Changes in the metabolic potential of the sponge microbiome under ocean acidification

Emmanuelle S. Botté¹, Shaun Nielsen², Muhammad Azmi Abdul Wahab¹,4, John Webster², Steven Robbins³, Torsten Thomas² & Nicole S. Webster¹,3

Anthropogenic CO₂ emissions are causing ocean acidification, which can affect the physiology of marine organisms. Here we assess the possible effects of ocean acidification on the metabolic potential of sponge symbionts, inferred by metagenomic analyses of the microbiomes of two sponge species sampled at a shallow volcanic CO₂ seep and a nearby control reef. When comparing microbial functions between the seep and control sites, the microbiome of the sponge Stylissa flabelliformis (which is more abundant at the control site) exhibits at the seep reduced potential for uptake of exogenous carbohydrates and amino acids, and for degradation of host-derived creatine, creatinine and taurine. The microbiome of Coelocarteria singaporensis (which is more abundant at the seep) exhibits reduced potential for carbohydrate import at the seep, but greater capacity for archaeal carbon fixation via the 3-hydroxypropionate/4-hydroxybutyrate pathway, as well as archaeal and bacterial urea production and ammonia assimilation from arginine and creatine catabolism. Together these metabolic features might contribute to enhanced tolerance of the sponge symbionts, and possibly their host, to ocean acidification.
N early one-third of anthropogenic CO$_2$ has been taken up by the oceans\textsuperscript{1}, resulting in lower concentrations of carbonic ions and higher concentrations of bicarbonate ions (HCO$_3^-$) and protons (H$^+\text{)}$). Such changes to ocean chemistry have led to a 0.1 unit reduction in seawater pH since pre-industrial times\textsuperscript{2}, a phenomenal known as ocean acidification (OA). As atmospheric CO$_2$ levels continue to rise, oceans are projected to acidify further\textsuperscript{3}, with significant consequences expected for sensitive ecosystems such as coral reefs, which depend on carbonate ions for future reef accretion\textsuperscript{4}. However, changes in biological and ecological processes resulting from OA can be difficult to predict, partly because the vast majority of studies have examined the biological and physiological effects of OA using short-term experimental exposures. An alternative approach is to investigate long-term effects of projected OA using volcanic CO$_2$ seeps which feature naturally reduced seawater pH\textsuperscript{5-6}. Research conducted at natural CO$_2$ seeps over the past decade has revealed substantial changes in the population structure of a wide range of organisms\textsuperscript{5,7,8}. Studies have assessed impacts on species diversity and abundance\textsuperscript{5,7,8}, gene expression\textsuperscript{9}, as well as functional potential of eukaryotic communities\textsuperscript{10}. From a microbial perspective, OA has also been shown to modify the composition of the microbiome in some reef species\textsuperscript{11-13}.

Microbes play vital roles in the health and survival of all marine organisms\textsuperscript{11,14,15}. One of the best understood reef microbial symbioses is that which occurs between sponges and their associated bacteria and archaea. Sponges often host abundant, diverse and species-specific microbial communities\textsuperscript{16}, thought to undertake essential roles in nutrient cycling, vitamin production and natural product biosynthesis\textsuperscript{17,18}. Shifts in microbial communities have been detected in sponges collected from a natural CO$_2$ seep at Upa-Upasina within the Milne Bay Province, Papua New Guinea (PNG). At this location, the shallow reef comprises a control site with “ambient” pH (median 8.01) representing a healthy and diverse reef system, and a seep site located 500 m away, featuring “reduced” pH (median 7.77) caused by the release of > 99% pure CO$_2$ gas, which has resulted in significantly reduced species diversity and habitat complexity\textsuperscript{5,19}. Two sponge species exhibit drastically different distribution and abundance patterns at Upa-Upasina. The seep-tolerant Coelocarteria singaporensis is 40-times more abundant at the seep, whereas the seep-sensitive Styliella flabelliformis is 6-times more abundant at the control site\textsuperscript{5}. Microbial community analysis revealed variation in the microbiome of C. singaporensis between sites, whereas the microbial communities of S. flabelliformis were largely conserved across both environments\textsuperscript{13}.

