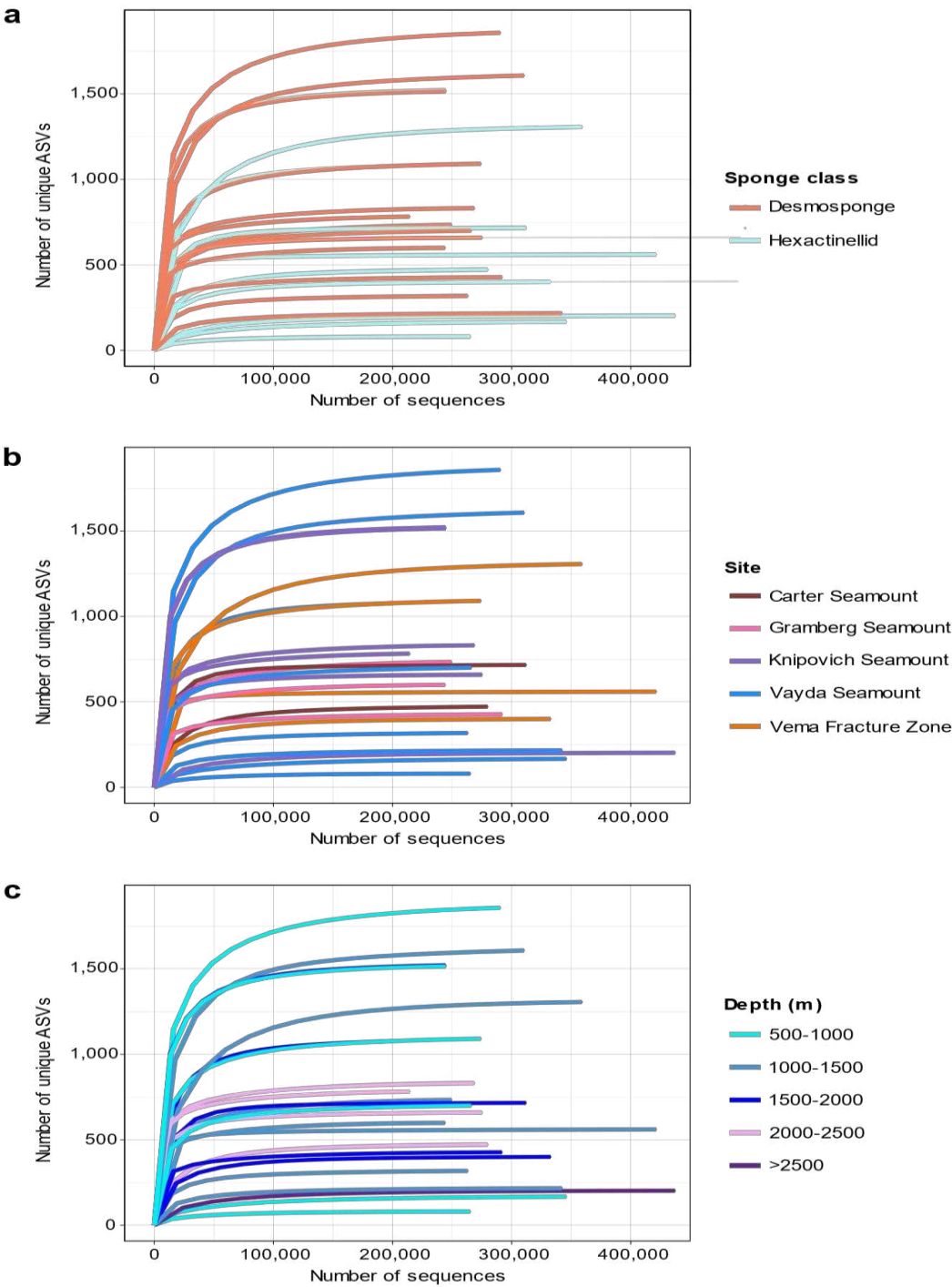
691

Figure S2. Rarefaction curves of 16S rRNA gene diversity for deep-sea Atlantic sponge samples. Displayed by a) sponge class b) site and c) depth.
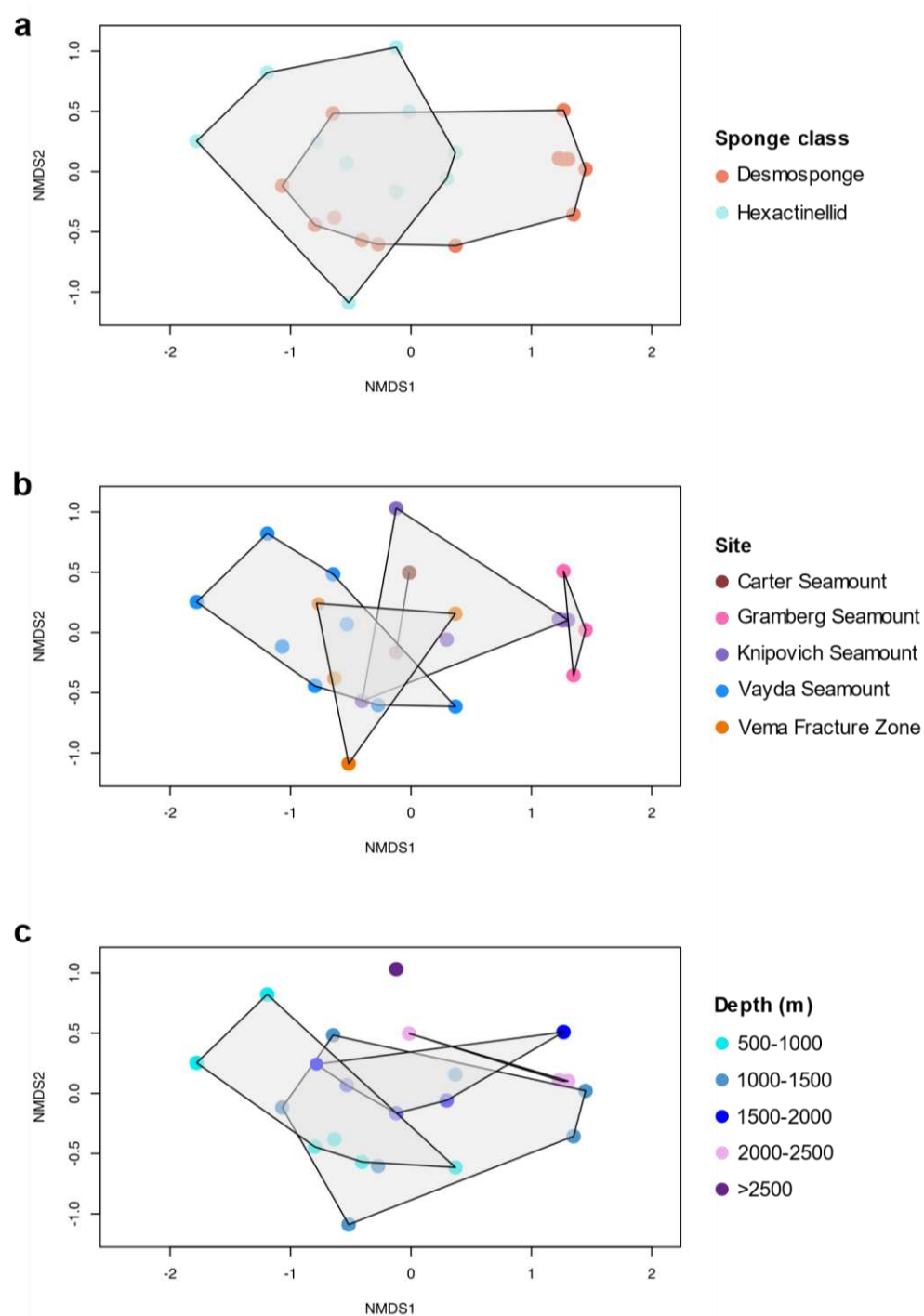
692

693

Figure S3. Compositional similarity of the deep-sea sponge microbiome in the Atlantic Ocean. NMDS performed on the Hellinger-transformed ASV dataset. Samples are grouped by (a) sponge class, (b) sampling site and (c) sampling depth.

694
695

27

696 **Table S1.** PCR primer sequences used in the study including the target gene and the

697 nucleotide sequence.

