The ISME Journal (2011) 5, 650–664

www.nature.com/ismej

ORIGINAL ARTICLE

Pyrosequencing reveals highly diverse and species-specific microbial communities in sponges from the Red Sea

On On Lee[1], Yong Wang[1,2], Jiangke Yang[1], Feras F Lafi[2], Abdulaziz Al-Suwailem[2] and Pei-Yuan Qian[1]

[1]*KAUST Global Partnership Program, Department of Biology, The Hong Kong University of Science and Technology, Clear Water Bay, Hong Kong and* [2]*The King Abdullah University of Science and Technology, Thuwal, Saudi Arabia*

Marine sponges are associated with a remarkable array of microorganisms. Using a tag pyrosequencing technology, this study was the first to investigate in depth the microbial communities associated with three Red Sea sponges, *Hyrtios erectus*, *Stylissa carteri* and *Xestospongia testudinaria*. We revealed highly diverse sponge-associated bacterial communities with up to 1000 microbial operational taxonomic units (OTUs) and richness estimates of up to 2000 species. Altogether, 26 bacterial phyla were detected from the Red Sea sponges, 11 of which were absent from the surrounding sea water and 4 were recorded in sponges for the first time. Up to 100 OTUs with richness estimates of up to 300 archaeal species were revealed from a single sponge species. This is by far the highest archaeal diversity ever recorded for sponges. A non-negligible proportion of unclassified reads was observed in sponges. Our results demonstrated that the sponge-associated microbial communities remained highly consistent in the same sponge species from different locations, although they varied at different degrees among different sponge species. A significant proportion of the tag sequences from the sponges could be assigned to one of the sponge-specific clusters previously defined. In addition, the sponge-associated microbial communities were consistently divergent from those present in the surrounding sea water. Our results suggest that the Red Sea sponges possess highly sponge-specific or even sponge-species-specific microbial communities that are resistant to environmental disturbance, and much of their microbial diversity remains to be explored.

The ISME Journal (2011) **5,** 650–664; doi:10.1038/ismej.2010.165; published online 18 November 2010
Subject Category: microbe–microbe and microbe–host interactions
Keywords: sponge; pyrosequencing; microbial community structure; sponge specific

Introduction

Marine sponges (phylum Porifera) are considered to be one of the oldest multicellular animals (metazoans) with a history of more than 600 million years, yet possess a very low degree of tissue differentiation and coordination (Simpson, 1984). They have attracted substantial research interest because of their ecological importance and their production of a wide range of bioactive compounds for pharmacological use (Vogel, 2008). One striking characteristic of sponges is their association with a remarkable array of microorganisms, including archaea (Preston *et al.*, 1996), bacteria (Hentschel *et al.*, 2001; Webster and Hill, 2001), cyanobacteria (Thacker and Starnes, 2003), unicellular algae (Vacelet, 1982), dinoflagellates (Garson *et al.*, 1998) and fungi (Maldonado *et al.*, 2005), which can constitute up to 60% of the tissue volume in some sponge species (Vacelet and Donadey, 1977; Hentschel *et al.*, 2003) and exceed a density of 10^9 microbial cells per ml of sponge tissue (Webster and Hill, 2001). The extraordinary highly abundant and diverse microorganisms in sponges have led to ecological questions on their role, consistency and specificity of their association, and it is still enigmatic how the association is established and maintained.

One fundamental but not-yet-resolved question on sponge–microbe association is the true diversity of microbial associates in sponges. On the basis of immunological evidence in the 1980s (Wilkinson, 1984), this type of association began in the Precambrian when a widespread, sponge-specific bacterium formed a single species group in a sponge, which was distinct from the isolates found in the surrounding sea water. Most of the early studies

Correspondence: P-Y Qian, Department of Biology, KAUST Global Partnership Program, The Hong Kong University of Science and Technology, Clear Water Bay, Hong Kong.
E-mail: boqianpy@ust.hk
Received 17 May 2010; revised 13 September 2010; accepted 21 September 2010; published online 18 November 2010

on sponge–microbe associations heavily relied on morphological descriptions using microscopy technique (Vacelet and Donadey, 1977) and numerical taxonomy based on cultivation (Santavy *et al.*, 1990). These studies are generally limited by the fact that microbes are largely indistinguishable by morphology and most of them are not cultivable (Amann *et al.*, 1995; Eilers *et al.*, 2000). In recent years, advances in molecular techniques, such as 16S rRNA gene-based clone library, fluorescence *in situ* hybridization, denaturing gradient gel electrophoresis and terminal restriction fragment length polymorphism (T-RFLP), provide a culture-independent means to improve resolution on revealing the microbial diversity and to make a more accurate phylogenetic affiliation assessment for microbes in a complex community such as sponges. Using these techniques, 20 bacterial and 2 archaeal phyla were discovered in sponges from different geographic locations up to 2008 (Taylor *et al.*, 2007; Webster *et al.*, 2008; Zhu *et al.*, 2008). In addition, sponges from geographically separated regions shared high similarity in their bacterial communities, and sequence clusters specific for sponges absent from the surrounding sea water were identified (Friedrich *et al.*, 2001; Hentschel *et al.*, 2002). These studies suggest a stable and specific relationship between sponges and bacteria. However, debate on the stability and specificity of sponge–microbe association arises because of contradictory results showing variations in bacterial communities associated with sponges from different genera, locations or time scale (Qian *et al.*, 2006; Wichels *et al.*, 2006; Lee *et al.*, 2007). In a more recent study, Lee *et al.* (2009b) demonstrated that consistent bacterial communities only hold for certain species of sponges. The discrepancies may be attributed to the methods used in different studies, yet there is concern that the resolution of all the above-mentioned methods may lead to inconcrete conclusions.

The massively parallel tag pyrosequencing technology is an extremely high throughput sequencing approach that can generate up to four million base pairs (bp) of sequences from multiple samples in a single 10-h reaction (Margulies *et al.*, 2005; Binladen *et al.*, 2007). Using this technology, Sogin *et al.* have demonstrated that the previous molecular methods underestimated the true bacterial diversity by one to two orders of magnitude, and most of the observed phylogenetic diversity revealed by this new technique was represented by thousands of low-abundance populations (Sogin *et al.*, 2006). Extremely high diversity of bacterial ribotypes in three sponge species from Australia's Great Barrier Reef and the vertical transmission of certain groups of 'sponge-specific' members from sponge adults to larvae were also successfully revealed using this method (Webster *et al.*, 2010). However, this technique has not been fully utilized to study the spatial variations of microbial communities in sponges, which may help to address the issue of stability and specificity of the microbial associates.

The Red Sea, famous for its perennial high temperature and high salinity of seawater bodies, is a unique and largely unexplored marine ecosystem in the world. An extensive area of coral reefs (up to 2000 km) and different habitats along the coast of the Red Sea nourish diverse communities of corals and sponges. The Red Sea sponges have been studied during the past two decades for their natural products and bioactive compounds, as well as for their ecological importance to coral reefs (Ilan *et al.*, 2004). There are about 240 sponge species recorded in the Red Sea (Radwan *et al.*, 2010), yet only a few with their microbial communities have been studied (Hentschel *et al.*, 2002; Oren *et al.*, 2005; Radwan *et al.*, 2010). In this study, we used the massively parallel tag pyrosequencing technology to better explore the diversity, stability and specificity of microbial communities associated with different sponge species from different locations of the Red Sea. We also compared the sponge-associated microbial communities with those present in the surrounding sea water to assess the influence of environmental conditions.