Understanding how the sponge microbiome responds to long-term environmental change has been recognised as critical for anticipating the fate of sponge symbioses under predicted climate scenarios\textsuperscript{20-22}. Yet, whilst changes in microbial community composition have been found under OA, the impact on the functional potential of symbiotic communities is currently unknown.

To address this question, we examined the microbiomes of C. singaporensis and S. flabelliformis at Upa-Upasina. We specifically investigated whether the functional potential of the symbiotic microbiomes varied between control and seep sites and if these changes could contribute to maintaining a successful symbiosis in acidified oceans.

Here we find that the microbiome of the seep-tolerant C. singaporensis exhibits greater variation in functional potential between the seep and control sites than the microbiome of the seep-sensitive S. flabelliformis. At the seep, the S. flabelliformis microbiome displays an overall loss of functional potential. In contrast, the microbiome of C. singaporensis shows an enhanced potential for energy-efficient carbon fixation and nitrogen metabolism, with a greater capacity to exploit excess inorganic carbon originating from the seep environment, as well as host-derived creatine. We hereby identify microbial functions potentially contributing to the success of a sponge species in naturally acidified waters.

Results

Microbiome overview. Shotgun metagenome sequencing revealed that both sponge species harboured microbial communities that were distinct from each other and the surrounding seawater (Supplementary Note 1 and Supplementary Fig. 1). The C. singaporensis microbiome was dominated by Proteobacteria, Chloroflexi and Thaumarchaeota whereas the S. flabelliformis microbiome was largely comprised of Proteobacteria and Thau- marchaeota. Microbial community composition varied between the control site and the adjacent CO$_2$ seep site, with statistically significant differences found at multiple taxonomic levels, particularly for C. singaporensis (Supplementary Note 1 and Supplementary Fig. 2). The differences in microbial communities between the two species and the major shift in the composition of the C. singaporensis microbiome between sites confirmed previous findings\textsuperscript{13}. However, the exclusion of Archaea from these previously published analyses precludes direct comparisons of the datasets.

Differences between the two species were also found at the functional level (Supplementary Fig. 3, Supplementary Tables 1–2, Supplementary Note 2 and Supplementary Note 3). Functions that were more abundant at either the control or seep site, regardless of p-value, are hereafter termed ‘enriched’. Twenty percent of significantly enriched genes (KO annotation) in the C. singaporensis microbiome were more abundant at the seep, as were 80% of significantly enriched genes in the S. flabelliformis microbiome (Fig. 1a and Supplementary Table 1). However, genes that were enriched at the seep were generally an abundant component of the metagenome of C. singaporensis but a minor component of the S. flabelliformis metagenome. The total relative abundance of seep-enriched genes (COG annotation) was 68 and 19% for C. singaporensis and S. flabelliformis, respectively (Fig. 1b). In contrast, the total relative abundance of genes enriched at the control site was 65% and 34%, respectively. A correlation vector of environmental variables over ordination of KOs revealed that the functional profiles were primarily driven by factors associated with inorganic carbon chemistry (Supplementary Note 4, Supplementary Fig. 4 and Supplementary Tables 3–6).

