OPEN

Metabolic profiles of prokaryotic and eukaryotic communities in deep-sea sponge *Neamphius huxleyi* indicated by metagenomics

Zhi-Yong Li[1], Yue-Zhu Wang[2], Li-Ming He[1] & Hua-Jun Zheng[2]

[1]Marine Biotechnology Laboratory, State Key Laboratory of Microbial Metabolism and School of Life Sciences and Biotechnology, Shanghai Jiao Tong University, Shanghai 200240, P. R. China, [2]Shanghai-MOST Key Laboratory of Health and Disease Genomics, Chinese National Human Genome Center at Shanghai, Shanghai 201203, P. R. China.

The whole metabolism of a sponge holobiont and the respective contributions of prokaryotic and eukaryotic symbionts and their associations with the sponge host remain largely unclear. Meanwhile, compared with shallow water sponges, deep-sea sponges are rarely understood. Here we report the metagenomic exploration of deep-sea sponge *Neamphius huxleyi* at the whole community level. Metagenomic data showed phylogenetically diverse prokaryotes and eukaryotes in *Neamphius huxleyi*. MEGAN and gene enrichment analyses indicated different metabolic potentials of prokaryotic symbionts from eukaryotic symbionts, especially in nitrogen and carbon metabolisms, and their molecular interactions with the sponge host. These results supported the hypothesis that prokaryotic and eukaryotic symbionts have different ecological roles and relationships with sponge host. Moreover, vigorous denitrification, and CO_2 fixation by chemoautotrophic prokaryotes were suggested for this deep-sea sponge. The study provided novel insights into the respective potentials of prokaryotic and eukaryotic symbionts and their associations with deep-sea sponge *Neamphius huxleyi*.

The symbiotic phenomenon of marine organisms is common[1,2]. Marine symbioses could significantly influence the ecology, physiology and evolution of partners, for example symbioses between the marine invertebrates (*e.g.* tubeworm and mussel) and bacteria may explain the high biomasses observed in the vicinity of deep-sea hydrothermal vents and around cold seeps[2]. As the oldest multicellular animals (more than 600 million years old)[3], marine sponges (phylum Porifera) often harbor dense and diverse microbial communities including bacteria, archaea, fungi and unicellular microalgae[4–11]. Sponge-microbes symbioses have been suggested by the presence of a core microbial community and sponge-specific microbial lineages as well as the microbial vertical transmission[6,8,9]. In particular, the adhesins, adhesion-related proteins, ankyrin repeat proteins (ARP), tetratricopeptide repeat domain-encoding proteins (TPR), and transposable insertion elements observed recently in sponge metagenome and sponge bacterial genome suggest a close association of bacterial symbionts with their sponge host[12–19]. Sponges probably represent one of the most complex symbioses on earth[4–11], therefore provide an ideal model system for the biological and ecological investigation of marine symbioses.

After learning more about the diversity of sponge microbial symbionts, the function evaluation of the microbial symbionts represents the frontier and hot issue of sponge symbioses. Investigations on single strain, functional gene and genome have suggested the functions of symbiotic microbes in sponges, such as producing bioactive compounds, nitrogen cycling and carbon fixation[12–15]. Modern omics provides a promising strategy for understanding the metabolic diversity of the sponge symbionts. In 2010, Thomas *et al.* first explored the functional genomic signature of bacteria associated with the sponge *Cymbastela concentrica* by shotgun sequencing[16]. Thereafter, Liu *et al.* analyzed the bacterial functional proteins in the sponge *Cymbastela concentrica* using metaproteogenomic technique[17]. Recently, Fan *et al.* investigated the metabolisms of the bacterial communities of six sponges using metagenomics, and suggested the functional equivalence and evolutionary convergence in complex microbial communities of sponge symbionts[18]. Omics investigations revealed previously unknown diversity and functions of sponge symbionts[16–19], but to date, only bacterial communities of shallow-water sponges were involved.

It is known that, besides bacteria and archaea[5,6,8–10], sponges also harbor fungi (*e.g.* Ascomycota and Basidiomycota) and phototrophic eukaryotes (*e.g. Chlorella* and dinoflagellate)[4,20,21]. Like coral holobiont, when all the components within a sponge are considered as a whole, the term 'sponge holobiont' can be used. A huge diversity of prokaryotes has been identified in many sponges[5–10]. Prokaryotes play important roles in nutrient and energy cycling, and secondary metabolism[12–19,22,23], but the eukaryotic actors of the symbioses are not well known[4]. For instance, marine protists including algae and protozoa are important components of the oceans[24,25], but we rarely know about the protists in sponges. The same problem exists for sponge fungal symbionts. Therefore the total metabolic profile of one sponge holobiont is not clear because of the lack of information on eukaryotic symbionts.

The metabolic analysis of sponge holobionts at the whole community level including prokaryotes and eukaryotes is helpful for understanding the biology and ecology of sponge symbioses. What are the respective potentials of prokaryotic and eukaryotic symbionts in one sponge holobiont as well as their relationships with the sponge host? It is an important scientific problem needs to be answered. Meanwhile, to date, omics explorations have only focused on shallow-water sponges[16–19], the community structures and metabolic profiles of deep-sea sponge holobionts have been rarely investigated[26,27].

Neamphius huxleyi is a kind of typical deep-sea water sponge in the oceans, the natural products of this deep-water sponge have been reported[28], but its community structure and function remain nearly unknown[29]. With the aim to understand the respective metabolic potentials of prokaryotic and eukaryotic symbionts and their relationships with the sponge host as well as the deep-sea sponge holobionts, in this study, sponge *Neamphius huxleyi* from the India Ocean (*ca.* 1,800 m, temp. *ca.*4°C, salinity *ca.*34.3‰ (see Supplementary Fig. S1 online) were investigated using metagenomics. Because the deep-sea water has its unique characteristics which are different from those of shallow waters[30], *e.g.* high concentrations of nitrate and nitrite (see Supplementary Fig. S1 online), and no light for CO_2 fixation by photosynthesis, therefore N and C related metabolisms were particularly highlighted. This study provided novel insights into the whole metabolism of one deep-sea sponge holobiont, particularly the different contributions of prokaryotic and eukaryotic symbionts as well as their relationships with the sponge host were suggested.

Results

Metagenome data. A total of 72 Gb (about 652 million reads) sequences were obtained from the metagenome of the deep-sea sponge *Neamphius huxleyi*. About 2.0 million reads identified as sponge genome sequences were excluded using BlastN with parameters: identity rate 90%, alignment coverage 80%. Finally, about 55 Gb clean data (565 million reads) were used for metagenomic analysis. In total, 88,301 contigs were assembled with the minimal length of 500 bp, among which 15,472 contigs (17.8%) were above 2 kb (see Supplementary Table S1 online). By gene prediction using Gene Mark and Meta Gene Annotator, 65,634 out of 113,332 genes were homologous to the known genes.

Prokaryotic and eukaryotic components in deep-sea sponge *Neamphius huxleyi*. A total of 300,723 paired reads and 163,858 non-paired reads matched 16S rRNA gene references. The prokaryotic community was composed of archaea (Crenarchaeota and Euryarchaeota) and bacteria (19 phyla including Acidobacteria, Actinobacteria, Aquificae, Bacteroidetes, Chlamydia, Chloroflexi, Cyanobacteria, Deferribacteres, Deinococcus-Thermus, Firmicutes, Fusobacteria, Gemmatimonadetes, Lentisphaerae, Nitrospira, Planctomycetes, Proteobacteria, Spirochaetes, Thermotogae, Verrucomicrobia, and 6 candidate divisions bacteria incertae sedis, BRC1, OD1, OP10, SR1 and TM7) (see Supplementary Fig. S2a online).

Proteobacteria, including Alpha-, Gamma-, Beta-, Delta-, Epsilon-proteobacteria, dominated the prokaryotic community (80.76% of the 16S rRNA gene reads)(see Supplementary Fig. S2b online), followed by Actinobacteria, Bacteroidetes, Planctomycetes and Firmicutes. In contrast, the prokaryotes in the seawater only include Acidobacteria, Actinobacteria, Firmicutes, Proteobacteria, Euryarchaeote and Crenarchaeote based on the 16S rRNA gene libraries of bacteria and archaea (see Supplementary Fig. S2c online).

Based on the 25,549 reads matching 18S rRNA gene references, phylogenetically diverse eukaryotic symbionts were observed in this deep-sea sponge (see Supplementary Fig. S2d online) including fungi, protophyte (Eccrinales, Alveolata, Rhodophyta, Viridiplantae, Cryptophyta and Stramenopiles) and protozoa (Amoebozoa, Apusozoa, Euglenozoa and Rhizaria). Eccrinales (31.04% of the 18S rRAN gene reads), which are the members of the opisthokont protophyte Mesomycetozoea, dominated the protophyte community followed by Alveolata (27.87%). Rhizaria represented the most predominant protozoa. Fungal community observed in this deep-sea sponge consisted of mitosporic Ascomycota, Pezizomycotina, and Saccharomycotina (Ascomycota) and Agaricomycotina, Pucciniomycotina, and Ustilaginomycotina (Basidiomycota).