Primer name	Target gene	Sequence	Reference
515F	16S rRNA	5′-GTG YCA GCM GCC GCG GTA A-3′	(1)
926R	16S rRNA	5′-CCG YCA ATT YMT TTR AGT TT-3′	(1)
Illumina-515F	16S rRNA	5′-TCG TCG GCA GCG TCA GAT GTG TAT AAG AGA CAG GTG YCA GCM GCC GCG GTA A-3′	(2)
Illumina-926R	16S rRNA	5′-GTC TCG TGG GCT CGG AGA TGT GTA TAA GAG ACA GCC GYC AAT TYM TTT RAG TTT-3′	(2)
LCO1490	COI gene	5′-GGT CAA CAA ATC ATA AAG ATA TTG G-3′	(3)
HCO2198	COI gene	5′-TAA ACT TCA GGG TGA CCA AAA AAT CA-3	(3)
NL4F	28S rRNA	5′-GAC CCG AAA GAT GGT GAA CTA-3′	(4)
NL4R	28S rRNA	5′-ACC TTG GAG ACC TGA TGC G-3′	(4)

698

699   **Table S2.** DADA2 ASV features identified by decontam 1.8.0 as contaminants which were then removed from the final data set

DADA2_Contaminant_Feature_ID	Taxonomic Assignment
50633c4b837aa1a3b21742c0765270a4	Unassigned;NA;NA;NA;NA;NA;NA
ec1735efd303dce8c3a0043d92ff93c8	D_0__Bacteria;D_1__Acidobacteria;D_2__Subgroup 26;NA;NA;NA;NA
15157ac5595c109ea66f313a3eb5eb7a	D_0__Archaea;D_1__Thaumarchaeota;D_2__Nitrososphaeria;D_3__Nitrosopumilales;D_4__Nitrosopumilaceae;NA;NA
f21dfad54cde3762dea1d00a41729507	D_0__Archaea;NA;NA;NA;NA;NA;NA
d838a3400b85af80da105b0545a6c789	D_0__Bacteria;D_1__Planctomycetes;D_2__Planctomycetacia;D_3__Pirellulales;D_4__Pirellulaceae;D_5__uncultured;NA
d0bffe0ce5e1b758547fe2b490066b94	D_0__Bacteria;D_1__Planctomycetes;D_2__Planctomycetacia;D_3__Pirellulales;D_4__Pirellulaceae;NA;NA
f1be6b0c6e4b3136ead03bc4d5c67c79	D_0__Bacteria;D_1__Acidobacteria;D_2__Subgroup 6;D_3__uncultured bacterium AD106-G4;D_4__uncultured bacterium AD106-G4;D_5__uncultured bacterium AD106-G4;D_6__uncultured bacterium AD106-G4
b9b990ccd54fd7edbe52563acac38b90	D_0__Bacteria;D_1__Acidobacteria;D_2__Subgroup 6;NA;NA;NA;NA
a04b27b43cec356541b84c9346d36125	D_0__Bacteria;D_1__Chloroflexi;D_2__Dehalococcoidia;D_3__SAR202 clade;NA;NA;NA
63eed8bec991106ef19f974d5f2063f2	D_0__Bacteria;D_1__Chloroflexi;D_2__Dehalococcoidia;D_3__SAR202 clade;D_4__uncultured Chloroflexus sp.;D_5__uncultured Chloroflexus sp.;D_6__uncultured Chloroflexus sp.
c055aacdfae9d0169fc1c6f884292c4c	D_0__Bacteria;D_1__Deinococcus-Thermus;D_2__Deinococci;D_3__Thermales;D_4__Thermaceae;D_5__Meiothermus;D_6__uncultured organism
3464a7e5d5c33c0a88af22e879eb89a4	D_0__Bacteria;D_1__Proteobacteria;D_2__Deltaproteobacteria;D_3__Desulfarculales;D_4__Desulfarculaceae;D_5__uncultured;D_6__uncultured delta proteobacterium
d2bad679382dfdf5b82f911661bf5dfb	D_0__Bacteria;D_1__Proteobacteria;D_2__Gammaproteobacteria;D_3__KI89A clade;D_4__uncultured bacterium;D_5__uncultured bacterium;D_6__uncultured bacterium
dd606ddbf01aa42616b1135ac87f3fea	D_0__Bacteria;D_1__Proteobacteria;D_2__Gammaproteobacteria;D_3__Alteromonadales;D_4__Moritellaceae;D_5__Moritella;NA
0d53eae08491290ff50a3c98788baf5b	D_0__Bacteria;D_1__Proteobacteria;D_2__Gammaproteobacteria;D_3__P0X4b2H11;D_4__uncultured bacterium;D_5__uncultured bacterium;D_6__uncultured bacterium
5921c63ef5ace7407df7613b0113ea46	D_0__Bacteria;D_1__Proteobacteria;D_2__Gammaproteobacteria;D_3__Betaproteobacteriales;D_4__Methylophilaceae;D_5__OM43 clade;NA
927ef3265e9e3355963d4408200e31f8	D_0__Bacteria;D_1__Actinobacteria;D_2__Actinobacteria;D_3__Micrococcales;D_4__Micrococcaceae;D_5__Micrococcus;NA