Transport of carbon and host-derived compounds. One of the most abundant functions at the seep in the microbiomes of both sponge species was related to ATP-Binding Cassette (ABC) transporters (Fig. 1c, Fig. 2, Supplementary Note 5, Supplementary Fig. 5 and Supplementary Fig. 6). Total organic carbon (TOC) in seawater was consistent between sites\textsuperscript{7} (Supplementary Table 4); however, microbiomes of both C. singaporensis and S. flabelliformis exhibited a reduced potential for uptake of carbohydrates at the seep, suggesting that the microbial community derives less benefit from external energy sources and carbon at the seep than at the control site\textsuperscript{22}. The C. singaporensis microbiome exhibited a reduction in the relative abundance of genes encoding ABC importers for D-xylene (xy1FGH, 4.9-times), ribose (RbsABC, 2.9-times), polysaccharides (putative transporter, lplABC, 15-times), alpha-1,4-digalactorurate (aflaeFG, 6.3-times), arabinosaccharide (araNPQ, 5.2-times), lactose/l-arabinose (lactEFGK, 2-times), raffinose/stachyose (msmEFG, 2-times), rhamnose (rhapQST, absent from the seep samples) and Sn-glycerol-3-phosphate (ugpABC, 4.3-times), which provides a
source of phosphate in addition to carbon\textsuperscript{23}. In the \textit{S. flabelliformis} microbiome, ABC transporter genes with significantly reduced relative abundance at the seep were involved in the import of glucose/mannose (\textit{gltEFG}, 1.5-times), multiple sugars (several systems, 1.8-times), Sn-glycerol-3-phosphate (\textit{ugpABCE}, 2.2-times), trehalose (\textit{thaEFG}, 2-times) and the osmoprotectant glycerol (\textit{glpPQSTV}, 2-times). The \textit{S. flabelliformis} microbiome also exhibited reduced capacity for the uptake of several amino acids and other compounds that could serve as energy sources, with a lower relative abundance of importers for arginine/ornithine (\textit{aatJMPQ}, 1.3-times), branched-chain amino acids (\textit{livFGHKM}, 1.4-times), general L-amino acids (\textit{appJMPQ}, 1.5-times), taurine (\textit{tuaABC}, 1.4-times), spermidine (\textit{potABCD}, 1.6-times), and thiamine (\textit{thiBPQ}, 2.3-times). Other genes related to transport and which differed between the two species are further discussed in Supplementary Note 5. The microbial communities of \textit{C. singaporenensis} and \textit{S. flabelliformis} at the seep appear to have a restricted capacity to use exogenous organic compounds for energy production and biosynthesis, suggesting that sponge hosts, in particular \textit{S. flabelliformis}, might retain carbohydrates and amino acids derived from filter-feeding and de novo biosynthesis, rather than providing them to the symbionts.

\textbf{Carbon fixation and carbonic anhydrase.} The concentrations of bicarbonate ion (HCO\textsubscript{3}⁻) at Upa-Upasina were significantly higher at the seep than at the control site (1887 ± s.d. 140 μmol kg\textsuperscript{-1} at the seep vs. 1661 ± s.d. 45 μmol kg\textsuperscript{-1} at the control, Supplementary Table 4). An increased potential for carbon fixation via the HP/HB cycle of Thaumarchaeota was found in \textit{C. singaporenensis} at the CO\textsubscript{2} seep, providing support for possible metabolic adaptation of the microbiome to the seep environment. The 3-hydroxypropionate/4-hydroxybutyrate (HP/HB) cycle typically uses bicarbonate ion to produce acetyl-coA and succinyl-coA, which can both be used for biosynthesis via further conversion into oxaloacetate, pyruvate or oxoglutarate-CoA\textsuperscript{24}. A two to four-times increase in the relative abundance of genes encoding key HP/HB cycle enzymes was found at the seep, including for malonic semialdehyde reductase (3.9-times, ANOVA, \textit{p}=0.03), \textit{3-hydroxypropionyl-coA synthase} (3-times, ANOVA, \textit{p}>0.05), \textit{3-hydroxypropionyl-coA dehydratase} (1.9-times, ANOVA, \textit{p}>0.05), \textit{4-hydroxybutyryl-CoA synthase} (2.8-times, ANOVA, \textit{p}>0.05), \textit{4-hydroxybutyryl-CoA dehydrogenase} (4-times, ANOVA, \textit{p}>0.05), \textit{3-hydroxybutyryl-CoA dehydrogenase} (3.8-times, ANOVA, \textit{p}<0.01) (Fig. 3a and Supplementary Fig. 8). These genes were almost all assigned to the Thaumarchaeota genus \textit{Nitrosopumilus} (Supplementary Table 7 and Supplementary Note 13), explaining the lack of genes encoding malonyl-coA reductase, acryloyl-coA reductase, succinyl-coA reductase and succinate semialdehyde reductase, which are yet to be identified from any Thaumarchaeota\textsuperscript{25}. \textit{Nitrosopumilus maritimus} has been reported to use the most energy-efficient version of the HP/HB cycle described to date\textsuperscript{26}. These findings suggest that...
ATP-Binding Cassette (ABC) transporters significantly increased on rare occasions in carbonic anhydrase (COG0288, 1.5-times and COG0338, 5.2-times, respectively), reducing relative abundance of two genes encoding carbonic anhydrase in C. singaporensis microbiome.