Materials and methods

Sample collection and DNA extraction

Sponge specimens and sea water were collected from four different locations (habitats) along the coast of the Red Sea in Saudi Arabia in April 2009 (Table 1): site 1 is close to the outlet of Sharm Obhur, whereas site 2 is situated within the same Sharm but with residential and industrial buildings along both sides of shore; site 4 is about 50 km away from sites 1 and 2 and is at Abu Madafi Reef—one of the best coral reef areas in the Red Sea; site 5 serves as a non-reef reference site of 25 km away from site 4 and 60 km away from sites 1 and 2. At each site, sponge specimens ($n=3$) were collected by scuba divers or snorkelers, and 4 l of sea water in close proximity to the sponges was also collected. *In situ* environmental parameters were measured ($n=3$) using a multi-parameters water quality monitoring sonde (YSI 6600, YSI, USA). Samples were immediately transported back to the laboratory in separate sterile plastic bags. On arrival at the laboratory, sponge specimens were flushed with 0.22 μm-membrane-filtered sea water to remove loosely attached microbes and debris. About 10 ml of sponge tissues for each sponge species was preserved in 70% of ethanol for species identification ($n=1$), whereas 0.5 ml of the internal tissues from each replicate was cut into small pieces with sterile razor blades, and then frozen in 0.8 ml of extraction buffer (100 mM of Tris-HCl, 100 mM of Na_2-EDTA, 100 mM of Na_2HPO_4, 1.5 M of NaCl, 1% of CTAB; at pH 8) for DNA extraction ($n=3$). Suspended particles and eukaryotes in the seawater samples were removed by

Table 1 Sponge species collected from different locations along the coast of the Red Sea

Sites	*Location*	*Depth (m)*	*Sponge species*	*Sample ID*	*Taxonomy*		
					Class	*Order*	*Family*
1	21°42.32'N 39°04.23'E	12	*Hyrtios erectus*	HE-1	Demospongiae	Dictyoceratida	Thorectidae
			Stylissa carteri	SC-1	Demospongiae	Halichondrida	Dictyonellidae
			Xestospongia testudinaria	XT-1	Demospongiae	Haplosclerida	Petrosiidae
2	21°42.64'N 39°05.69'E	8	*Hyrtios erectus*	HE-2	Demospongiae	Dictyoceratida	Thorectidae
			Stylissa carteri	SC-2	Demospongiae	Halichondrida	Dictyonellidae
			Xestospongia testudinaria	XT-2	Demospongiae	Haplosclerida	Petrosiidae
4	22°03.66'N 38°46.07'E	19	*Hyrtios erectus*	HE-4	Demospongiae	Dictyoceratida	Thorectidae
			Stylissa carteri	SC-4	Demospongiae	Halichondrida	Dictyonellidae
5	22°10.18'N 38°57.40'E	15	*Hyrtios erectus*	HE-5	Demospongiae	Dictyoceratida	Thorectidae
			Stylissa carteri	SC-5	Demospongiae	Halichondrida	Dictyonellidae

filtering 1 l of sea water through a 1.6 μm-pore-sized glass fiber membrane (GF/A, diameter 125 mm, Whatman, Wisconsin, UK), and the microbial cells therein were subsequently retained on a 0.22 μm polycarbonate membrane (polycarbonate, dia. 47 mm, Millipore, Massachusetts, UK). The polycarbonate membrane was then frozen in 0.8 ml of extraction buffer for DNA extraction ($n = 3$). Total genomic DNA was extracted according to the modified sodium dodecyl sulfate-based method described by Lee *et al.* (2009b) and further purified by the Mo Bio soil DNA isolation kit (Mo Bio Laboratories, Carlsbad, CA, USA) following the manufacturer's protocols. The quality and quantity of DNA were checked with a NanoDrop device (ND-1000, Thermo Fisher, USA) and kept at −20 °C until use. The remaining 1 l of seawater sample was used for the determination of microbial cell density by DAPI staining, followed by counting under an epifluorescent microscope ($n = 3$), and for the measurement of nutrient content by using a TOC machine (TOCV-CPH, Shimadzu, Japan) and a nutrient analyzer (Skalar 4000, Skalar, The Netherlands) following the manufacturers' protocols ($n = 3$).

Terminal restriction fragment length polymorphism (T-RFLP) analysis of bacterial community structure

The 16S rRNA genes (rDNA) in the total genomic DNA were amplified by PCR using the universal primers 341F (5′-CCTACGGGAGGCAGCAG-3′) and 926R (5′-CCGTCAATTCCTTTRAGTTT-3′) (Liu *et al.*, 1997) and analyzed by T-RFLP following the protocols stated by Lee *et al.* (2009b). Cluster analysis and one-way analysis of similarity (ANOSIM) of the T-RFLP patterns were performed with the computer program PRIMER v5.2.4 (Plymouth Marine Laboratory, Plymouth, UK). The cluster analysis was based on the similarity matrices created by comparing the relative abundance, and on the presence or absence (determined by threshold relative intensities of 0.5 and 1%) of individual T-RFs among samples using the Bray–Curtis method (Clarke and Warwick, 1994). Global and pairwise differences among individual samples were tested by ANOSIM (Clarke, 1993). The ANOSIM R statistic is calculated on the basis of difference in mean ranks between and within groups.

Sample preparation for pyrosequencing

The 16S rDNA was also amplified by PCR for multiplexed pyrosequencing using barcoded primers. A set of primers was designed by adding a 6-nucleotide barcode (Supplementary Table S1) to the universal forward primer U789F (5′-TAGATACC CSSGTAGTCC-3′) and the reverse primer U1068R (5′-CTGACGRCRGCCATGC-3′) for amplification of bacteria and archaea (Baker *et al.*, 2003). These primers cover about 94.8–97.7% of publicly available 16S rDNA (Wang and Qian, 2009) and target at the hypervariable V6 region. A 100-μl PCR reaction mixture contained 5 U of *Pfu* Turbo DNA polymerase (Stratagene, La Jolla, CA, USA), 1 × *Pfu* reaction buffer, 0.2 mM of dNTPs (TaKaRa, Dalin, China), 0.1 μM of each barcoded primer and 20 ng of genomic DNA template. PCR was performed with a thermal cycler (Bio-Rad, USA) under the following conditions: initial denaturation at 94 °C for 5 min; 26 cycles at 94 °C for 30 s, 53 °C for 30 s and 72 °C for 45 s; and a final extension at 72 °C for 6 min. PCR products were purified using the TaKaRa Agarose Gel DNA Purification Kit (TaKaRa, China) and quantified with the NanoDrop device. A mixture of PCR products was prepared by mixing 200 ng of the purified 16S amplicons from each sponge and seawater sample and then pyrosequenced on the ROCHE 454 FLX Titanium platform (Roche, Basel, Switzerland) at the National Human Genome Centre of China at Shanghai, China, according to the manufacturer's manual.

Species richness estimators and taxonomic assignment of pyrosequencing reads

Raw pyrosequencing data that resulted from this study were submitted at the NCBI Short Reads Archive database (SRA012874.2). Pyrosequencing

reads with ambiguous nucleotides, shorter than 90-nucleotides or without a complete barcode and primer at one end were removed and excluded from further analysis. Qualified reads belonging to different samples were sorted out according to the corresponding barcodes by using a bespoken Java program. Multiple alignments of the reads from each sample were performed with MUSCLE (Edgar, 2004) and MAFFT (Katoh, 2008), and then input into DNAdist in the Phylip package (version 3.6) to produce a matrix of genetic distance. This matrix served as an input to DOTUR (Schloss and Handelsman, 2005) to assign reads into operational taxonomic units (OTUs) to generate rarefaction curves and to calculate the non-parametric abundance-based coverage estimator (Chao and Lee, 1992) and the bias-corrected species richness estimator Chao1 (Chao, 1984) at different dissimilarity levels.