700

701

702

Class	Order	Family	Genus and species (Closest relative on NCBI GenBank)	Name in this study	Samples	Sampling location	Sampling depths (m)	Method of identification (BLASTN %)
Demosponge	Desmacellida	*Desmacellidae*	*Desmacella cf. annexa*	*Desmacella sp. B1951, Desmacella sp. B01623*	2	Vema, Vayda	710, 926	CO1 (92.86, 92.15%)
Demosponge	Desmacellida	*Desmacellidae*	*Desmacella hyalina*	*Desmacella hyalina*	1	Vayda	1153	CO1 (98.37%)
Demosponge	Poecilosclerida	*Hymedesmiidae*	*Phorbas bihamiger*	*Phorbas bihamiger*	1	Vayda	1483	CO1 (97.50%)
Demosponge	Biemnida	*Rhabderemiidae*	*Rhabderemia sp. 1 RAD-2021*	*Rhabderemia sp. 1 RAD-2021*	1	Gramberg	1127	CO1 (98.47%)
Demosponge	Polymastiidae	*Polymastiidae*	*Polymastia corticata*	*Polymastia corticata*	1	Gramberg	1869	CO1 (99.85%)
Demosponge	Tetractinellida	*Vulcanellidae*	*Vulcanella aberrans*	*Vulcanella aberrans B01671, Vulcanella aberrans B01116*	2	Vayda, Knipovich	1150, 701	CO1 (99.36%, 99.38%)
Demosponge	Tetractinellida	*Geodiidae*	*Geodia barretti*	*Geodia barretti*	1	Knipovich	2257	CO1 (99.69%)
Demosponge	Tetractinellida	*Geodiidae*	*Geodia megastrella*	*Geodia megastrella*	1	Knipovich	2307	CO1 (100%)
Demosponge	Tetractinellida	*Ancorinidae*	*Stelletta tuberosa*	*Stelletta tuberosa B1955, Stelletta tuberosa B01165*	2	Gramberg, Knipovich	1460, 2307	CO1 (100%, 98.76%)
Demosponge	Trachycladida	*Trachycladidae*	*Trachycladus sp. NCI325*	*Trachycladus sp. NCI325*	1	Vayda	569	28S (97.51%)
Hexactinellid	Sceptrulophora	*Aphrocallistidae*	*Aphrocallistes beatrix*	*Aphrocallistes beatrix B0790, Aphrocallistes beatrix B01657*	2	Vayda	742, 865	CO1 (99.54, 99.69%)
Hexactinellid	Lyssacinosida	*Euplectellidae*	*Hertwigia sp. MD-2008*	*Hertwigia sp. MD-2008*	1	Vema	1140	28S (97.09%)
Hexactinellid	Sceptrulophora	*Farreidae*	*Farrea sp. USNM 1122183*	*Farrea sp. USNM 1122183*	1	Carter	1544	CO1 (99.11%)
Hexactinellid	Lyssacinosida	*Euplectellidae*	*Euplecterlla sp. HBOI*	*Euplecterlla sp. B1101*	1	Knipovich	1505	CO1 (96.11%)
Hexactinellid	Sceptrulophora	*Farreidae*	*Aspidoscopulia ospreya*	*Aspidoscopulia sp. B1963*	1	Vayda	1706	CO1 (86.91%)
Hexactinellid	NA	NA	NA	NA	4	Vema, Carter, Knipovich	1302, 1648, 2318, 2618	Spicule morphology

Table S4. DADA2 denoising of raw sequencing to amplicon sequence variants for each sponge sample and two negative controls