Energy-convertible via creatininase, for which the encoding gene was previously reported to be abundant in the microbiomes of both species, favoring energy-efficient autotrophy. However, genes encoding a series of enzymes involved in creatinine degradation were less abundant at the seep site for both species. In the microbiome of C. singapurensis these genes encode (i) cytosine deaminase (1.4-times, ANOVA, \( p < 0.05 \)), which hydrolyses creatine into N-methylhydantoin with the release of a molecule of ammonia, (ii) the two subunits of N-methylhydantoinase (3.1-times and 3.4-times increase, ANOVA, both \( p < 0.01 \)), which uses a molecule of ATP to hydrolyse N-methylhydantoin into N-carbamoylsarcosine and releases a second molecule of ammonia, (iii) N-carbamoylsarcosine oxidase (absent from all seep samples), which deaminates N-carbamoylsarcosine into sarcosine and releases a second molecule of ammonia (Fig. 3b). In the S. flabelliformis microbiome, the genes encoding N-methylhydantoinase subunits were also significantly less abundant at the seep site (1.3-time and 1.9-time, respectively, ANOVA, \( p < 0.01 \)) for both (Fig. 3b, Supplementary Fig. 13 and Supplementary Table 11). Simultaneously, C. singapurensis symbionts displayed a greater potential for creatine degradation, but 2-times decrease in creatinine concentrations at the CO2 seep may favour energy-efficient autotrophy by Thaumarchaeota in the C. singapurensis microbiome.

Another potential marker for adaptation to the high bicarbonate concentrations at the seep was the increased bicarbonate concentration at the CO2 seep, supporting our data for each sponge species. Asterisk denotes \( p < 0.05 \) according to ANOVA. The higher HCO\(_3\) concentrations at the CO2 seep may favour energy-efficient autotrophy by Thaumarchaeota in the C. singapurensis microbiome.

Nitrogen metabolism. OA had a large effect on nitrogen cycling genes in the microbiomes of both species, favouring energy-efficient biochemical pathways with creatine and urease degradation in C. singapurensis and causing a decrease in the capacity to use key nitrogenous compounds in S. flabelliformis (Fig. 3b, c, Supplementary Note 8, Supplementary Fig. 10 and Supplementary Fig. 13). Sponges can produce creatine and creatinine, with previous reports suggesting that these sponge-derived compounds can fuel carbon and nitrogen metabolism in symbiotic Archaea. Both compounds are interconvertible via creatinase, for which the encoding gene did not vary in relative abundance between sites for either species. However, genes encoding a series of enzymes involved in creatinine degradation were less abundant at the seep site for both species. In the microbiome of C. singapurensis these genes encode (i) cytosine deaminase (1.4-times, ANOVA, \( p < 0.05 \)), which hydrolyses creatine into N-methylhydantoin with the release of a molecule of ammonia, (ii) the two subunits of N-methylhydantoinase (3.1-times and 3.4-times increase, ANOVA, both \( p < 0.01 \)), which uses a molecule of ATP to hydrolyse N-methylhydantoin into N-carbamoylsarcosine and (iii) N-carbamoylsarcosine oxidase (absent from all seep samples), which deaminates N-carbamoylsarcosine into sarcosine and releases a second molecule of ammonia (Fig. 3b). In the S. flabelliformis microbiome, the genes encoding N-methylhydantoinase subunits were also significantly less abundant at the seep site (1.3-time and 1.9-time, respectively, ANOVA, \( p < 0.01 \)) for both (Fig. 3b, Supplementary Fig. 13 and Supplementary Table 11). Simultaneously, C. singapurensis symbionts displayed a greater potential for creatine degradation at the seep with a 5-times increase in the relative abundance of the gene encoding creatinase (ANOVA, \( p < 0.01 \)), which produces urease and sarcosine, and a non-significant, but 2-times increase (from 0.02% to 0.04%) in genes encoding all urease subunits and accessory proteins, which release two molecules of ammonia and one molecule of CO2, without using ATP. C. singapurensis thus might preferentially transfer creatine to its symbionts over creatinine at the CO2 seep, supporting our findings that energy conservation is an important feature of successful sponge-microbes symbiosis in an acidified ocean.
Fig. 3 Microbial metabolic pathways related to carbon, nitrogen and sulfur metabolism. a Carbon fixation via HP/HB cycle (see Supplementary Methods for detail of KO assignment to functions). b and c Degradation and utilisation of nitrogen compounds. d Sulfide assimilation following taurine degradation. P-values (represented by asterisks) were derived from ANOVA. Line colours represent taxonomic origin and coloured circles represent ratios of relative abundances, all according to the colour scale described in the figure.
In contrast, the *S. flabelliformis* microbiome exhibited a considerable proportion of total free amino acids in sponges. OA negatively impacted the metabolic sulfur metabolism, more so in sponges than in archaeal species. These findings are consistent with previous studies showing that OA can impact the sulfur cycle in aquatic environments, particularly in aquatic animals that rely on sulfur metabolism for energy production.