The tag pyrosequence reads were classified into different taxonomic groups by comparison with those entries in the RDP database version 10.8, which includes >64 916 bacterial and >33 082 archaeal sequences (of which 275 057 and 6624 are in full length, respectively), using RDP classifier (Cole *et al.*, 2008). A 50% confidence threshold was applied for assigning reads at phylum level. This threshold was demonstrated to be sufficient for accurate assignment of short sequences (Liu *et al.*, 2008; Claesson *et al.*, 2009). Reads assigned to a known phylum were further classified to lower taxonomic levels, yet the 50% threshold was not used for the assignment. The number of representative reads for each OTU determined at a 3% dissimilarity at each taxonomic level was counted and the proportions of each group in a sample were calculated.

For the assignment of reads to previously described sponge-specific clusters, 16S rRNA sequences were first extracted from an ARB database containing the sponge-specific cluster sequences (Taylor *et al.*, 2007). V5 and V6 regions of the recruited cluster sequences were split by recognizing the U789F and U1068R primers, and then aligned with the tag pyrosequence reads using MUSCLE (Edgar, 2004). Similarity matrix was calculated as mentioned above and a minimum of 75% sequence similarity threshold was applied for assigning the tag sequences into the cluster showing the highest similarity (Webster *et al.*, 2010). The relative abundance and occurrence of tag assigned into sponge-specific clusters were visualized as a heatmap using JColorGrid (Joachimiak *et al.*, 2006).

Comparison of microbial communities

The similarity among the microbial communities in sponges and sea water from different sampling sites was determined using UniFrac analysis (http://bmf.colorado.edu/unifrac/) (Lozupone and Knight, 2005; Lozupone *et al.*, 2006). The input file was parsed from the OTU output determined at a dissimilarity level of 3%. The number of reads and the sequence identity of a representative read from each OTU were defined according to the origin, and used as the environmental file for the UniFrac analysis. Multiple alignments for DNA sequences from each representative OTU were performed by MUSCLE (Edgar, 2004), and a matrix of genetic distance was produced by DNAdist in the Phylip package (version 3.6). The matrix was then used to construct a neighbor-joining tree with the Phylip package. The tree and environment files were processed with ClusterEnvs in UniFrac to group the samples by the unweighted-pair group method using average linkages (UPGMA) clustering.

Results

Environmental parameters of the study sites

Sponge samples were collected from different depths, ranging from 8 to 19 m, at four different locations along the coast of the Red Sea, Saudi Arabia (Table 1). The temperature, salinity and pH of sea water from those sampling sites did not differ substantially, but variations were observed for dissolved oxygen content, turbidity, nutrient contents (Table 2) and microbial cell density. Site 1 had the highest concentrations of chlorophyll a, total nitrogen, nitrate and ammonium, but the lowest turbidity, whereas site 2 recorded the highest concentrations of dissolved oxygen and total phosphorus but the lowest nitrate concentration. These two sites were generally characterized by high nutrient concentration and considered as relatively polluted areas. In contrast, low concentrations of dissolved oxygen, chlorophyll a, total nitrogen, ammonium and phosphorus were observed at sites 4 and 5, which were considered as clean sites but characterized by the lowest and highest total organic carbon content, respectively. The microbial cell density in sea water was the lowest at site 4 (4×10^5 cells ml^{-1}) but the highest at site 5 (8×10^5 cells ml^{-1}).

Bacterial community structure revealed by T-RFLP analysis

A total of 164 T-RFs of different sizes were obtained from the sponge and seawater samples and the occurrence of T-RFs in each sample was summarized in Supplementary Table S2. The number of T-RFs detected from each sample ranged from 14 to 40, and the numbers were generally higher among samples from sponges *Hyrtios erectus* and *Xestospongia testudinaria* (with an average of 31 T-RFs) than from those from sea water (an average of 22 T-RFs), whereas the lowest was observed for the sponge *Stylissa carteri* (an average of 17 T-RFs). The numbers of T-RFs for the same type of sample (that is, sea water or different species of sponge) from different sampling sites showed little variations. Cluster analysis indicated that the bacterial

Table 2 Environmental parameters of the sampling locations in this study

Site	*Sample ID*	*Temp (°C)*	*Sal (ppt)*	*DO (mg l^{-1})*	*PH*	*Turbidity (NTU)*	*Chl a (μg l^{-1})*	*TOC (mg l^{-1})*	*TC (mg l^{-1})*	*TN (mg l^{-1})*	*NO_3 (mg l^{-1})*	*NH_4 (mg l^{-1})*	*TP (μg l^{-1})*	*Microbial density ($\times 10^3$ cell per ml)*
1	W-1	/	39.15	6.64	7.62	0.233	0.50	3.591	30.39	0.496	0.210	1.256	2.29	573.61
2	W-2	26.79	39.21	6.72	7.80	0.833	0.33	2.247	28.92	0.192	0.116	1.195	3.01	772.82
4	W-4	25.80	39.57	5.70	7.89	1.233	0.23	1.767	26.58	0.124	0.126	0.888	1.58	408.47
5	W-5	26.46	39.70	4.99	7.95	0.700	0.40	6.151	30.74	0.119	0.168	0.291	0.01	817.81

Abbreviations: NH_4, ammonium; Chl a, chlorophyll a; DO, dissolved oxygen; NO_3, nitrate; Sal, salinity; TC, total carbon; Temp, temperature; TN, total nitrogen; TOC, total organic carbon; TP, total phosphorus.
Seawater samples were collected from each site for measurement. Temp, Sal, DO, pH, turbidity and Chl a content were measured by YSI; TOC and TC were measured by the TOC machine, whereas TN, NO_3, NH_4 and TP were measured by the nutrient analyzer; microbial density was determined by DAPI staining followed by counting on an epifluorescent microscope.

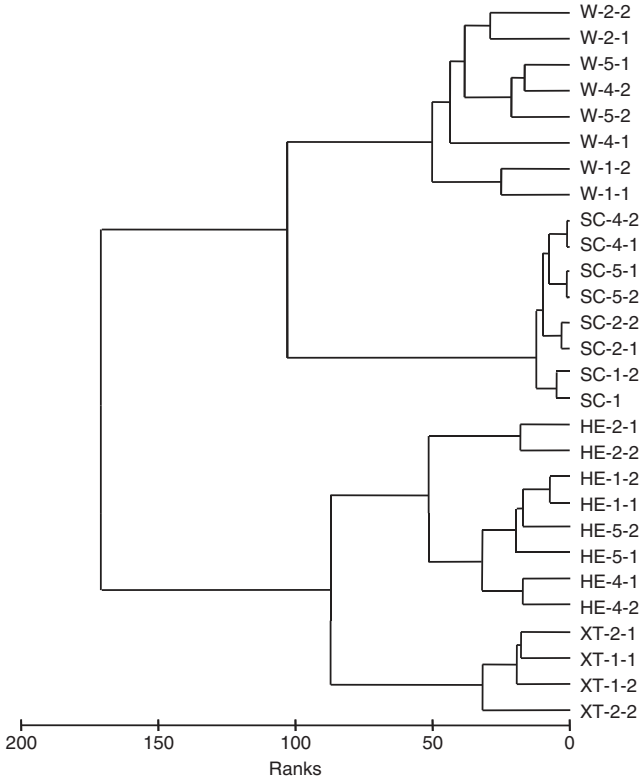

Figure 1 Dendrogram showing the similarity of bacterial communities, revealed by T-RFLP analysis of the 16S rDNA, from sponge and seawater samples from different sites. To simplify the dendrogram, only two randomly chosen replicates for each type of sample were shown here. Refer to Tables 1 and 2 for sample abbreviations.

communities in replicate samples were highly similar (Figure 1). The T-RFLP profiles of the seawater samples from different sites were grouped into one large cluster, with no clear distinction among the sites (Figure 1; R values ranged from 0.495 to 0.995, $P<0.05$, Supplementary Table S3), and were distantly related to other three clusters formed by the three sponge species (Figure 1; $R=1$, $P<0.05$, Supplementary Table S3). The samples from each sponge species, regardless of the sampling site, also shared high similarity in their bacterial communities, as shown by individual clusters formed for each sponge species (Figure 1), and generally low R values in ANOSIM comparison (Supplementary Table S3). However, those from the sponges *H. erectus* and *X. testudinaria* were more similar to each other than to those from *S. carteri* (Figure 1; Supplementary Table S3). Cluster analyses based on the relative abundance of T-RFs and on different levels of threshold intensities showed similar clustering patterns (data not shown).