Overall metabolism and the respective metabolic profiles of prokaryotic and eukaryotic communities in deep-sea sponge *Neamphius huxleyi*. Fig. 1 exhibits the protein categories based on COG and KOG database. There are abundant proteins related to nutrient metabolism of amino acids, carbohydrates, lipids and inorganic ions. In the case of energy metabolism, the symbionts could obtain energy from nitrogen (PATH:ko00910) and sulfur (PATH:ko00920) metabolisms.

Interestingly, the relative abundance of functional genes related to amino acids, carbohydrates, respiration, cold shock, membrane transport, signal transduction and energy metabolisms were different between prokaryotic and eukaryotic communities based on SEED and KEGG analyses (Fig. 2). As depicted in Table 1, prokaryotes and eukaryotes show different gene enrichment (*i.e.* the related gene is mainly involved in some pathways) in some terms of SEED subsystem, *e.g.* amino acids and derivatives, nitrogen metabolism in prokaryotes, while fatty acids, lipids, transposable elements, regulation and cell signaling, disease and defense metabolisms in eukaryotes. In the case of stress response, membrane transport and virulence related genes, prokaryotic and eukaryotic communities also have different gene enrichment trends for detailed metabolic categories (Table 1).

Total nitrogen metabolism and the respective potentials of prokaryotic and eukaryotic communities. Genes involved in nitrogen fixation (*nos*Z), assimilation (*glt*, GDH2), DNRA (*nrf*A), ammonia oxidation (*amo*A) and denitrification (*nar*, *nap*A, *nir*K, *nor*, *nos*Z) were detected in the metagenome indicating a complex nitrogen cycle in the deep-sea sponge (Fig. 3). Particularly, *nrf* gene was detected for the first time in sponges suggesting a DNRA (dissimilatory nitrite reduction to ammonium) process in this deep-sea sponge. There are 25,374 reads of *amo* (ammonia monooxygenase gene), 8,461 reads of *nar* (nitrate reductase gene) and 4,401 reads of *glt* (glutamate synthase gene), suggesting ammonia oxidization and assimilation, and nitrate reduction might be the important processes in the nitrogen cycling of sponge *Neamphius huxleyi*.

MEGAN analysis showed that mainly prokaryotes such as Proteobacteria, Actinomycetales and archaea might probably be involved in the nitrogen cycling (Fig. 4), for example *Nitrosopumilus* and *Cenarchaeum* in ammonia oxidization, *Aeromonas* in nitrogen fixation, *Agrobacterium* and *Sphingomonadaceae* in ammonia assimilation, and *Ochrobacterium* in denitrification. In contrast, very few eukaryotes were involved in the nitrogen cycle based on Fig. 4.
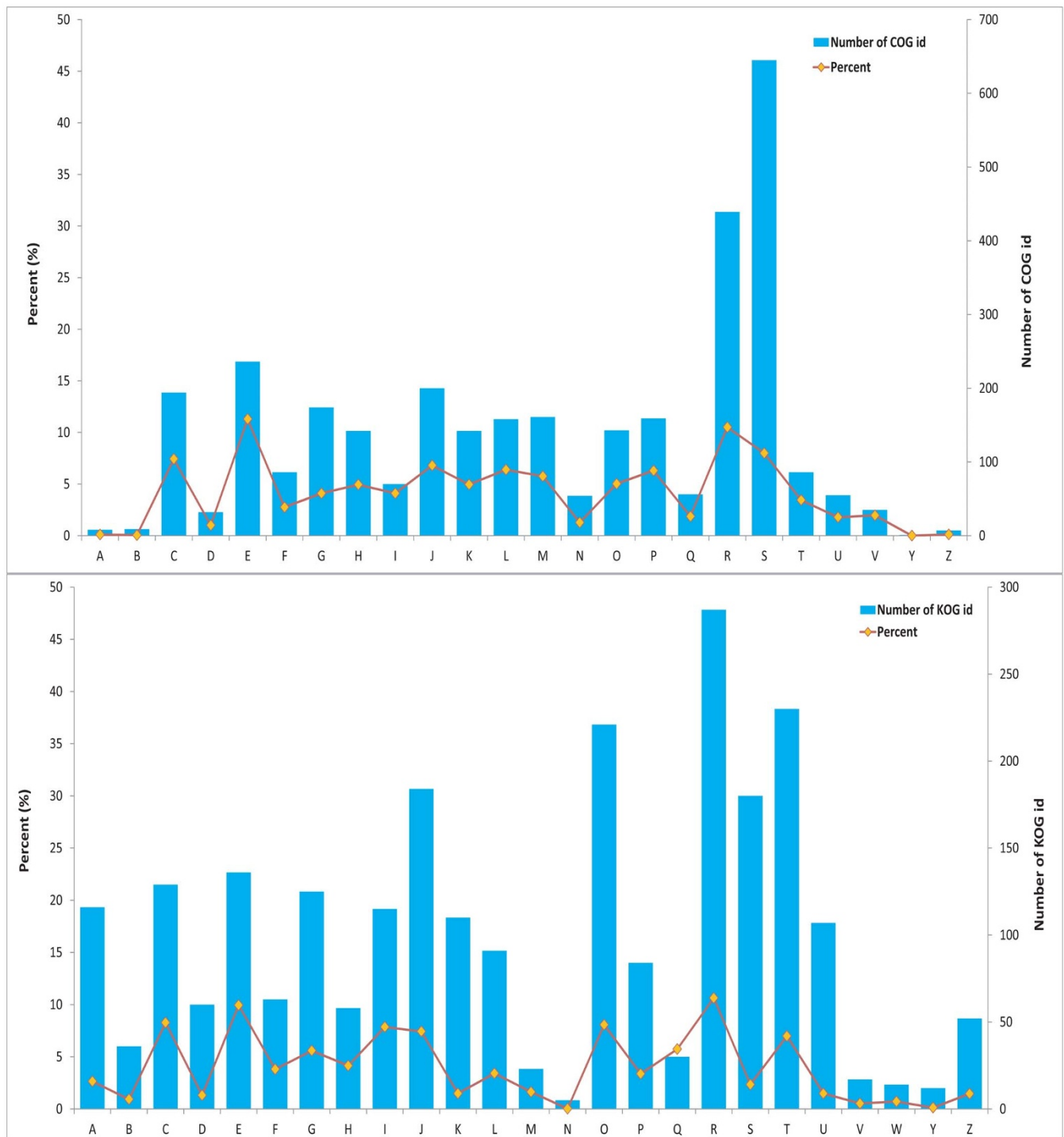
Figure 1 | The illustration showed the types of COG and KOG classification and percentage of each protein category ((a): COG database, prokaryotes; (b): KOG database, Eukaryotes). A: RNA processing and modification; B: chromatin structure and dynamics; C: energy production and conversion; D: cell cycle control, cell division, chromosome partitioning; E: amino acid transport and metabolism; F: nucleotide transport and metabolism; G: carbohydrate transport and metabolism; H: coenzyme transport and metabolism; I: lipid transport and metabolism; J: translation, ribosomal structure and biogenesis; K: transcription; L: replication, recombination and repair; M: cell wall/membrane/envelope biogenesis; N: cell motility; O: posttranslational modification, protein turnover, chaperones; P: inorganic ion transport and metabolism; Q: secondary metabolites biosynthesis, transport and catabolism; R: general function prediction only; S: Function unknown; T: signal transduction mechanisms; U: intracellular trafficking, secretion, and vesicular transport; V: defense mechanisms; W: extracellular structures; Y: nuclear structure; Z: cytoskeleton.

Carbon metabolism and the respective potentials of prokaryotic and eukaryotic communities. The carbon metabolic pathway was constructed according to the metagenome data (Fig. 5; see Supplementary Table S2 online). Accordingly, the symbionts probably use two alternative pathways, the Calvin-Benson cycle and Reverse Krebs cycle (Reductive Citric Acid Cycle) which does not depend directly on the presence of light, for the synthesis of organic matter using CO_2 and water as materials (PATH:ko00720).