Together these results indicate that at the control site ammonia production via urea is predominantly supported by bacterial-driven creatinine metabolism, whereas archaeal-driven arginine metabolism and energy-efficient bacterial-driven creatine catalysis predominates in the *C. singaporensis* microbiome at the seep site.

**Sulfur metabolism.** OA negatively impacted the metabolic potential for taurine dissimilation and sulfur assimilation in *S. flabelliformis*, which may reflect reduced host production of taurine in acidified seawater. Indeed, taurine can represent a considerable proportion of total free amino acids in sponges and the microbiome of the seep-sensitive species presented a significant decrease in the relative abundance of genes encoding three enzymes in the taurine degradation pathway: taurine-pyruvate aminotransferase (*tpa*, 2.3-times, ANOVA, *p* < 0.01), which degrades taurine into sulfoacetaldehyde and L-alanine, sulfoacetaldehyde acetyltransferase (*xsc*, 1.6-times, ANOVA, *p* < 0.01), which produces sulfit and acetate phosphate, and cysteine (1.4-times, ANOVA, *p* = 0.02), which converts sulfit into sulfide (Fig. 3d and Supplementary Fig. 14). Most of the scaffolds carrying *tpa* had sequence similarity to Thiotrichales and unclassified Gammaproteobacteria, corroborating the taxonomic results (Supplementary Fig. 2, Supplementary Table 12 and Supplementary Note 9). Scaffolds carrying *xsc* were assigned to diverse taxa, as sulfoacetaldehyde can also result from the decarboxylation of sulfopyruvate in the sulfolactate degradation pathway. The bacterial orders Burkholderiales (Betaproteobacteria) and Oceanospirillales (Gammaproteobacteria) represented a lower proportion of the taxa carrying the *xsc* gene at the seep and therefore appear to be primarily responsible for the decline in relative gene abundance of *xsc* (Supplementary Table 12). *S. flabelliformis* microbiomes also exhibited a lower potential for sulfur assimilation at the seep, as genes composing two pathways for the conversion of L-homoserine into L-homocysteine were significantly less abundant in the CO23 seep (Fig. 3c). Three of the four genes encoding these pathways’ enzymes, namely homoserine O-succinyl-transferase (*meta*), cystathionine-β-lyase (*metC*) and O-succinyl-homoserine sulphydrylase (*metZ*), were significantly less abundant at the seep, with 1.9-times, 1.6-times and 1.5-times changes, respectively (ANOVA, *p* < 0.01, *p* = 0.01 and *p* < 0.01 respectively) (Fig. 3c). Taxa potentially contributing to these pathways include the orders Burkholderiales (Betaproteobacteria), Methylococcales (methanotrophic Gammaproteobacteria), Oceanospirillales (heterotrophic Gammaproteobacteria) and Chromatiales (anoxicogenic photosynthetic Gammaproteobacteria) (Supplementary Table 12). Ocean acidification thus reduces the potential for the *S. flabelliformis* microbiome to use taurine to produce sulfit and subsequently assimilate it into biomass. Whether driven by pH-induced reductions in host taurine production or through a reduced potential for microbial dissimilation, this reduced capacity for sulfur metabolism could contribute to the poor performance of *S. flabelliformis* in acidified waters.