Diversity and species richness estimators of microbial communities

Approximately 200 000 raw pyrosequencing reads of the 16S rDNA spanning the hypervariable regions V5 and V6 were obtained from the seawater and sponge samples. After screening out the noise and poor-quality reads, about 140 000 reads with an average read length of 290–300 bp were used for subsequent analyses, of which 80% were obtained from the sponge samples (Supplementary Table S4). Rarefaction analysis, based on OTUs at 3% dissimilarity, indicated that most of the bacterial libraries, for instance, W-2, W-4 SC-1 and HE-5, and some of the archaeal libraries, including SC-1, SC-2 and HE-4, well represented the microbial communities as the rarefaction curves were approaching plateaus; yet, some libraries, particularly for XT-1, XT-2 and HE-2, may require deeper sequencing to avoid underestimation of microbial diversity in the samples (Figures 2a and 3a). Rank-abundance curves indicated that a majority of the reads belonged to rare organisms represented by only a few sequence tags, whereas all samples contained relatively low proportions of highly abundant bacteria or archaea (Figures 2b and 3b).

The number of OTUs and estimated species richness at 3%, 5% and 10% dissimilarity are provided in Supplementary Table S4. At a 3% dissimilarity level, the number of bacterial OTUs

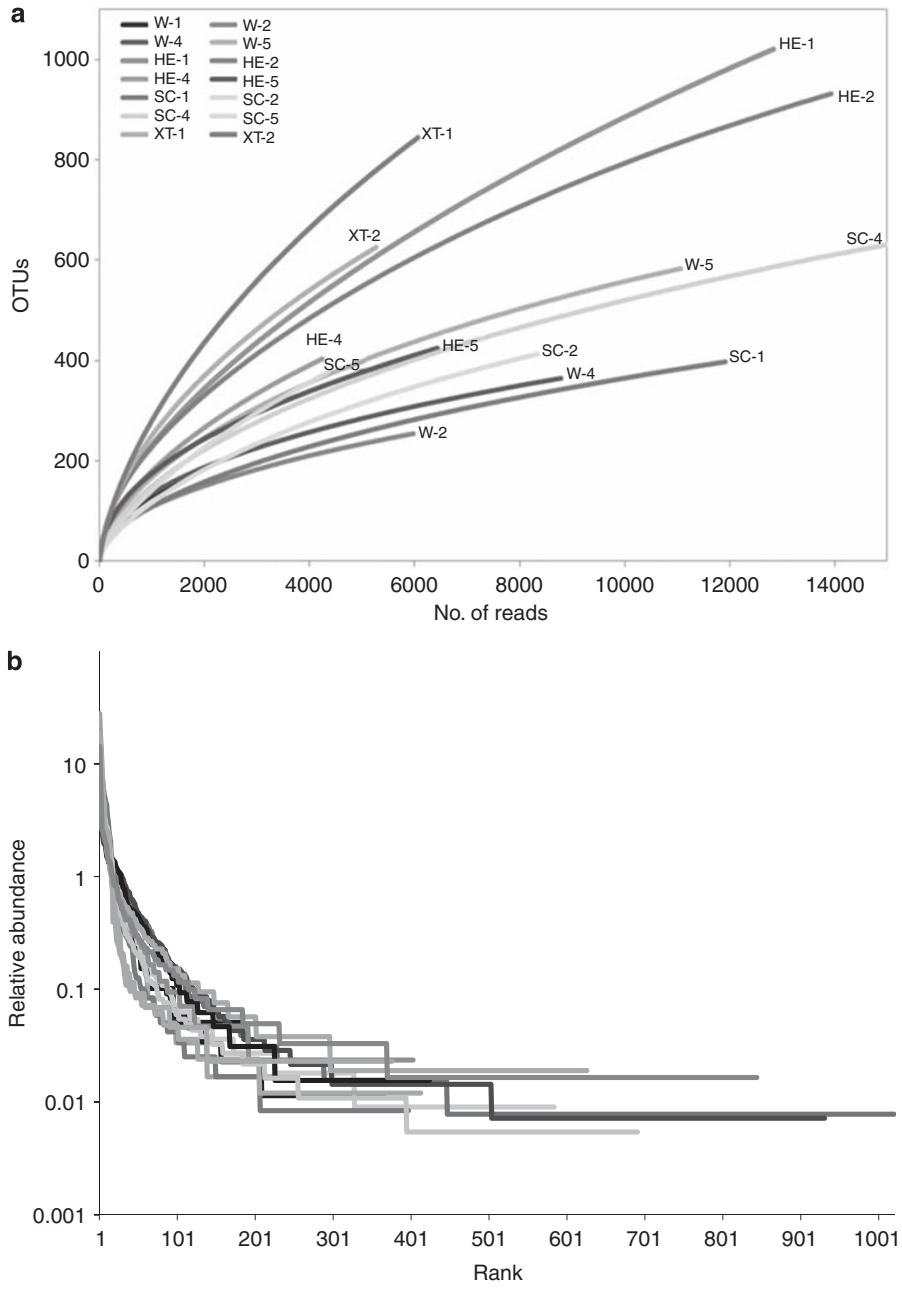

Figure 2 Diversity of bacterial communities in sponges and sea water from the Red Sea. (**a**) Rarefaction curve and (**b**) rank-abundance curve based on bacterial OTUs at a dissimilarity level of 3%. Refer to Tables 1 and 2 for sample abbreviations.

was generally higher in the sponge samples, with the highest number of 1020 OTUs detected in *H. erectus* from site 1 than in the seawater samples, with the lowest of 182 OTUs recovered from sea water from site 1 (Figure 4, Supplementary Table S4). Abundance-based coverage estimator and Chao1 estimated up to 2386 and 2054 bacterial species in *H. erectus* from site 1, whereas the two values for the seawater samples were as low as 331 and 292 bacterial species from the same site (Supplementary Table S4). For the seawater samples, the highest number of bacterial OTUs and species richness were observed at site 5, followed by those at site 4. Among the sponge samples, the number of bacterial OTUs and estimated diversity were generally lower in *S. carteri*, except for SC-4. For *H. erectus*, those at sites 1 and 2 exhibited higher bacterial diversity than those at sites 4 and 5, which were in contrast to the case in the seawater samples in which higher diversity was found at sites 4 and 5 (Figure 4). The bacterial diversity in *X. testudinaria* from both sites 1 and 2 was comparable to that in *H. erectus* from site 2. No clear relationship was observed between bacterial diversity and sampling site. Similar observations were made for OTUs and estimators determined at 5% and 10% dissimilarity (Supplementary