The 29 processes of carbon metabolism-linked genes can be found as Supplementary Table S2 online, genes encoding enzymes 3.1.3.11
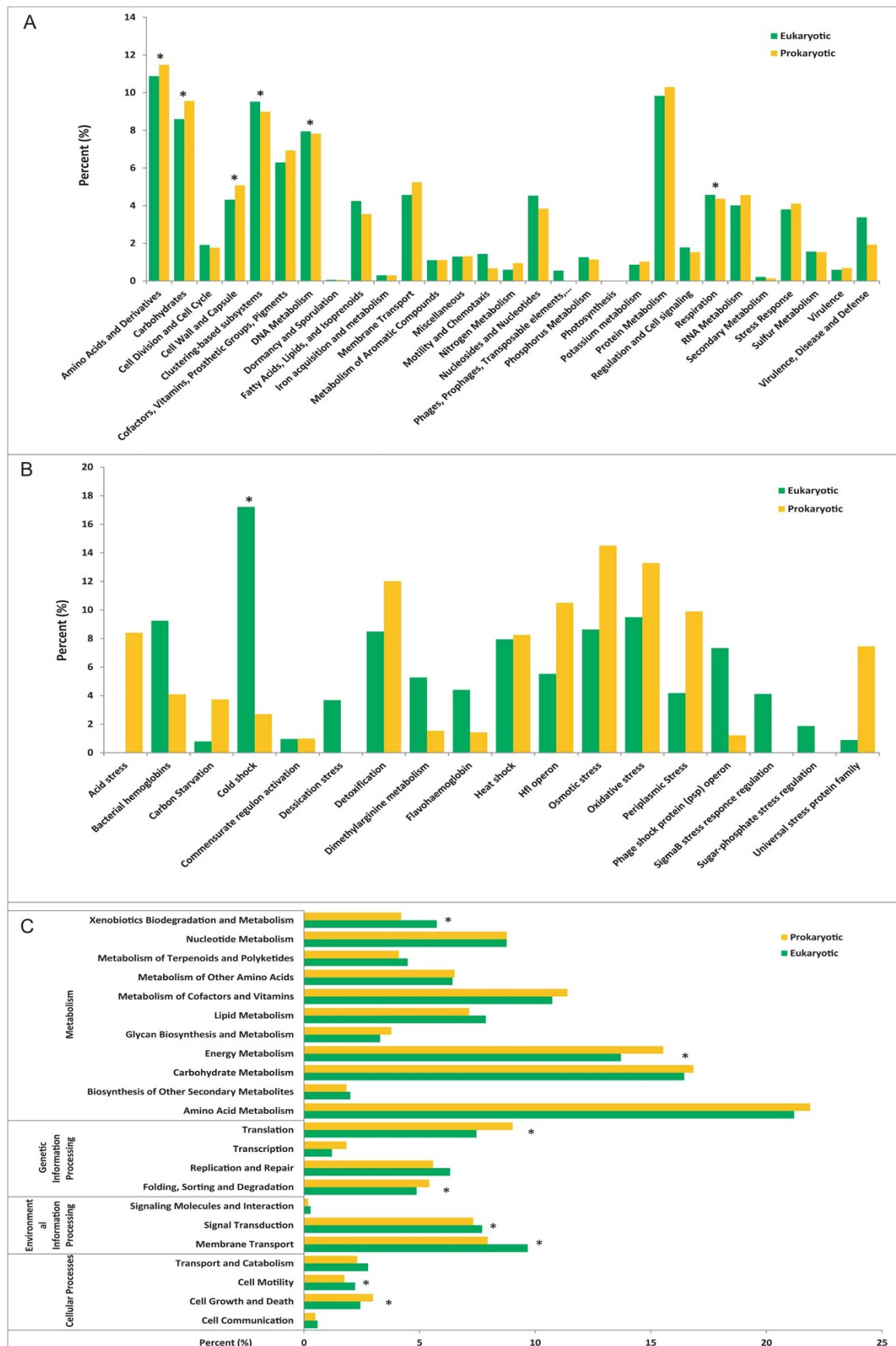
Figure 2 | (A): Genes that matched the major (all) metabolic categories (E-value < 10^{-3}) based on SEED subsystem. (B): Genes that matched stress response (E-value < 10^{-3}) based on SEED subsystem. (C): KEGG analysis of the metabolic difference between eukaryotes and prokaryotes. "*" indicated a statistical significance with a P-value < 0.001 based on the mapped-read numbers in t-test. Yellow: prokaryotic symbionts, Green: eukaryotic symbionts.

**Table 1 | Enrichment analysis of eukaryotic and prokaryotic genes**

	Enriched category	P value (%)	
		Enriched in eukaryotes	Enriched in prokaryotes
Eukaryotic genes	Fatty acids, lipids, and isoprenoids	4.24E-06 (28.4)	
	Motility and chemotaxis	1.17E-36 (40.0)	
	Transposable elements, plasmids	1.58E-99 (73.7)	
	Regulation and cell signaling	4.67E-25 (34.4)	
	Respiration	1.66E-12 (29.3)	
	Disease and defense	1.77E-107 (40.5)	
Prokaryotic genes	Amino acids and derivatives		1.45E-11 (34.6)
	(Arginine; urea cycle, polyamines)		(2.01E-05 (36.7))
	(Histidine metabolism)		(2.63E-32 (48.2))
	(Lysine, threonine, methionine, and cysteine)		(1.57E-07 (36.3))
	Metabolism of aromatic compounds		3.19E-21 (43.6)
	Nitrogen metabolism		4.19E-15 (42.4)
	(Ammonia assimilation)		(1.24E-27 (48.2))
	Potassium metabolism		2.08E-28 (45.8)
Virulence, disease and defense	Phage shock protein (psp) operon	5.38E-09 (80.0)	
	Type III, Type IV, Type VI, ESAT secretion systems	2.21E-25 (45.0)	
	Fimbriae of the chaperone/usher assembly pathway		7.06E-31 (86.8)
Stress response	Cold shock	2.21E-25 (69.3)	
	Osmotic stress		4.98E-21 (44.4)
	Oxidative stress		4.26E-16 (39.6)
	Periplasmic stress		2.93E-13 (50.9)
Membrane transport	Protein secretion system, Type II	2.46E-06 (37.7)	
	Protein translocation across cytoplasmic membrane	1.62E-10 (45.4)	
	Protein and nucleoprotein secretion system, Type IV		5.54E-06 (44.1)
	Uni- Sym- and Antiporters		4.61E-126 (69.2)

Enrichment means the related gene is mainly involved in some pathways. P means the reliability, P < 0.001; ''%'' represents the percentage of different genes in detailed category.

(FBP → F6P), 4.1.1.31 (PEP → oxaloacetate), 4.1.2.13 (FBP → GAP + DHAP), 5.1.3.1 (X5P → RuBP), 5.3.1.6 (R5P → Ru5P) and 6.2.1.5 (succinate → succinyl-CoA), showed different gene enrichment trends between prokaryotic and eukaryotic communities. MEGAN analysis indicated that prokaryotes including bacteria (Actinobacteria, Acidobacteria, Bacteroidetes, and Proteobacteria) and archaea (*Nitrosopumilus* and *Cenarchaeum*) were mainly involved in these six processes (Fig. 6). It is Proteobacteria, e.g. *Sphingopyxis*, *Agrobacterium* and *Brevundimonas*, who mainly participate in the six carbon metabolic processes (see Supplementary Fig. S3 online), whereas only few species eukaryotes, *e.g. Schizophyllum*, *Schizosaccharomyces* and *Tetrahymena*, might be involved.

This finding was supported in depth by the carbon monoxide dehydrogenases analysis. Multiple kinds of aerobic-type carbon monoxide dehydrogenases (*e.g.* COG1319, CoxM; COG1529, CoxL; COG2080, CoxS) involved in carbon fixation were detected in the sponge metagenome. MEGAN analysis suggested that bacteria (Proteobacteria, Actinomycetales and Sphingobacteriales), particularly Alphaproteobacteria, and archaea (*e.g. Picrophilus* of Euryarchaeota) rather than eukaryotes might be involved in the CO_2 fixation of deep-sea sponge *Neamphius huxleyi* (Fig. 7).