**Discussion**

Our study characterises the potential effects of OA on marine microbial symbioses. However, metagenomic sequencing has several limitations that need to be acknowledged. Whilst differences in the relative abundance of specific pathways indicate a shift in functional potential within the microbiome, these pathways need to be expressed to effectively alter metabolism. For example, despite considerable evidence that Thaumarchaeota oxidise ammonia and emerging research confirming they also metabolise urea, it is possible that increased relative abundances of these pathways within the *C. singaporensis* microbiome do not translate into changes in metabolism. Future studies should therefore utilise complimentary transcriptomic, proteomic, metabolomic approaches or stable isotope probing to unequivocally validate altered symbiont metabolism under OA. Similarly, we acknowledge limitations in using results derived from CO2 seeps for inferring impacts under future acidified oceans. Firstly, changes in oceanic carbonate chemistry will occur concomitantly with other stressors including ocean warming, pollution, overfishing and habitat destruction, with the potential for additive, mitigating or synergistic effects. However, the Upa-Upasina location is not exposed to anthropogenic pressures such...
as pollution or eutrophication and occasional mixing of the water column during high-wind seasons means that organisms may experience periods of respite from low pH conditions at the seep. Secondly, in addition to changes in functional potential of the microbiome, other factors such as habitat structure, competition for space, predation, host metabolism, reproductive success or larval settlement rates could contribute to the differential abundance of the two sponge species across sites, particularly the striking 40-times increase in C. singaporensis at the seep. Additional information on ecological and physiological characteristics of C. singaporensis is now needed to tease apart the relative contributions of each of these factors in the success of this species at the Upa-Upasina seep.

Mindful of these considerations, we nevertheless uncover an enhanced functional potential for energy-efficient carbon and nitrogen metabolism in a sponge species that is tolerant of OA conditions projected for 2100 and show a general loss of functional potential in an OA-sensitive species (Fig. 4). Host metabolic adaptation has been detected as a stress response of sponges. Remarkably, a similar compensatory mechanism might also be occurring at the microbial scale, with the microbiome of OA tolerant species enriched in energy-efficient pathways that could enable excess energy to be utilised elsewhere as needed. Given the key roles that microbes play in host health, metabolic adaptation of the microbiome could also confer tolerance to future acidified oceans.

**Methods**

**Sample collection.** Samples were collected in the Milne Bay province of Papua New Guinea at Upa-Upasina in March 2014 following published methods and as part of Expedition 37. In detail, three biological replicates of C. singaporensis and S. flabelliformis were collected on SCUBA at the seep and control sites (500 m apart) and brought back to the surface in individual plastic bags filled with seawater from the collection point. Small tissue samples from each specimen were preserved in 70% ethanol for species identification and the remaining tissue was snap-frozen for microbial enrichment and metagenomic analysis. Twelve litre seawater samples were also collected in triplicate from the control and seep sites and successively filtered through 5 μm and 0.22 μm Sterivex filters, which were subsequently stored at −20 °C until processing. All samples were transported to the Australian Institute of Marine Science according to AQIS regulations (permit number 14002493). Sponge species were identified on the basis of morphological characters by taxonomic specialists at Western Australian Museum and conformed to original descriptions for C. singaporensis and S. flabelliformis.