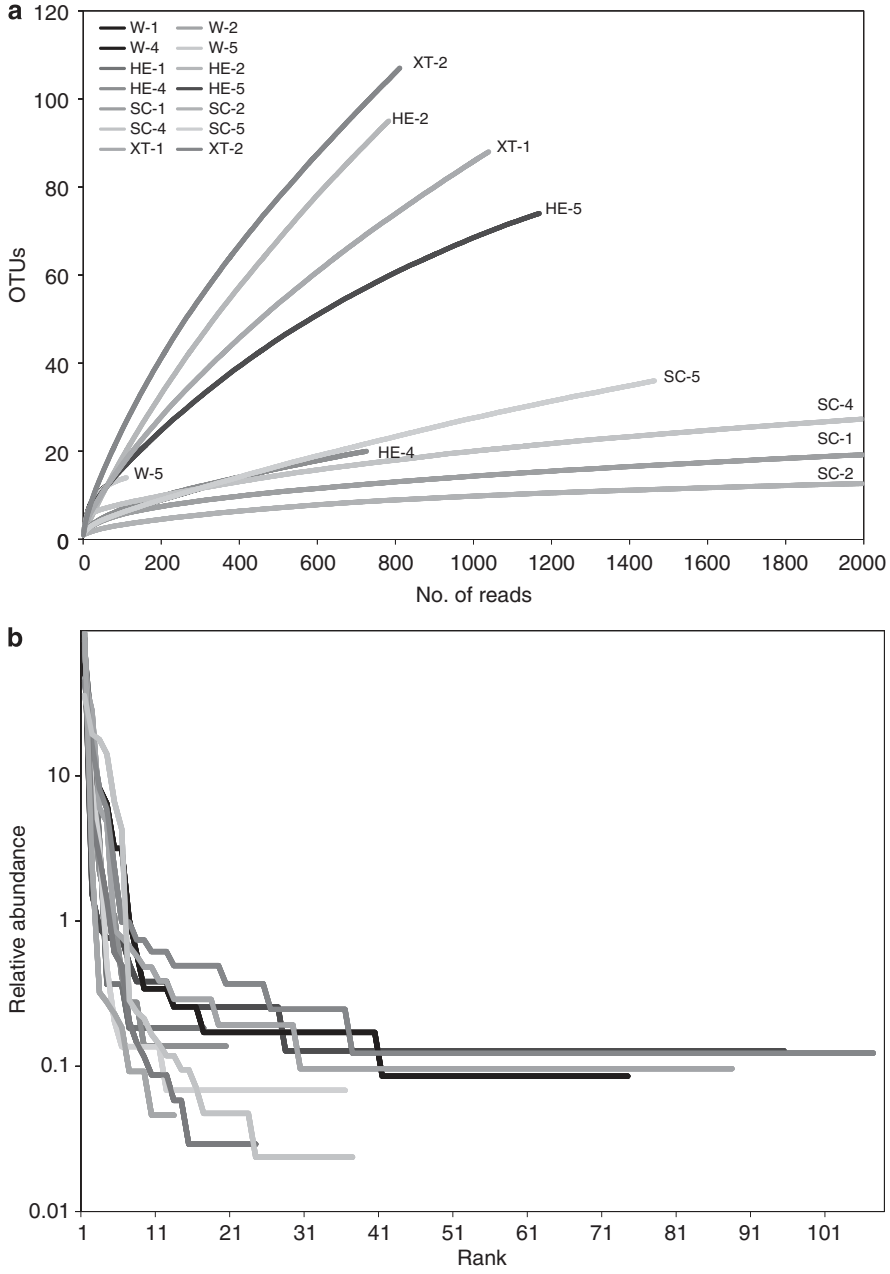

Figure 3 Diversity of archaeal communities in sponges and sea water from the Red Sea. (**a**) Rarefaction curve and (**b**) rank-abundance curve based on archaeal OTUs at a dissimilarity level of 3%. Refer to Tables 1 and 2 for sample abbreviations.

Table S3). Although the archaeal diversity of the seawater and sponge samples was far lower than the bacterial diversity, the following similarities between them were observed: (1) archaeal diversity was generally higher in sponge samples than in seawater samples; (2) *S. carteri* showed the lowest archaeal diversity among sponge samples; and (3) archaeal diversity of the seawater samples from sites 4 and 5 was higher than that from sites 1 and 2, yet this was not the case for *H. erectus*, suggesting no clear relationship between archaeal diversity and sampling site (Supplementary Table S4).

Taxonomic assignment of bacterial and archaeal pyrosequencing reads

At a confidence threshold of 50%, 132 837 out of 136 183 qualified reads (that is, 97.5%) could be assigned to a known phylum using the RDP classifier, of which 115 612 belonged to bacteria and 17 225 were affiliated with archaea (Supplementary Table S4). Most of the unclassified reads were associated with sponge samples (ranging from 0.6 to 6% of the qualified reads from a sponge sample), whereas few (<0.02% of the qualified reads) from the seawater samples were unclassified (Supplementary Figure S1). Almost all of the

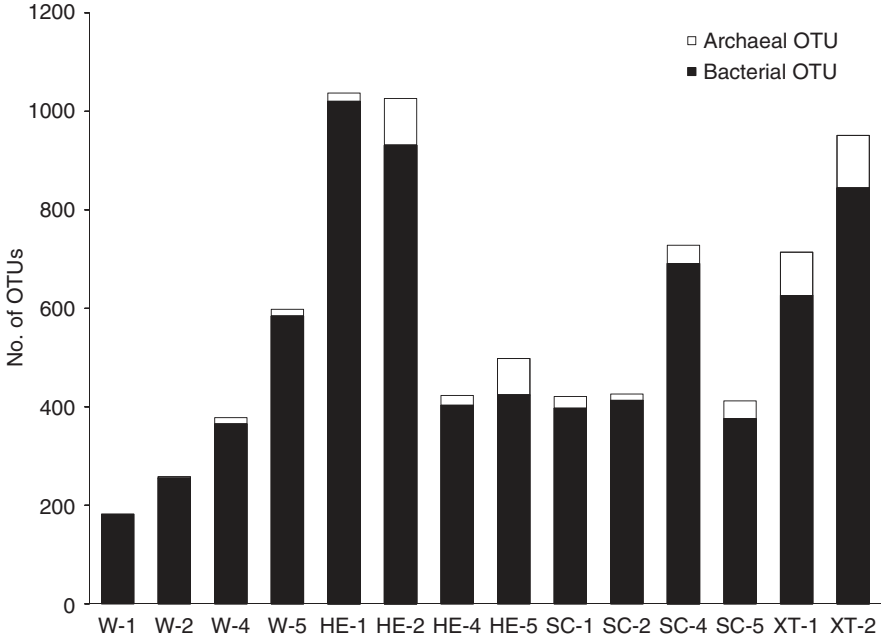

Figure 4 Number of similarity-based OTUs detected in the indigenous planktonic and sponge-associated microbial communities. Each qualified read obtained from pyrosequencing was compared with the entries in the RDP database to classify it into 'bacteria' or 'archaea' at a confidence threshold of 50%. OTU was determined at a dissimilarity level of 3%. Refer to Tables 1 and 2 for sample abbreviations.