Molecular interactions of prokaryotic and eukaryotic symbionts with the sponge host. Genes encoding transposases, ankyrin repeat
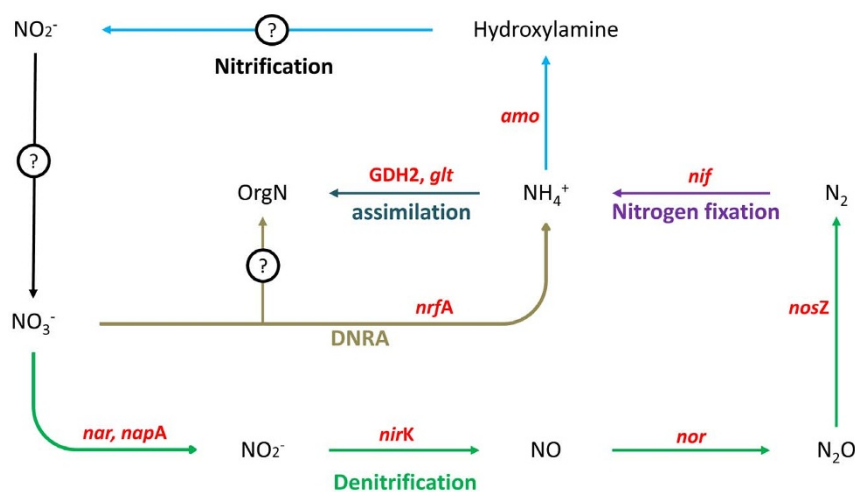
Figure 3 | Nitrogen cycling of the sponge symbionts. GDH: glutamate dehydrogenase (1,698 reads); *amo*: ammonia monooxygenase (25,374 reads); *nar*: nitrate reductase (8,461 reads); *glt*: glutamate synthase (4,401 reads); *nor*: nitric oxide reductase (736 reads); *nap*A: periplasmic dissimilatory nitrate reductase (243 reads); *nir*K: Copper-containing nitrite reductase (151 reads); *nos*Z: nitrous oxide reductase (110 reads); *nif*: nitrogenase (146 reads); *nrf*A: cytochrome c nitrite reductase (7 reads). The question mark means the functional gene is not detected in the metagenome.
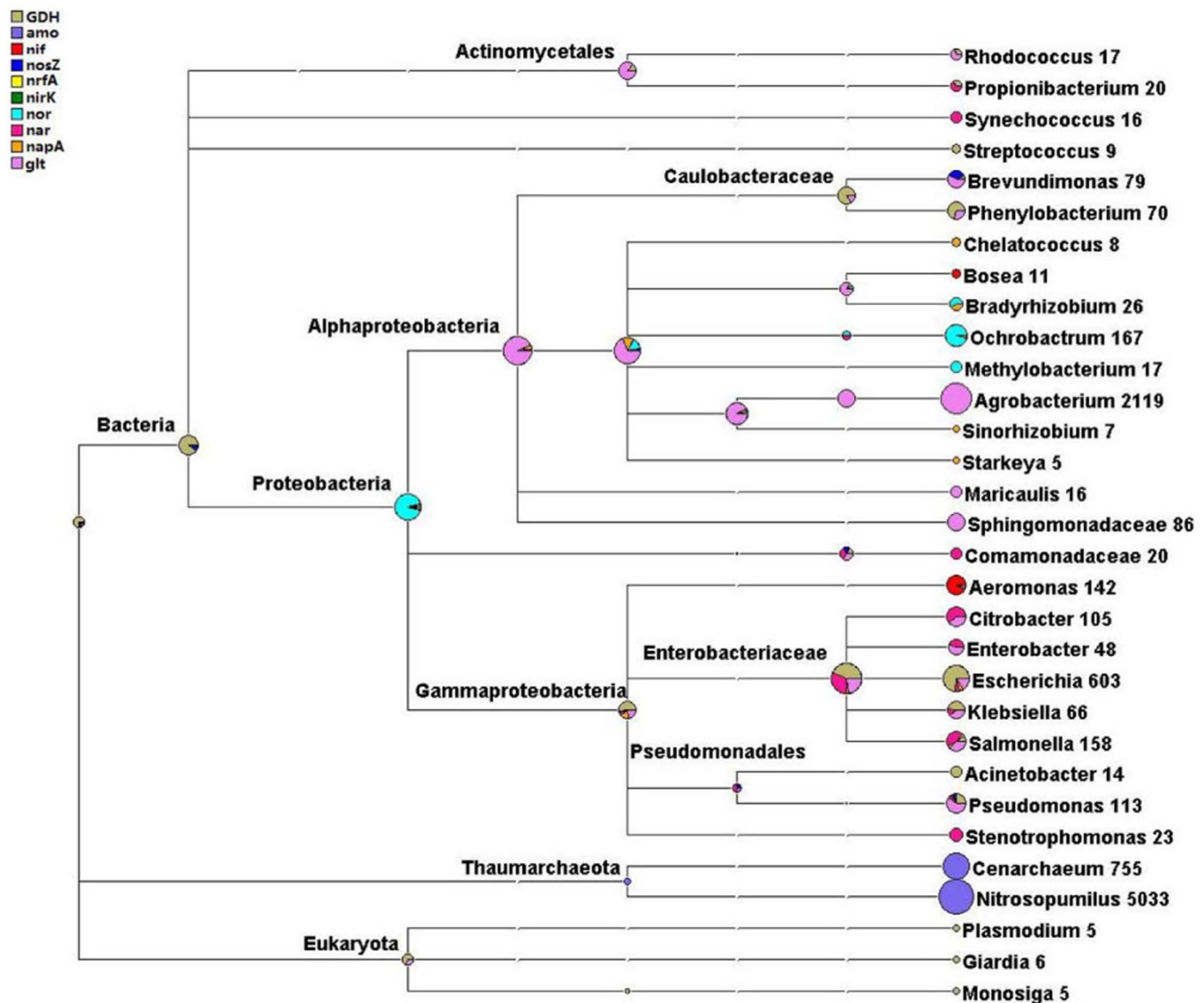
Figure 4 | MEGAN analysis of nitrogen cycling genes assigned to microbes. The size of the circles is relative to the number of genes assigned to each node (also indicated in numbers), and taxonomy is displayed with the lowest level predicted.

proteins (ARPs) and tetratricopeptide repeat proteins (TPRs) were observed in both prokaryotic community (*e.g.* COG5433, COG5659 and COG5421; COG0666; COG0790 and COG0457) and eukaryotic community (*e.g.* KOG4585; KOG0522, KOG2505, KOG0510, KOG-0508, KOG0509 and KOG4177; KOG2471, KOG4555, KOG0553 and KOG1125). The statistical difference in the transposase genes enrichment of eukaryotes from prokaryotes was observed, *e.g.* PF08722, PF02914, PF01609, PF12017, PF05598, PF04986, PF01548, PF02992 and PF07592 (see Supplementary Fig. S4 online). Taking the type IV secretion system as another example (see Supplementary Table S3 online), eukaryotic symbionts may mainly use DNA uptake competence system protein and Cag pathogenicity island (PAI) protein, while prokaryotic symbionts may use the conjugal DNA-protein transfer protein. This suggests the different molecular interactions of prokaryotic and eukaryotic symbionts with their sponge host.

Discussion

The metagenome data show that the deep-sea sponge *Neamphius huxleyi* harbors phylogenetically diverse prokaryotes *e.g.* at least 19 bacterial phyla and 2 archaeal phyla (see Supplementary Fig. S2a online), most of which belong to sponge-associated or specific microbial lineages, *i.e.* Proteobacteria, Actinobacteria,

Bacteroidetes, Planctomycetes, Firmicutes, Acidobacteria and Chloroflexi[5–10]. In contrast, only 4 bacterial phyla Acidobacteria, Actinobacteria, Firmicutes and Proteobacteria were detected in the seawater around the sampling site (see Supplementary Fig. S2c online). Although Proteobacteria was also the most predominant bacteria in deep-sea water, Beta- and Epsilon-proteobacteria which were present in sponge *Neamphius huxleyi* were not detected in the seawater. Crenarchaeote was the second predominant prokaryotic group (22%) in the deep-sea water, however it only account to 2.08% of the microbial community of sponge *Neamphius huxleyi*. All these results showed obvious difference of prokaryotes in sponge *Neamphius huxleyi* from that in the deep-sea water, and suggested the sponge-specific microbial lineages which were different from that of the environment seawater. Thus, the detected microbes in the deep-sea sponge *Neamphius huxleyi* could be considered as the components of sponge microbiome. Besides prokaryotes, phylogenetically diverse fungi and protists including protophyte and protozoa were detected in this deep-sea sponge (see Supplementary Fig. S2d online), therefore the exploration of the physiological functions of eukaryotic symbionts is an important task ahead.

Compared with the findings from the metagenomics analysis of shallow-water sponges[16–19], similar overall metabolism and asso-
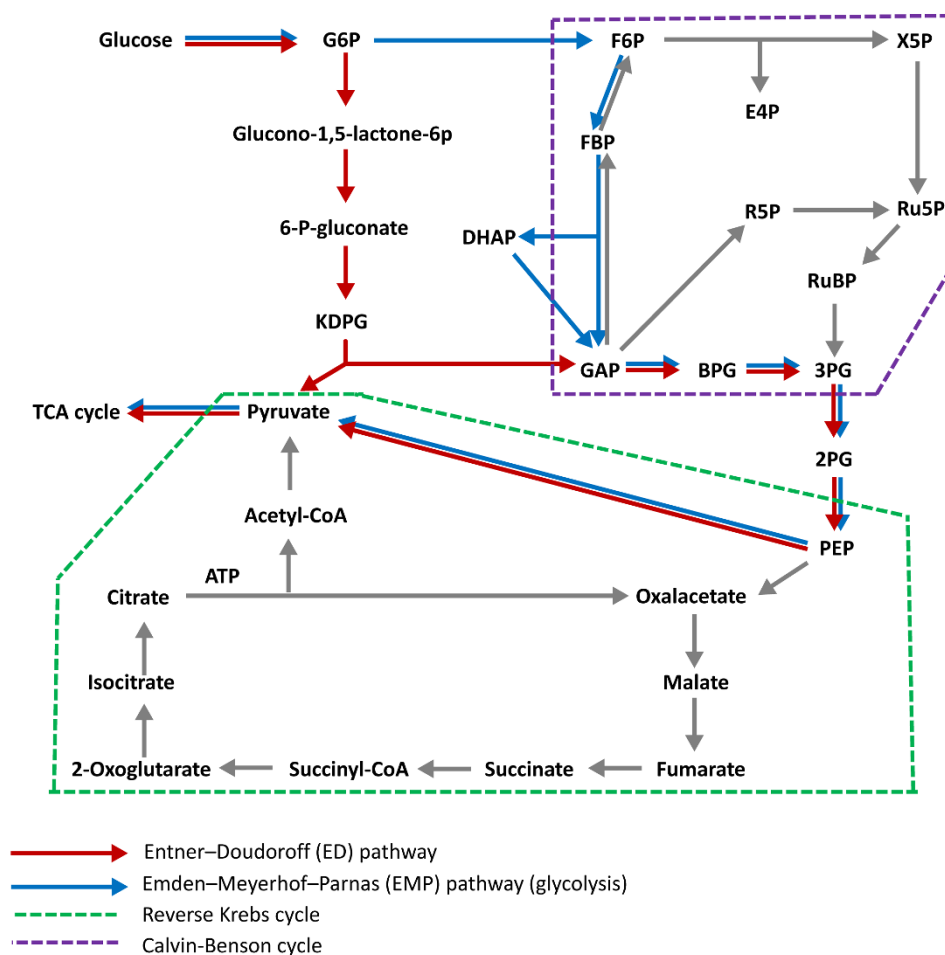
Figure 5 | Carbon metabolic pathway. GAP: glyceraldehyde-3-phosphate; FBP: fructose 1,6-bisphosphate; DHAP: dihydroxyacetone phosphate; KDPG: 2-keto-3-dehydro-6-phosphogluconate; G6P: glucose 6-phosphate; F6P: fructose 6-phosphate; FBP: fructose 1,6-bisphosphate; E4P: erythrose 4-phosphate; X5P: xylulose 5-phosphate; GAP: glyceraldehyde-3-phosphate; Ru5P: ribulose 5-phosphate; S7P: sedoheptulose 7-phosphate; R5P: ribose 5-phosphate; RuBP: ribulose 1,5-bisphosphate; BPG: 1,3-Bisphospho-glycerate; 3PG: 3-phosphoglycerate; 2PG: 2-phosphoglycerate; PEP: phosphoenolpyruvate.