**Sample processing and sequencing.** Microbial cells were enriched following established procedures with minor modifications. In detail, sponge tissue samples transferred to sterile Falcon tubes (Corning, New York, USA) were rinsed twice in a shaker set to 200 rotations per minute (rpm) for 5 min in ~10 mL of sterile Calcium/Magnesium-Free Sea Water (CMFSW) containing 0.4 M NaCl, 10 mM KCl, 7 mM NaHCO3 and 0.05 mM Na2SO4 and 0.05 mM NaHCO3 in milliQ (mQ) water. The rinsed tissue was cut into small (~1 cm³) pieces using a 70% ethanol-rinsed and CMFSW-rinsed Tupperware TurboChef (Orlando, Florida, USA) and homogenised in ~10–12 mL of CMFSW for a maximum of 10 min with a SilentCrusher M homogeniser (Heidolph, Schwabach, Germany). All instruments and equipment used were thoroughly washed with 70% ethanol and CMFSW between samples. Sterile collagenase (Sigma-Aldrich, St Louis, Missouri, USA) diluted into sterile CMFSW was added to each sample at a concentration of 0.5 g L⁻¹ before incubation at 150 rpm for 30 min. Each sample was filtered through a 100 μm Nylon sterile cell strainer (Corning, New York, USA) and...
centrifuged for 15 min at 100 × g at 4 °C. The supernatant was collected and further centrifuged for 15 min at 300 × g at 4 °C. The newly formed supernatant was diluted with sterile sodium pyrophosphate buffer (10 mM Na2HPO4, pH 7.8) and centrifuged again at 5000 × g for 20 min at 4 °C. The pellet was resuspended in 80% ethanol in 0.5 M NaCl. The DNA was precipitated by adding 0.2 M sodium acetate and incubating at 4 °C for 30 min. After precipitation, the DNA was resuspended in 10 mM Tris(HCl) pH 8 and 0.5 M NaCl and centrifuged at 16,000 × g for 5 min at 4 °C. The DNA contained in the supernatant was precipitated in a fresh tube by adding 0.1 volume sodium acetate 3 M and 0.6 volume 100% isopropanol and incubating for 10 min at room temperature. The DNA was pelleted by centrifugation at 16,000 × g for 30 min. After centrifugation, the supernatant, 1% of 70% ethanol was added. A final centrifugation at 16,000 × g for 15 min was conducted. Ethanol was removed and left to completely evaporate. The clean DNA was resuspended in 20 μL of MQ water and left at 4 °C overnight, then stored at −20 °C. DNA quality and concentrations were checked using Nanodrop and Qubit. Nextera XT library preparation was performed on all samples at the Ramaciotti Centre for Genomics (University of New South Wales, Australia) before multiplexing and sequencing across four lanes of Illumina MiSeq (2 × 250 bp).

Metagenomes assembly. Raw reads were quality filtered with Prinseq to discard sequences with ambiguous bases or those that were shorter than 30 bp. Reads were assembled using IDBA-UD with default parameters and the pre-correction option. Reads were mapped back to the assembled genome using Bowtie2 to assess assembly quality and obtain coverage information. The resulting assembly was processed in Supplementary Table 18. Raw reads were submitted to the Sequence Read Archive (accession number SRP159543) and can be found under SRA accession number SRP159543 (Supplementary Table 19).

Taxonomic assignment. Taxonomy was assigned to 16S rRNA subunit genes extracted from the metagenomic assemblies using GraffT. Taxonomic resolution at the species level was not always possible, resulting in some scaffolds only being assigned to higher ranks. GraffT packages can be found at [https://data.aace.ues.edu.au/public/grafft/7/].

Functional annotation. Scaffolds were submitted to IMG/MER for annotation along with coverage information to enable the estimation of gene copy numbers. Scaffolds identified by IMG to be of Eukaryotic origin were removed. Estimated copy numbers of clusters of Orthologous Genes (COGs) and KEGG Orthology (KO) terms were retrieved from the KEGG database after annotation through the IMG pipeline using the “standard gene calling methods” protocol from IMG. Counts of individual COGs and KEGGs were used for statistical analyses. Analysis of the KOs and COGs was performed over three stages: (i) for each species, KO and COG terms were assigned; (ii) significantly enriched KOs were grouped into KEGG pathways; (iii) KO and COG terms were assigned to several categories and were not “dereplicated” but left in all categories assigned. KO and COG terms were assigned to several categories and were not “dereplicated” but left in all categories assigned. KO and COG terms were assigned to several categories and were not “dereplicated” but left in all categories assigned. KO and COG terms were assigned to several categories and were not “dereplicated” but left in all categories assigned.