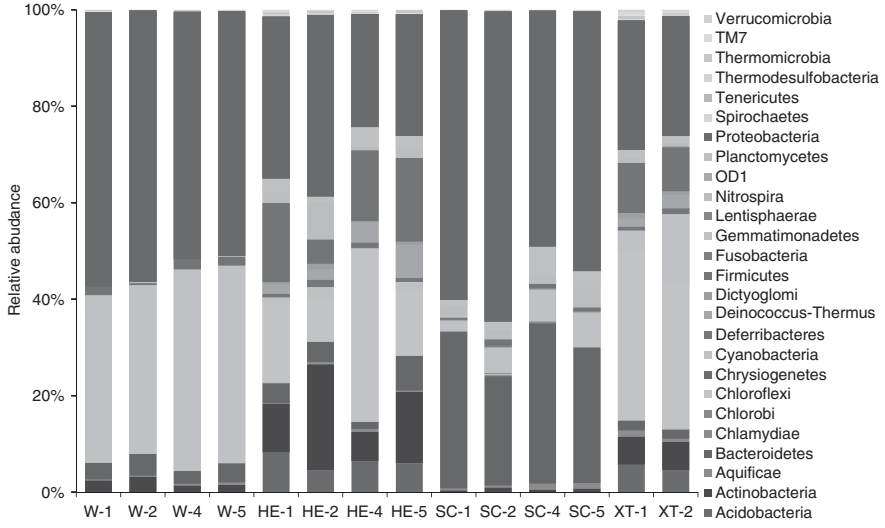

Figure 5 Taxonomic classification of bacterial reads retrieved from different sponge species and sea water at phylum level using RDP classifier with a confidence threshold of 50%. Refer to Tables 1 and 2 for sample abbreviation.

classified reads from the seawater samples (>99%) belonged to bacteria. The percentage of bacterial reads in the sponge samples varied from 70% (in SC-5) to 96% (in HE-1) (Supplementary Figure S1). In general, more archaeal reads were retrieved from the sponge *S. carteri* (19–28%) than from the other two sponge species (4–15%) (Supplementary Figure S1).

Altogether, 26 bacterial phyla were recovered from our samples, with 9–12 different phyla found in seawater samples and 19–24 retrieved from sponge samples (Figure 5; Supplementary Table S5). More than 90% of the bacterial reads from the seawater samples were affiliated with two dominant phyla, Proteobacteria and Cyanobacteria, whereas reads belonging to Actinobacteria, Aquificae, Bacteroidetes, Chlamydiae, Chlorobi, Chloroflexi, Dictyoglomi, Firmicutes, Deinococcus-Thermus, Planctomycetes, Thermomicrobia, Verrucomicrobia and the candidate phylum TM7 were found to be the minor groups. In addition to all these phyla, another 11 phyla including Acidobacteria, Chrysiogenetes, Deferribacteres, Fusobacteria, Gemmatimonadetes, Lentisphaerae, Nitrospira, Spirochetes, Tenericutes, Thermodesulfobacteria and the candidate phylum

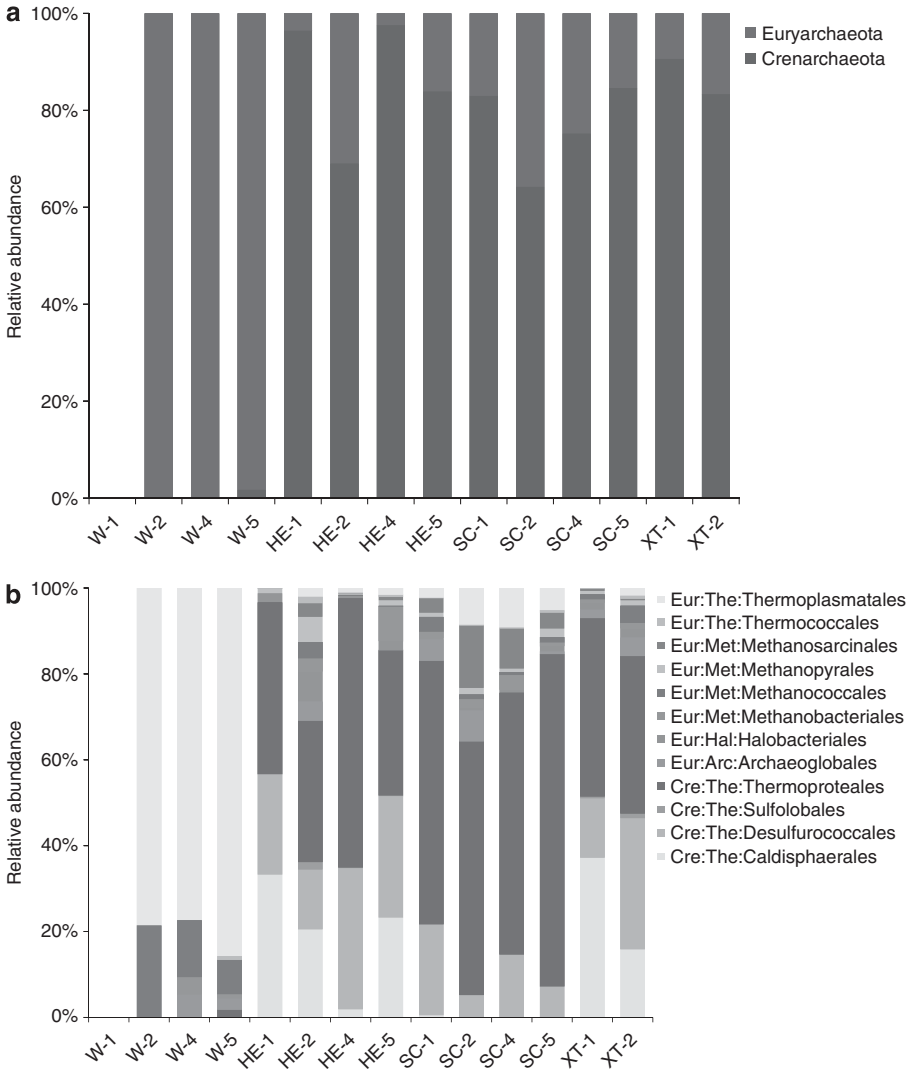

Figure 6 Taxonomic classification of archaeal reads retrieved from different sponge species and sea water at (**a**) phylum and (**b**) order levels using RDP classifier with a confidence threshold of 50%. Refer to Tables 1 and 2 for sample abbreviation.

OD1 were recovered from sponge samples. When comparing the bacterial communities in different sponge species, Proteobacteria, Firmicutes and Chloroflexi, altogether constituting up to 52–73% of the reads from a sponge sample, dominated the communities in *H. erectus* and *X. testudinaria,* whereas Proteobacteria was the major phylum in *S. carteri*, which made up to 60% of the reads, followed by Bacteroidetes (up to 33%) (Figure 5). When the bacterial communities at lower taxonomic levels were examined, 35 classes and 83 orders were recovered (Supplementary Figure S2). The bacterial communities in sea water were heavily loaded with Rhodospirillales in Alphaproteobacteria and Cyanobacteria Family II, yet their proportions diminished drastically in sponge samples. Gammaproteobacteria, particularly Oceanospirillales and Sphingobacteria, were the most dominant bacteria in *S. carteri,* whereas, in contrast, Anaerolineae, Clostridia and Deltaproteobacteria constituted a significant portion of reads from *H. erectus* and *X. testudinaria*. At all the three levels of taxonomic classification, the bacterial community in the seawater samples was less diverse than that in sponges, and those associated with sponges *H. erectus* and *X. testudinaria* were more complex and contained more rare bacteria. In addition, the bacterial compositions in sea water or the same sponge species from different sites did not vary substantially.

Two archaeal phyla, Euryarchaeota and Crenarchaeota, which have been previously reported in sponges and sea water from other places, were found in our samples, but their relative abundance in different samples varied (Figure 6a). Except for W-1, in which no archaeal read was found, nearly all the archaeal reads from seawater samples were classified into Euryarchaeota; yet, a high proportion of the archaeal reads from the sponge samples (ranging from 64 to 98%) was affiliated with Crenarchaeota, and the proportion of the two

archaeal phyla also varied among the same species of sponges from different locations. At the order level, the archaeal communities in sea water from sites 2, 4 and 5 were largely composed of Thermoplasmatales, followed by Methanococcales, but both groups reduced substantially in sponges (Figure 6b). The archaeal communities in *H. erectus* and *X. testudinaria* were dominated mainly by Thermoproteales, Desulfurococcales and Caldisphaerales, whereas the former two orders were also found in *S. carteri* and the latter was almost absent in *S. carteri*. Similar to the bacterial communities, the archaeal communities were less diverse in seawater samples than in sponges, whereas those in *H. erectus* and *X. testudinaria* showed a greater diversity and contained more archaeal groups of low abundance.