ciation of prokaryotic symbionts with sponge host were suggested for the deep-sea sponge *Neamphius huxleyi*, e.g. providing the sponge host with amino acids and vitamins, ammonia removal and molecular interaction between the sponge host and its symbionts. This supported the hypothesis of functional equivalence and evolutionary convergence in complex communities of sponge microbial symbionts suggested by Fan *et al.* to some extent[18]. Because of the unique characteristics of deep-sea environment from shallow water, some metabolic characteristics of the deep-sea sponge *Neamphius huxleyi* were suggested in this study, particularly for nitrogen and carbon metabolisms.

Nitrogen cycle is a critical component of the biogeochemical cycles in the oceans[31]. Hoffmann *et al.* revealed a complex nitrogen cycle in sponge *Geodia barrette*, however only nitrification, denitrification and anammox processes were included[22]. In the metagenomic investigation of six species of shallow-water sponges by Fan *et al.*[18], only denitrification and ammonia oxidization were analyzed. In this study, nitrogen fixation, assimilation, DNRA, ammonia oxidation and complete denitrification were suggested in the deep-sea sponge *Neamphius huxleyi* (Fig. 3). Based on Fig. 4, in this deep-sea sponge, ammonia oxidizing archaea (AOA), *Nitrosopumilus* and *Cenarchaeum* of Thaumarchaeota, are probably mainly involved in the ammonia oxidation. Siegl *et al.* suggested the assimilation of NH_3 by Poribacteria through the single-cell genomics[15]. As shown in Fig. 4, Proteobacteria, especially Alphaproteobacteria e.g.

Agrobacterium, are probably involved in the ammonia assimilation of this deep-sea sponge. The vigorous denitrification, particular nitrate and nitrite reduction, in this deep-sea sponge is probably because the higher nitrate and nitrite concentration in the deep-sea water, for example in the sampling site, the concentration of nitrate increases from 0 at the water surface to 20 µM/L at approx 900 m water depth (see Supplementary Fig. S1 online). Meanwhile, the low oxygen condition in the deep sea favors the denitrification process. Besides the nitrogen form the seawater, nitrogenous waste secreted by the host could be metabolized and recycled by the sponge microbial symbionts as earthworms[32].

For deep-sea invertebrates including sponges, carbon source maybe a limiting factor, which could be supported by the carbon starvation stress-related proteins detected in the metagenome (Fig. 2). Based on carbon monoxide dehydrogenase and $NaHC^{14}O_3$ utilization, CO_2 fixation in marine invertebrates has been early suggested[15,33]. Many deep-sea invertebrates live in association with microbes which ensure a light-independent type of primary production named 'chemoautotrophy'[34]. CO_2 fixation by chemoautotrophic prokaryotes may be a common phenomenon in deep-sea symbioses[34,35]. The symbionts of deep-sea sponges cannot use photosynthesis for producing organic carbon molecules, instead, which might be provided by the CO_2 fixation of chemoautotrophic prokaryotes such as Proteobacteria, Actinomycetales and Sphingobacteriales in sponges (Fig. 7).
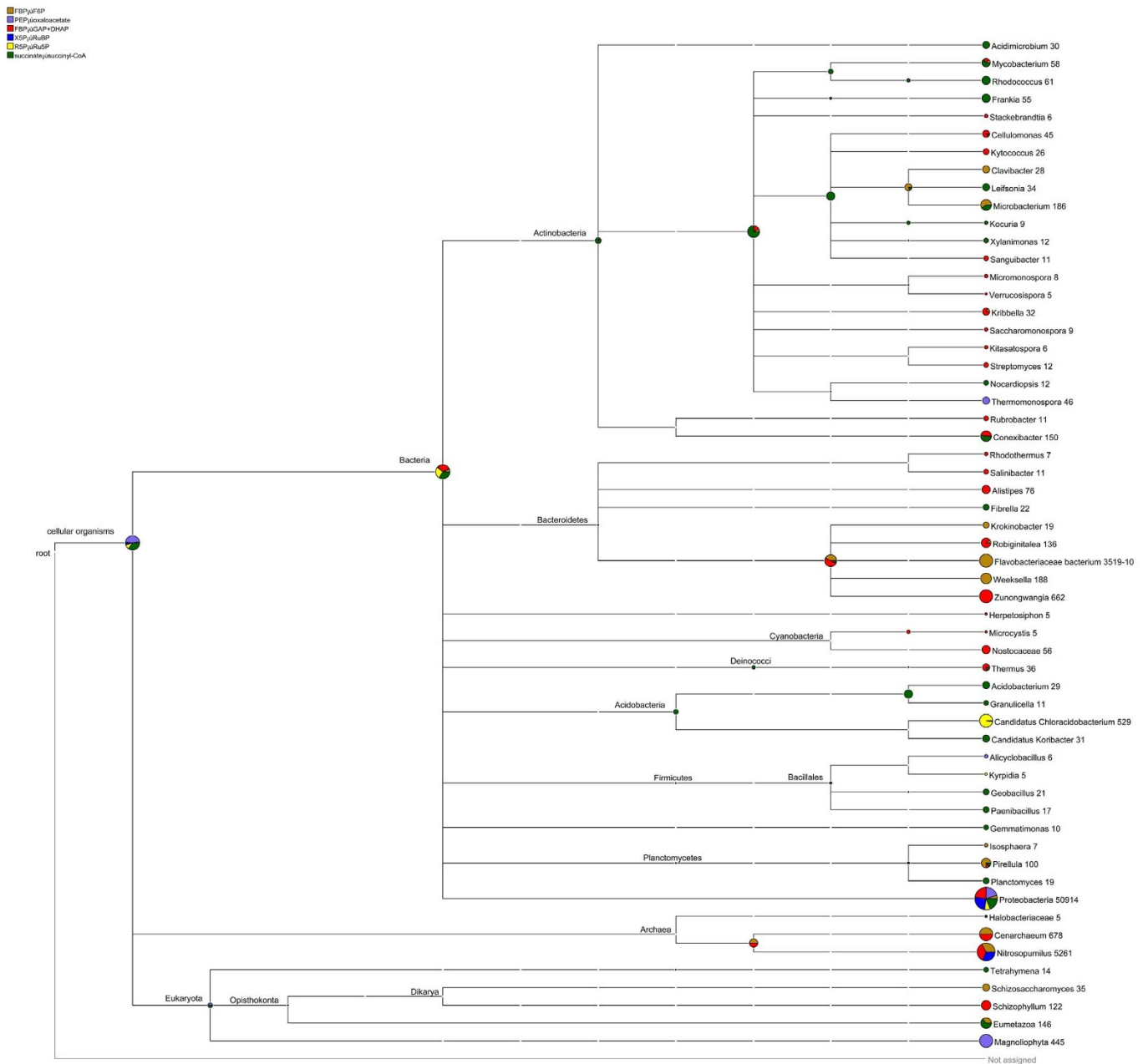
Figure 6 | MEGAN analysis of carbon metabolism-related microbes. 3.1.3.11: FBP → F6P; 4.1.1.31: PEP → oxaloacetate; 4.1.2.13: FBP → GAP + DHAP; 5.1.3.1: X5P → RuBP; 5.3.1.6: R5P → Ru5P; 6.2.1.5: succinate → succinyl-CoA. For Proteobacteria involved in the carbon metabolism please see Supplementary Fig. S3 online.