Sample-code	Sample-ID	input	filtered	denoised	merged	non-chimeric
B0958	Hexactinellid B0958	362870	306353	306353	293968	279823
B0507	*Hertwigia* sp. MD-2008	454548	402398	402398	392163	359254
B01671	*Vulcanella aberrans* B01671	451350	370365	370365	341796	310478
B01165	*Stelletta tuberosa* B01165	411178	355792	355792	303745	214299
B1101	*Euplectella* sp. B1101	346965	281666	281666	262303	244673
B01637	*Desmacella hyalina*	326105	276846	276846	271328	263138
B0768	*Farrea* sp. USNM 1122183	403221	326306	326306	323577	312162
B01623	*Desmacella* sp. B01623	408675	353907	353907	328635	274158
B01171	*Geodia barretti*	397309	334193	334193	300019	275125
B1951	*Desmacella* sp. B1951	446100	370664	370664	327932	291343
B1963	*Aspidoscopulia sp.* B1963	317991	256963	256963	248992	200223
B01657	*Aphrocallistes beatrix* B01657	304879	272175	272175	269622	265216
B01116	*Vulcanella aberrans* B01116	433479	356167	356167	331611	244735
B0790	*Aphrocallistes beatrix* B0790	419338	362530	362530	354738	345848
B01683	*Phorbas bihamiger*	455059	358571	358571	355041	342180
B01609	Hexactinellid B01609	424806	375249	375249	370658	332808
B01175	Hexactinellid B01175	512952	456317	456317	454292	437157
B01686	*Trachycladus* sp. NCI325	402187	334188	334188	317666	266038
B1954	*Rhabderemia* sp. 1 RAD-2021	400168	345122	345122	288800	251398
B1955	*Stelletta tuberosa* B1955	474722	374922	374922	309185	244069
B01140	*Geodia megastrella*	498304	415785	415785	349302	268611
B01698	Hexactinellid B01698	564951	435208	435208	432029	423233
B1983	*Polymastia corticata*	401621	346304	346304	319266	291842
Blank	-	429850	364245	364245	355614	350173
Gel-extract-neg	-	2269	662	662	462	398

31

708 **Table S5.** Alpha diversity metrics of each deep-sea sponge sample ordered by Observed
709 ASVs.

Sponge sample	Observed	Chao1	Shannon	InvSimpson
Desmacella sp. B1951	1864	1887.90	4.50	11.35
Vulcanella aberrans B01671	1610	1644.71	2.77	3.55
Euplectella sp. B1101	1534	1552.55	4.29	7.97
Vulcanella aberrans B01116	1530	1537.48	4.29	9.52
Hertwigia sp. MD 2008	1271	1294.95	1.30	1.44
Desmacella sp. B01623	1098	1109.23	2.96	3.32
Aspidoscopulia sp. B1963	1086	1102.60	2.80	4.33
Geodia megastrella	851	853.76	4.30	19.17
Stelletta tuberosa B01165	784	788.36	4.28	16.75
Rhabderemia sp. 1 RAD 2021	739	741.25	3.87	6.80
Farrea sp USNM 1122183	734	736.15	1.81	2.13
Trachycladus sp. NCI325	706	710.50	2.20	3.35
Geodia barretti	669	670.06	2.54	2.63
Stelletta tuberosa B1955	595	595.61	3.96	10.67
Hexactinellid B01698	566	568.14	3.24	8.48
Hexactinellid B0958	481	495.44	1.05	1.69
Polymastia corticata	439	440.00	3.14	5.34
Hexactinellid B01609	411	412.27	1.78	3.19
Desmacella hyalina	312	315.06	1.18	1.92
Phorbas bihamiger	216	218.00	1.43	3.22
Hexactinellid B01175	202	202.79	0.22	1.06
Aphrocallistes beatrix B0790	150	151.00	0.93	1.94
Aphrocallistes beatrix B01657	80	80.43	0.84	1.99

710 **References:**
711
712 1. Parada AE, Needham DM, Fuhrman JA. Every base matters: assessing small
713 subunit rRNA primers for marine microbiomes with mock communities, time series and
714 global field samples. Environ Microbiol. 2016;18(5):1403-14.
715 2. Thompson LR, Sanders JG, McDonald D, Amir A, Ladau J, Locey KJ, et al. A
716 communal catalogue reveals Earth's multiscale microbial diversity. Nature. 2017.
717 3. Folmer O, Black M, Hoeh W, Lutz R, Vrijenhoek R. DNA primers for amplification of
718 mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates. Mol Mar
719 Biol Biotechnol. 1994;3(5):294-9.
720 4. Nichols SA. An evaluation of support for order-level monophyly and interrelationships
721 within the class Demospongiae using partial data from the large subunit rDNA and
722 cytochrome oxidase subunit I. Molecular Phylogenetics and Evolution. 2005;34(1):81-96.
723
724