Assignment of tag sequences into 'sponge-specific' clusters

About 35.9–64.8% of the tag sequences from an individual sponge sample could be assigned to one of the sponge-specific clusters previously described (Taylor *et al*., 2007) (Figure 7, Supplementary Table S6). Among these sponge tags, most of them were assigned to clusters for the phyla Cyanobacteria (SC1 and SC2), Chloroflexi (SC13, SC15 and SC17), Poribacteria (SC25), Actinobacteria (SC34), Gammaproteobacteria (SC60, SC63 and SC65), Crenarchaeota (SC40) and some uncertain affiliations (SC26 and SC28), whereas the rest were found at low abundance in other sponge-specific clusters. For the tag sequences from seawater samples, only about an average of 8% of the reads resembled those in the sponge-specific clusters and most of them were affiliated with Cyanobacteria (SC1 and SC2), Nitrospira (SC23), Spirochaetes (SC29) and Alphaproteobacteria (SC47 and SC52 at a very low abundance (Figure 7; Supplementary Table S6).

Similarity among microbial communities

The results of UniFrac UPGMA cluster analysis, based on pyrosequencing, on the similarity of bacterial communities among seawater and sponge samples from different locations were in good agreement with those revealed by T-RFLP analysis. UniFrac cluster analysis showed that the bacterial communities in sea water from different locations were highly similar but drastically distant from those associated with sponges (Figure 8). For sponge samples, the samples from *H. erectus* and *X. testudinaria* from different locations were grouped together, whereas all the samples from *S. carteri* fell into a separate group, suggesting a higher similarity in the bacterial communities associated with *H. erectus* and *X. testudinaria*, regardless of the sampling site (Figure 8). For the archaeal communities, UniFrac analysis could only differentiate the communities in sea water from those associated with sponges, yet the communities among different

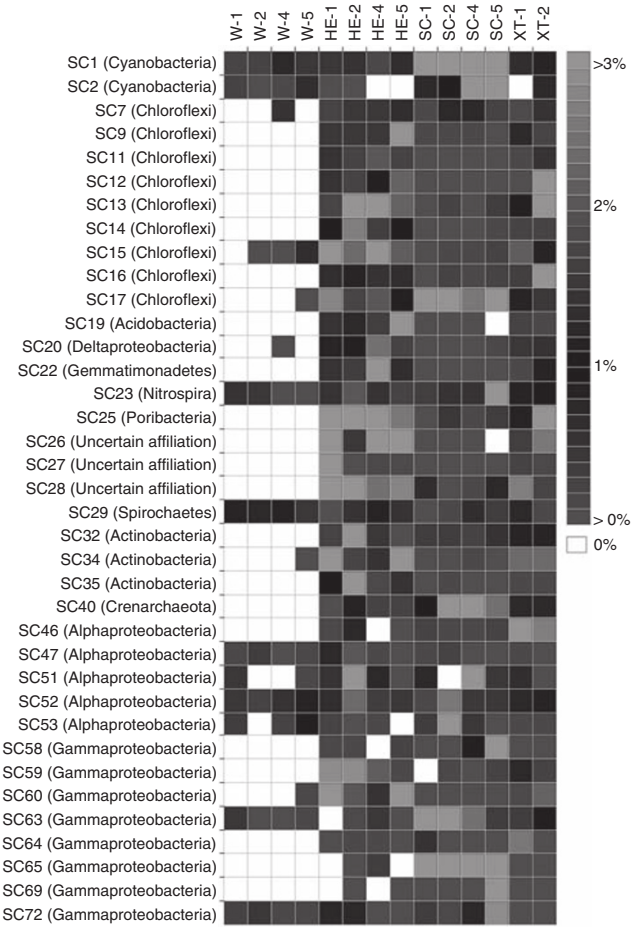

Figure 7 Heatmap showing the relative abundance and distribution of representative 16S rDNA tag sequences assigned into selected sponge-specific clusters. The divergence of >0.5% was used to filter out the clusters having small differences among the samples. The color code indicates relative abundance, ranging from blue (low abundance) to black to red (high abundance). White box indicates that no read was assigned to the corresponding cluster. Refer to Tables 1 and 2 for sample abbreviations.

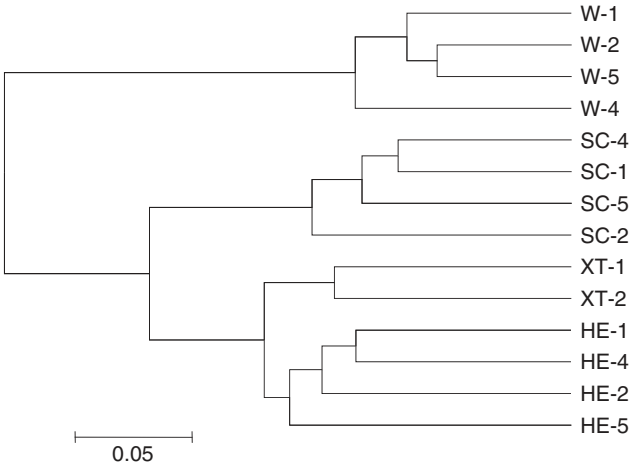

Figure 8 UniFrac UPGMA cluster of bacterial communities associated with different sponge species from different sampling locations in comparison with those from indigenous sea water. The figure was constructed on the basis of tag pyrosequencing data. Refer to Tables 1 and 2 for sample abbreviations.

species of sponges from different locations were largely indistinguishable (data not shown).

Discussion

It has long been recognized that sponges possess a wide variety of microorganisms and generally two scenarios have been proposed for the establishment and maintenance of sponge–microbe association (Taylor *et al.*, 2007). Microorganisms can be either acquired from the surrounding sea water during filter-feeding process of sponges or transferred from the parental sponges to progeny through reproductive stages. The former (horizontal or environmental transmission) requires certain degrees of recognition, possibly via the innate immune system in sponges (Müller and Müller, 2003), to differentiate symbionts from food microbes. Under this scenario, there would be some overlap between the microbes in the surrounding sea water and sponges. In contrast, for the vertical transmission scenario, as the symbionts are transferred from parent to daughter cells, and may have coevolved with sponge hosts over a long period of time and/over many generations, they would be sponge specific or species specific and may no longer be present in sea water. Thus, understanding how the sponge-associated microbial community differs from the one in sea water and among different sponge species would provide a better understanding of the development of sponge-associated microbial community. Using tag pyrosequencing technology in this study, we were the first to reveal the microbial communities associated with three Red Sea sponges, *Hyrtios erectus*, *Stylissa carteri* and *Xestospongia testudinaria*, from different locations along the coast of Saudi Arabia and to compare them with the indigenous planktonic microbial communities in the surrounding sea water. We demonstrated that the sponge-associated microbial communities were highly diverse and consistently diverged from the surrounding planktonic communities, suggesting highly sponge-specific microbial communities in the Red Sea sponges that are resistant to environmental influence. In addition, substantial variations found in the bacterial communities associated with different species of sponge supported the fact that the bacterial communities were even sponge-species specific. Furthermore, we observed high similarity but low diversity of planktonic microbial communities at different sampling sites with different environmental settings.