The existing sponge omics investigations were only focused on the prokaryotic symbionts of sponges[16–19], in this study, the community structure and function of both prokaryotic and eukaryotic symbionts in deep-sea sponge *Neamphius huxleyi* were revealed. Meanwhile, by the comparison of metabolic potentials, different contributions of prokaryotic and eukaryotic symbionts of sponge *Neamphius huxleyi* as well as their associations with the same sponge host were highlighted. For example, MEGAN analysis showed that prokaryotes rather than eukaryotes are mainly involved in the nitrogen cycling (*e.g.* Proteobacteria, Thaumarchaeota and Actinomycetales), and CO_2 fixation (*e.g.* Proteobacteria, Bacteroidetes and Acidobacteria) (Fig. 3, 4, 7; Table 1). On the other hand, metagenomic analysis suggested the role of eukaryotes cannot be ignored, for example, protozoa *Giardia*, *Monosiga* and *Plasmodium* may participate in the ammonia assimilation process (Fig. 4), fungi *Schizophyllum*

and *Schizosaccharomyces* may be involved in some processes of CO_2 fixation (*e.g.* 4.1.2.13;3.1.3.11) (Fig. 7).

It is known that different Ankyrin repeat proteins (ARPs) and tetratricopeptide repeat proteins (TPRs) are often found in facultative or obligate symbionts and could modulate the host's behavior[36]. The observed abundant ARPs and TPRs in the prokaryotes and eukaryotes in the deep-sea sponge *Neamphius huxleyi* indicated the possible molecular communication between prokaryotes/eukaryotes and their sponge host, which was similar to the finding from shallow water sponges[17,18]. But, different interaction strategies between prokaryotic and eukaryotic symbionts were indicated by the metagenomic analysis. For instance the different enrichment of genes related to type IV secretion system (see Supplementary Table S3 online) indicated different strategies of prokaryotes and eukaryotes in delivering ARPs and TPRs into the cells of sponge host.
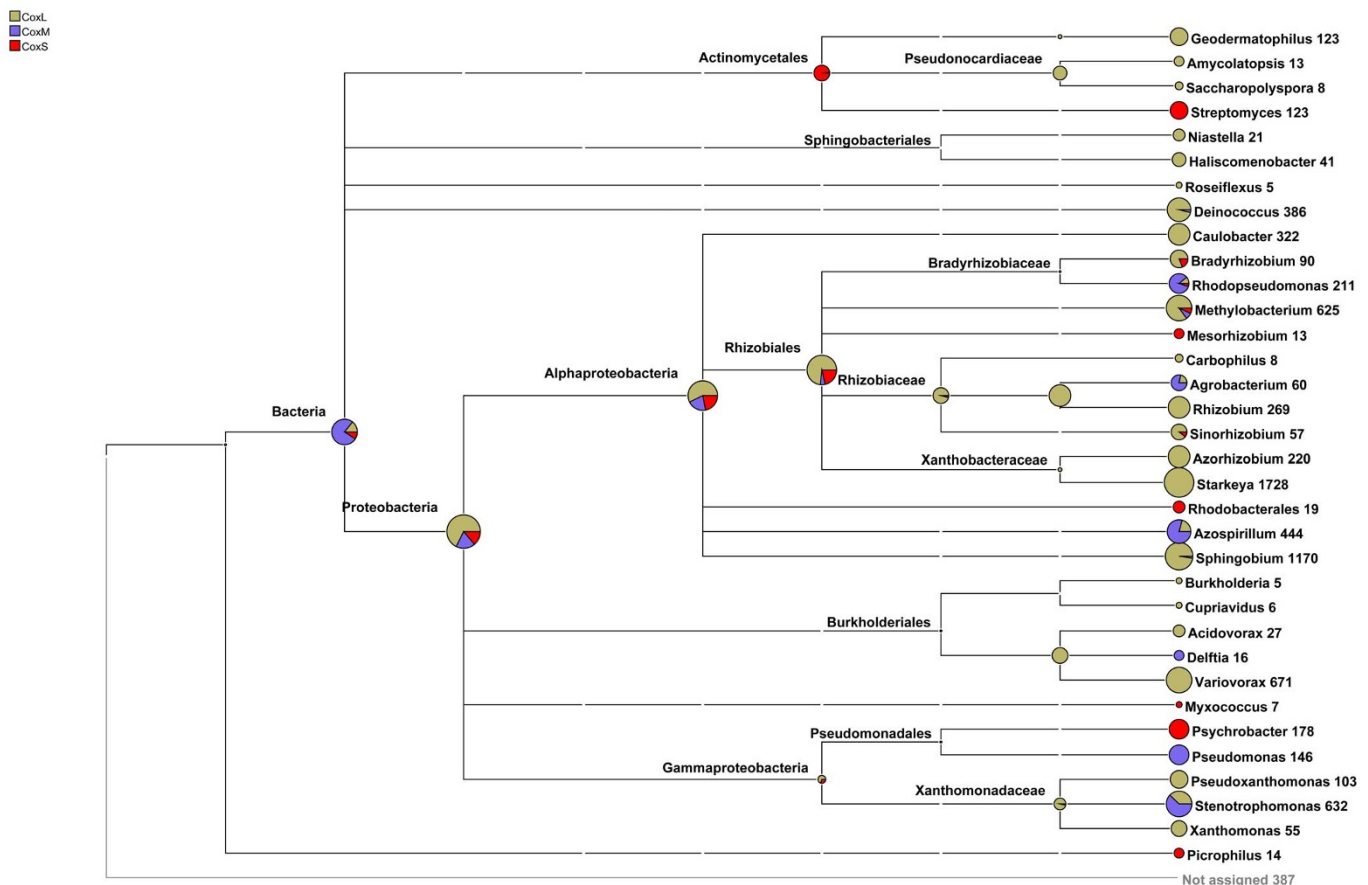
Figure 7 | MEGAN analysis of monoxide dehydrogenase. CoxM: aerobic-type carbon monoxide dehydrogenase, middle subunit CoxM/CutM homologs; CoxL: aerobic-type carbon monoxide dehydrogenase, large subunit CoxL/CutL homologs; CoxS: aerobic-type carbon monoxide dehydrogenase, small subunit CoxS/CutS homologs. The size of the circle is relative to the number of genes assigned to each node (also indicated in numbers), and taxonomy is displayed with the lowest level predicted.

Meanwhile, it is known that transcriptases play an important role in genetic exchange and rearrangement, and consequently facilitate the evolutionary adaptation of microbial populations to specific niches[37]. Like shallow water sponges[18], genes encoding transcriptases were detected in this deep-sea sponge. But, the different abundance of transcriptase genes between prokaryotic and eukaryotic communities (see Supplementary Fig. S4 online) indicated the prokaryotes and eukaryotes might have their own distinct set of transcriptase systems, and therefore have different coevolution strategies with the sponge host and other community members.

In summary, prokaryotic and eukaryotic symbionts in the deep-sea sponge *Neamphius huxleyi* show different metabolic potentials especially in nitrogen and carbon metabolisms and molecular interactions with their sponge host according to the MEGAN and gene enrichment analyses. Based on this study, it could be hypothesized that prokaryotic and eukaryotic symbionts have different ecological roles and associations with the same sponge host. In addition, though similar overall metabolism as that of shallow water sponges was suggested, some unique metabolic profiles of deep-sea sponge were indicated, for example vigorous denitrification, CO_2 fixation by chemoautotrophic prokaryotes, indicating the adaptation of sponge microbial symbionts to the deep-sea environment besides to the sponge host. In future, more species of sponges should be investigated at the whole community level including prokaryotes and eukaryotes for the biological and ecological evaluation of the sponge holobionts. Particularly, investigations using metatranscriptomics or metaproteomics will contribute significantly to our understanding of the whole community functions of the deep-sea sponge holobionts,

as well as the relationships of symbionts with the sponge host and deep-sea environment.

Methods

Sponge and seawater sampling. Sponge samples were collected at a depth of *ca*. 1,800 m (temp. *ca*.4°C, salinity *ca*.34.3‰) from the Indian Ocean (36°80′N, 52°76′E) adjacent to the Southwest Indian Ocean Ridge (see Supplementary Fig. S1 online), during the 2010 scientific investigation on the Ocean No 1. Research Ship, China. Samples were transferred directly to Zip-lock bags containing seawater to prevent the contact of sponge tissue with the air, and then stored at −70°C immediately before use. The sponge was identified as *Neamphius huxleyi* with a similarity of 99% according to 28S rRNA gene sequence amplified with the primer set NL4F/NL4R[38]. The metagenome data were deposited in the GenBank/EMBL/DDBJ database under accession number SRA052801. Seawater samples were collected using 10L-Niskin bottles mounted on a conductivity-temperature-depth-oxygen (CTDO) rosette sampler. Concentrations of nitrate and nitrite in sea water samples were determined on board by ion chromatographyspectrophotometry[39]. After membrane (0.2 μm) filtration, the diversity of prokaryotes in the seawater around the sampling site (*ca*. 1800 m) was analyzed by 16S rRNA gene library construction using primer sets 27F and 1492R for bacteria[40], 21F and 958R for archaea[41].