Statistical analysis of biological data. Rarefaction curves were generated for taxonomic and gene datasets to assess differences in coverage between samples and for quality control. Bray-Curtis similarity matrices were generated for genes (KOs) using the R package Vegan. Variation amongst samples was visualised using hierarchical clustering, as well as Principal Coordinate Analysis using the “PERMANOVA+ for Primer” software package (Supplementary Fig. 3). Cluster analysis confirmed clear separation between seawater and sponge samples and greater partitioning between control and seep samples for sponges compared to seawater. A Permutational Multivariate Analysis of Variance (PERMANOVA) was performed on the Bray-Curtis matrix of KOs and showed an interaction between site and species in the partitioning of the samples (p = 0.0237, Supplementary Table 2), which enabled us to perform a more constrained analysis of microbial functional diversity in relation to environmental variables (see section here). Given the focus of this manuscript, no further analyses were performed on the seawater samples. Gene copy numbers were rarefied to the lowest number across samples for each species to account for differences in sequencing depth. Mean and standard deviations were calculated on percentage relative abundance of rarefied data per species. Counts of taxa and genes were modelled using negative binomial generalised linear models with site (control vs. seep) as a fixed explanatory covariate and the total gene copy number as an offset term.

Statistical analysis of environmental data. The control and seep sites at the Upa-Upasina location exhibited near-identical temperatures (annual averages of 26.9 °C ± 1.37 and 28.82 °C ± 1.38, respectively) and salinities (annual averages of 35.01 ppt ± 0.63 and 34.77 ppt ± 0.54, respectively). Metrics for the complete suite of environmental variables and water chemistry recorded from the Upa-Upasina site are shown in Supplementary Tables 3–5, along with results from individual t-tests performed on each of these variables to determine any significant difference between sites.

Microbial functions in relation to environmental variables. To determine the role of environmental variables in driving the partitioning of samples between sites, the mean values derived from twenty-one environmental parameters were assessed against the mean Bray-Curtis matrix of KOs using a canonical analysis of principle coordinates (CAP). Mean values of KOs within each sample were used for the analysis to align with environmental data derived at the site level. The CAP function in PRIMER v7 performed distance-based canonical correlation of the KOs to the environmental data, thereby constraining the ordination according to variables most significantly driving differences between samples (Supplementary Note 4, Supplementary Fig. 4 and Supplementary Table 6).

Received: 31 March 2019 Accepted: 15 August 2019
Published online: 12 September 2019

ARTICLE NATURE COMMUNICATIONS | https://doi.org/10.1038/s41467-019-12156-y | www.nature.com/naturecommunications
Acknowledgements

The authors thank Katharina Fabricius for facilitating this study within the March–April 2014 expedition she led. We thank crew members of the M/V Chertan for transport to the field site and the community of Upa-Upasina for welcoming us and allowing us to carry science in their reef. Sam Noonan provided logistics and field assistance. The Ramaciotti Centre for Genomics generated sequencing data. Pam Engelberts facilitated analyses on the ACE server. Glynn Gorick and E.S.B. produced Fig. 4. We are very grateful to Jean-Baptiste Raina for constructive criticism and feedback on the paper. This study was funded by the Australian Institute of Marine Science.

Author contributions

E.S.B. and N.S.W. designed the study. E.S.B. and M.A.A.W. collected field samples. M.A.A.W. carried out sponge species identification and analysis of environmental variables. E.S.B. conducted laboratory work. E.S.B. and J.W. conducted metagenomic assemblies. E.S.B. performed all bioinformatics analyses with guidance from S.R. S.N. performed statistical analyses. E.S.B. conceptualised Fig. 4. E.S.B., N.S.W. and T.T. interpreted the data and wrote the paper. All authors edited the paper before submission.