DNA extraction and deep sequencing. Three samples of *Neamphius huxleyi* were used for investigation after washing three times with artificial seawater (ASW) (1.1 g $CaCl_2$, 10.2 g $MgCl_2\cdot 6H_2O$, 31.6 g NaCl, 0.75 g KCl, 1.0 g Na_2SO_4, 2.4 g Tris-HCl, 0.02 g $NaHCO_3$, 1 L distilled water, pH 7.6) to eliminate the microbes loosely attached on the sponge surface and canals. Two strategies were used to extract sponge metagenomic DNA: (1) QIAGEN DNeasy® Kit (Qiagen) following the manufacturer's protocol; (2) CTAB (Cetyltrimethyl Ammonium Bromide)-based method according to Taylor *et al*.[42]. DNA samples extracted by different methods were pooled and mixed thoroughly before deep sequencing.

Deep sequencing was carried out on Genome Analyzer IIx system and Highseq 2000 of Illumina Company using paired-end technology (2 × 120, 2 × 100). Total metagenomic DNA was broken up into 300 bp fragments by Covaris and extracted using QIA quick PCR purification kit (QIAGEN, part # 28104). Adaptors were ligated

9

to the extracted 300 bp fragments. Then, 300 bp fragments were enriched using Phusion DNA Polymerase through low cycle PCR under default conditions with 5 cycles. Cluster generation and sequencing were performed according to the manufacturer's manual.

Phylogenetic affiliation of sponge symbionts. The sequences containing the ambiguous base (N) were removed using customized perl scripts. Each sequence was searched against Greengenes reference 16S rRNA database and SILVA Eukaryota database (version 104) using MEGABLAST[43–46] for identifying 16S rRNA and 18S rRNA genes, respectively. 16S rRNA gene and 18S rRNA gene references were used as templates, all gaps of 16S rRNA (V3 and V6) and 18S rRNA (V4) alignments were treated as single evolutionary events for calculating distance. For prokaryotic symbionts, the online software RDP classifier was used to assign sequences to phylogenetic taxonomy based on the Ribosomal Database Project under the condition of confidence threshold 50%[47,48]. In the case of eukaryotic symbionts, BLASTN was used to find the best hit against 18S rRNA genes from SILVA database (version 104) following parameters: 97% identity and 90% coverage[43].

Analysis of functional gene and metabolic pathway. These reads, which had ambiguous base, or adapter contamination, or were shorter than 50 bp, were abandoned. The sponge sequences (txid6040) from NCBI were contained as a reference for removing possible sponge sequences from metagenome data by MEGABLAST with E-value 1e-10 and "-W 7". After removing the ribosomal RNA which were identified by comparing with Greengenes reference 16S rRNA database and SILVA Eukaryota database, the qualified reads were assembled using Velvet[49], with the parameters of hash 35 and minimum length of 500 bp. Open reading frames (ORFs) were identified using two programs, MetaGeneAnnotator for prokaryotic gene prediction and GeneMark for eukaryotic gene prediction[50,51], with ORF length set as minimum 60 bp for prokaryote gene and 150 bp for eukaryotic gene, respectively. Based on the deduced amino acid sequences, the annotation was performed through BLASTP against SEED database[52], with parameters set at E-value 1e-3. Protein KOG (Eukaryotic Orthologous Groups) and COG (Clusters of Orthologous Groups) assignments were predicted through RPS-BLAST with the Conserved Domain Database (CDD) with E-value 1e-3[53]. The metabolic pathway was constructed based on the KEGG database by SBH (single-directional best hit) method[54]. For understanding the composition of symbiotic community which contributed to certain functions, sequences were compared with non-redundant NCBI nucleotide database using BLASTN and analyzed with MEGAN 4.69[55], which uses a last common ancestor (LCA) algorithm to assign sequences to the NCBI taxonomy.

Enrichment analysis of prokaryotic and eukaryotic genes. The reads number of each gene was firstly calculated using Bowtie and SAMtools[56,57]. The genes number was quantified by reads Per Kilo bases per Million reads (RPKM)[58], and the RPKM measure of reads density reflected the molar concentration of a gene by normalizing for gene length and for the total read number in the measurement. The reads were assumed to be belongs of prokaryotic or eukaryotic genes if they were mapped to predicted prokaryotic or eukaryotic genes. Then, differently genes were identified by DEGseq package using the method of MARS (MA-plot-based method with Random Sampling model)[59]. "FDR ≤ 0.001 and the absolute value of log2Ratio ≥ 1" were used as the threshold to judge the significance of gene abundance difference. Abundance analysis of PFAM and enzymes was performed at the parameter of p-value < 0.001. T-test was used to test the statistical significance between prokaryotic and eukaryotic genes based on SEED hierarchical categories at p < 0.001. Enrichment of SEED subsystem (three hierarchy systems) and KEGG pathways for a given gene list was calculated by a classical hypergeometric distribution statistical comparison using R (http://www.r-project.org/), which was based on the relative proportions of the entire SEED or KEGG hierarchical categories.

1. Bernhard, J. M., Buck, K. R., Farmer, M. A. & Bowser, S. S. The Santa Barbara Basin is a symbiosis oasis. *Nature* **403**, 77–79 (2000).
2. Duperron, S., Lorion, J., Samadi, S., Gros, O. & Gaill, F. Symbioses between deep-sea mussels (Mytilidae: Bathymodiolinae) and chemosynthetic bacteria: diversity, function and evolution. *C. R. Biol.* **332**, 298–310 (2009).
3. Srivastava, M. et al. The *Amphimedon queenslandica* genome and the evolution of animal complexity. *Nature* **466**, 720–724 (2010).
4. Gao, Z., Li, B., Zheng, C. & Wang, G. Molecular detection of fungal communities in the Hawaiian marine sponges *Suberites zeteki* and *Mycale armata*. *Appl. Environ. Microbiol.* **74**, 6091–6101 (2008).
5. Lee, O. O. et al. Pyrosequencing reveals highly diverse and species-specific microbial communities in sponges from the Red Sea. *ISME J.* **5**, 650–664 (2011).
6. Schmitt, S. et al. Assessing the complex sponge microbiota: core, variable and species-specific bacterial communities in marine sponges. *ISME J.* **6**, 564–574 (2012).
7. Simister, R. L., Deines, P., Botté, E. S., Webster, N. S. & Taylor, M. W. Sponge-specific clusters revisited: a comprehensive phylogeny of sponge-associated microorganisms. *Environ. Microbiol.* **14**, 517–24 (2012).
8. Taylor, M. W., Radax, R., Steger, D. & Wagner, M. Sponge-associated microorganisms: Evolution, ecology, and biotechnological potential. *Microbiol. Mol. Biol. Rev.* **71**, 295–347 (2007).
9. Webster, N. S. et al. Deep sequencing reveals exceptional diversity and modes of transmission for bacterial sponge symbionts. *Environ. Microbiol.* **12**, 2070–2082 (2010).
10. Webster, N. S. & Taylor, M. W. Marine sponges and their microbial symbionts: love and other relationships. *Environ. Microbiol.* **14**, 335–346 (2012).
11. Vogel, G. The inner lives of sponges. *Science* **320**, 1028–1030 (2008).
12. Hallam, S. J. et al. Genomic analysis of the uncultivated marine crenarchaeote *Cenarchaeum symbiosum*. *Proc. Nat. Acad. Sci. U S A* **103**, 18296–18301 (2006).
13. Hentschel, U., Piel, J., Degan, S. M. & Taylor, M. W. Genomic insights into the marine sponge microbiome. *Nat. Rev. Microbiol.* **10**, 641–654 (2012).
14. Kamke, J. et al. Single-cell genomics reveals complex carbohydrate degradation patterns in poribacterial symbionts of marine sponges. *ISME J.* **7**, 2287–2300 (2013).
15. Siegl, A. et al. Single-cell genomics reveals the lifestyle of *Poribacteria*, a candidate phylum symbiotically associated with marine sponges. *ISME J.* **5**, 61–70 (2011).
16. Thomas, T. et al. Functional genomic signatures of sponge bacteria reveal unique and shared features of symbiosis. *ISME J.* **4**, 1557–1567 (2010).
17. Liu, M., Fan, L., Zhong, L., Kjelleberg, S. & Thomas, T. Metaproteogenomic analysis of a community of sponge symbionts. *ISME J.* **6**, 1515–1525 (2012).
18. Fan, L. et al. Functional equivalence and evolutionary convergence in complex communities of microbial sponge symbionts. *Proc. Natl. Acad. Sci. U S A* **109**, E1878–E1887 (2012).
19. Trindade-Silva, A. E. et al. Taxonomic and functional microbial signatures of the endemic marine sponge *Arenosclera brasiliensis*. *PLoS ONE* **7**, e39905 (2012).
20. Garson, M. J., Flowers, A. E., Webb, R. I., Charan, R. D. & McCaffrey, E. J. A sponge/dinoflagellate association in the haplosclerid sponge *Haliclona* sp.: cellular origin of cytotoxic alkaloids by Percoll density gradient fractionation. *Cell Tissue Res.* **293**, 365–373 (1998).
21. Reisser, W. The taxonomy of green algae endosymbiotic in ciliates and a sponge. *Brit. Phycol. J.* **19**, 309–318 (1984).
22. Hoffmann, F. et al. Complex nitrogen cycling in the sponge *Geodia barretti*. *Environ. Microbiol.* **11**, 2228–2243 (2009).
23. Mohamed, N. M., Colman, A. S., Tal, Y. & Hill, R. T. Diversity and expression of nitrogen fixation genes in bacterial symbionts of marine sponges. *Environ. Microbiol.* **10**, 2910–2921 (2008).
24. Caron, D. A., Countway, P. D., Jones, A. C., Kim, D. Y. & Schnetzer, A. Marine protistan diversity. *Annu. Rev. Mar. Sci.* **4**, 467–93 (2012).
25. Sauvadet, A. L., Gobet, A. & Guillou, L. Comparative analysis between protist communities from the deep-sea pelagic ecosystem and specific deep hydrothermal habitats. *Environ. Microbiol.* **12**, 2946–2964 (2010).
26. Meyer, B. & Kuever, J. Phylogenetic diversity and spatial distribution of the microbial community associated with the Caribbean deep-water sponge *Polymastia cf. corticata* by 16S rRNA, *apr*A, and *amo*A gene analysis. *Microb. Ecol.* **56**, 306–321 (2008).
27. Radax, R. et al. Metatranscriptomics of the marine sponge *Geodia barretti*: tackling phylogeny and function of its microbial community. *Environ. Microbiol.* **14**, 1308–1324 (2012).
28. Dumdei, E. J., Blunt, J. W., Munro, M. H. & Pannell, L. K. Isolation of calyculins, calyculinamides, and swinholide H from the New Zealand deep-water marine sponge *Lamellomorpha strongylata*. *J. Org. Chem.* **62**, 2636–2639 (1997).
29. Han, M., Li, Z. & Zhang, F. The ammonia oxidizing and denitrifying prokaryotes associated with sponges from different sea areas. *Microb. Ecol.* **66**, 427–436 (2013).
30. Smedile, F. et al. Metagenomic analysis of hadopelagic microbial assemblages thriving at the deepest part of Mediterranean Sea, Matapan-Vavilov Deep. *Environ. Microbiol.* **15**, 167–182 (2013).
31. Zehr, J. P. & Kudela, R. M. Nitrogen cycle of the open ocean: from genes to ecosystems. *Ann. Rev. Mar. Sci.* **3**, 197–225 (2011).
32. Schramm, A. et al. Acidovorax-like symbionts in the nephridia of earthworms. *Environ. Microbiol.* **5**, 804–809 (2003).
33. van Duyl, F. C., Hegeman, J., Hoogstraten, A. & Maier, C. Dissolved carbon fixation by sponge–microbe consortia of deep water coral mounds in the northeastern Atlantic Ocean. *Mar. Ecol. Prog. Ser.* **358**, 137–150 (2008).
34. Horst, F. CO_2 fixation in the hydrothermal vent tube worm *Riftia pachyptila* (Jones). *Physiol. Zool.* **58**, 272–281 (1985).
35. Hügler, M. & Sievert, S. M. Beyond the Calvin cycle: Autotrophic carbon fixation in the ocean. *Annu. Rev. Mar. Sci.* **3**, 261–89 (2011).
36. Pan, X. et al. Ankyrin repeat proteins comprise a diverse family of bacterial type IV effectors. *Science* **320**, 1651–1654 (2008).
37. Moliner, C., Fournier, P. E. & Raoult, D. Genome analysis of microorganisms living in amoebae reveals a melting pot of evolution. *FEMS Microbiol Rev.* **34**, 281–294 (2010).
38. Nichols, S. A. An evaluation of support for order-level monophyly and interrelationships within the class Demospongiae using partial data from the large subunit rDNA and cytochrome oxidase subunit I. *Mol. Phylogenet. Evol.* **34**, 81–96 (2005).
39. Kim, S. W., Miyahara, M., Fushinobu, S., Wakagi,T. & Shoun, H. Nitrous oxide emission from nitrifying activated sludge dependent on denitrification by ammonia-oxidizing bacteria. *Bioresour. Technol.* **101**, 3958–3963 (2010).
40. O'Sullivan, L. A. et al. Distribution and culturability of the uncultivated 'AGG58 cluster' of the Bacteroidetes phylum in aquatic environments. *FEMS Microbiol. Ecol.* **47**, 359–370 (2004).

41. DeLong, E. F. Archaea in coastal marine environments. *Proc. Natl. Acad. Sci. USA* **89**, 5685–5689 (1992).

42. Taylor, M. W., Schupp, P. J., Dahllöf, I., Kjelleberg, S. & Steinberg, P. D. Host specificity in marine sponge-associated bacteria, and potential implications for marine microbial diversity. *Environ. Microbiol.* **6**, 121–130 (2004).

43. Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. Basic local alignment search tool. *J. Mol. Biol.* **215**, 403–410 (1990).

44. DeSantis, T. Z. *et al*. Greengenes, a chimera-checked 16S rRNA gene database and workbench compatible with ARB. *Appl. Environ. Microbiol.* **72**, 5069–5072 (2006).

45. Pruesse, E. *et al*. SILVA: a comprehensive online resource for quality checked and aligned ribosomal RNA sequence data compatible with ARB. *Nucl. Acids Res.* **35**, 7188–7196 (2007).

46. Schloss, P. D. & Handelsman, J. Toward a census of bacteria in soil. *PLoS Comput. Biol.* **2**, e92 (2006).

47. Cole, J. R. *et al*. The ribosomal database project: improved alignments and new tools for rRNA analysis. *Nucl. Acids Res.* **37**, 141–145 (2009).

48. Wang, Q., Garrity, G. M., Tiedje, J. M. & Cole, J. Naive Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. *Appl. Environ. Microbiol.* **73**, 5261–5267 (2007).

49. Zerbino, D. R. & Birney, E. Velvet: Algorithms for de novo short read assembly using de Bruijn graphs. *Genome Res.* **18**, 821–829 (2008).

50. Lomsadze, A., Ter-Hovhannisyan, V., Chernoff, Y. & Borodovsky, M. Gene identification in novel eukaryotic genomes by self-training algorithm. *Nucl. Acids Res.* **33**, 6494–6506 (2005).

51. Noguchi, H., Park, J. & Takagi, T. MetaGene: prokaryotic gene finding from environmental genome shotgun sequences. *Nucl. Acids Res.* **34**, 5623–5630 (2006).

52. Overbeek, R. *et al*. The subsystems approach to genome annotation and its use in the project to annotate 1000 genomes. *Nucl. Acids Res.* **33**, 5691–5702 (2005).

53. Marchler-Bauer, A. *et al*. CDD: a conserved domain database for protein classification. *Nucl. Acids Res.* **33**, 192 (2005).

54. Kanehisa, M. & Goto, S. KEGG: kyoto encyclopedia of genes and genomes. *Nucl. Acids Res.* **28**, 27–30 (2000).

55. Huson, D. H., Mitra, S., Weber, N., Ruscheweyh, H. & Schuster, S. C. Integrative analysis of environmental sequences using MEGAN4. *Genome Res.* **21**, 1552–1560 (2011).

56. Langmead, B., Trapnell, C., Pop, M. & Salzberg, S. L. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol.* **10**, R25 (2009).

57. Li, H. *et al*. The sequence alignment/map format and SAM tools. *Bioinformatics* **25**, 2078–2079 (2009).

58. Mortazavi, A., Williams, B. A., McCue, K., Schaeffer, L. & Wold, B. Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nat. Meth.* **5**, 621–628 (2008).

59. Wang, L., Feng, Z., Wang, X. & Zhang, X. DEGseq: an R package for identifying differentially expressed genes from RNA-seq data. *Bioinformatics* **26**, 136–138 (2010).

Acknowledgments

Author contributions

Additional information

Supplementary information accompanies this paper at http://www.nature.com/scientificreports

Competing financial interests: The authors declare no competing financial interests.

How to cite this article: Li, Z.-Y., Wang, Y.-Z., He, L.-M. & Zheng, H.-J. Metabolic profiles of prokaryotic and eukaryotic communities in deep-sea sponge *Neamphius huxleyi* indicated by metagenomics. *Sci. Rep.* **4**, 3895; DOI:10.1038/srep03895 (2014).

CORRIGENDUM: Metabolic profiles of prokaryotic and eukaryotic communities in deep-sea sponge *Neamphius huxleyi* indicated by metagenomics

Zhi-Yong Li, Yue-Zhu Wang, Li-Ming He & Hua-Jun Zheng

The original version of this Article contained an error in the naming of the species '*Neamphius huxleyi*', which was incorrectly given as '*Lamellomorpha* sp'. This has now been corrected in both the HTML and PDF versions of the Article.