Influence of environment on microbial communities

On the basis of the environmental parameters measured in this study (Table 2) and some coherent data from a previous study (Al-Farawati *et al.*, 2008), the four sampling locations could be categorized into two types of environmental settings: sites 1 and 2 were generally considered as polluted sites with a high nutrient concentration, whereas sites 4 and 5 represented relatively clean areas. Sites 1 and 2 are located at the mouth and within Sharm Obhur, which is a shallow and narrow coastal inlet located about 10 km long at the north of Jeddah city on the eastern coast of the Red Sea. Many marinas, hotels, other tourist attractions and residential areas are densely packed along Sharm Obhur, which are the major sources of pollutants and sewage discharge into the area. The nutrient levels at the Sharm have been increasing substantially over the last 10 years (Al-Harbi and Khomayis, 2005; Al-Farawati *et al.*, 2008). Although the relatively open and deep nature of site 1 in comparison with site 2 may have diluted the pollutants accumulated along the Sharm, the nutrient concentrations at site 1 were much higher than those at sites 4 and 5 (Table 2), which are relatively deep, open and oligotrophic environments compared with sites 1 and 2. We hypothesized that different microbes may adapt to these two contrasting environments differentially, leading to compositional variations in microbial community in the sea water. However, our findings rejected this hypothesis and showed a different trend from those in other water bodies in which bacterioplankton communities were reported to be highly diverse and environment driven (Giovannoni and Stingl, 2005; Martiny *et al.*, 2006; Pommier *et al.*, 2007; Ghiglione *et al.*, 2008). In our study, high similarities in the indigenous bacterial communities among different sampling sites were strongly supported by tag pyrosequencing and T-RFLP analysis (Figure 1; Supplementary Table S3). Proteobacteria (the α-subdivision), Cyanobacteria and Euryarchaeota were the most dominant groups in all our seawater samples (Figures 5 and 6). Although these phyla also dominated other planktonic systems (Morris *et al.*, 2002; Giovannoni and Stingl, 2005), the microbial diversity in our samples was low in comparison with that of other oceans (Venter *et al.*, 2004; Sogin *et al.*, 2006; Webster *et al.*, 2010). All these contrasting results may suggest that the Red Sea is a very unique environment, different from other marine habitats. The microbes found in different environments in the Red Sea are more likely to be generalists, rather than specialists, which can adapt to a wide range of environmental conditions, yet their diversity is low. Another possibility may be because of different scales of geographical separations among our sampling sites in comparison with those in other studies. Our sampling sites are at most 60 km apart, which is considered as a relatively close distance at a global scale of hundreds or thousands of km.

The bacterial communities in sea water and sponge have been studied and compared with other studies based on molecular techniques with lower resolution, such as denaturing gradient gel electrophoresis, T-RFLP and conventional clone library construction (Taylor *et al.*, 2005; Lee *et al.*, 2009a, b),

as well as the tag pyrosequencing technique (Webster *et al.*, 2009). All these studies demonstrated that highly distinct bacterial communities in sea water diverged from those associated with sponges, the same observation as ours. In our study, both T-RFLP and pyrosequencing revealed significant differences in the bacterial communities associated with sponges and their corresponding indigenous planktonic communities (Figures 1, 5 and 6). Only a small proportion of T-RFs/OTUs recovered from the seawater samples could be matched with those from the sponges, and seawater-specific T-RFs were observed. Thus, we conclude that there were clear compositional differences in the microbial communities from the sponges and the surrounding sea water.

Sponge-specific and sponge-species-specific microbial communities

Unique microbial communities specific for sponges were demonstrated in a number of previous studies (Friedrich *et al.*, 2001; Hentschel *et al.*, 2002; Taylor *et al.*, 2004; Webster *et al.*, 2009). Hentschel *et al.* defined sponge-specific clusters as groups of sequences that are derived from sponges of multiple sources (that is, different host sponges or the same host species but from different locations), and are more closely related to each other than to any other sequence from non-sponge sources (Hentschel *et al.*, 2002). To examine whether any of our tag sequences fell into the sponge-specific clusters described in early studies (Hentschel *et al.*, 2002; Taylor *et al.*, 2007), we performed multiple alignment and phylogenetic analysis using our tag sequences, together with those sponge-specific sequences, and found that a significant proportion of our reads from the sponge samples could be well matched (that is, $>75\%$ sequence similarity) with those in the sponge-specific clusters but those from the seawater samples could not (Figure 7). This result suggested that the microbes associated with the Red Sea sponges shared a certain degree of similarity with those associated with sponges from other oceans, and further supported the existence of sponge-specific groups across geographic and phylogenetic separations.

Considering the influence of species type on the sponge-associated microbial communities, we clearly observed higher similarities in bacterial communities among the sponges of the same species from different locations. This observation was supported by T-RFLP, tag pyrosequencing and statistical analyses (Figures 1 and 8; Supplementary Table S3), and suggested sponge-species-specific bacterial communities. However, it seemed that such high specificity was not observed for the sponge-associated archaeal communities. Although compositional differences were found in the archaeal communities associated with *S. carteri* when compared with those in *H. erectus* and *X. testudinaria*, these differences did not distinguish *S. carteri*-associated archaeal communities from the other two sponges in the UniFrac cluster analysis, indicating that the communities shared a certain degree of similarity.

There is some consensus that sponge-species-specific microbial communities are maintained by vertical transmission of specific microbes across different generations of sponges, which is well supported by a number of studies covering sponges of different species or from different oceans (Sharp *et al.*, 2007; Schmitt *et al.*, 2008; Steger *et al.*, 2008; Lee *et al.*, 2009a). Under this scenario, one would believe that ancestral symbionts might have colonized a sponge host before sponge speciation and then coevolved with the host. As such, one would expect to observe similar symbiont communities among closely related sponges. Among the different sponge species investigated in this study, the microbial communities in *H. erectus* and *X. testudinaria* were more similar and diverse than those in *S. carteri* (Figures 1, 5, 6 and 8). All these three sponge species belong to Demospongiae (subclass Ceractinomorpha) but in different orders: *H. erectus* to Dictyoceratida, *S. carteri* to Halichondrida, *X. testudinaria* to Haplosclerida. Molecular evidence suggested a closer phylogenetic relationship between the orders Halichondrida and Haplosclerida (Borchiellini *et al.*, 2004), but this was not reflected in the associated microbial communities of *S. carteri* and *X. testudinaria*. The phylogeny of sponge, especially for Demospongiae, remains far from being clearly resolved (Boury-Esnault, 2006). Therefore, the associated microbes cannot serve as evolutionary biomarkers for sponges unless the sponge phylogeny is well resolved and the coevolution theory is well supported by unequivocal evidence.

Comparison with previous studies on sponge-associated microbial communities

Before the establishment of pyrosequencing technology, molecular techniques with lower resolution have recovered up to 20 bacterial phyla and both major archaeal lineages, Crenarchaeota and Euryarchaeota, from sponges (Taylor *et al.*, 2007; Hardoim *et al.*, 2009). The number of bacterial phyla reported in sponges was increased to 26 with the addition of Fusobacteria, Tenericutes, Deferribacteres, Fibrobacteres, BRC1 and WS3 by a study using tag pyrosequencing (Webster *et al.*, 2009). Our study probably further increased this number to 30 with the discovery of reads assigned to the phyla Chlorobi, Chrysiogenetes, Thermodesulfobacteria and the candidate phylum OD1 reported in sponges for the first time (Figure 5; Supplementary Table S5). Culturable bacteria for these phyla have been reported in lake, open water, arsenic-contaminated and high-temperature environments; some of them are related to sulfur or sulfate metabolism (Macy *et al.*, 1996; Garrity and Holt, 2001; Ludwig and

Klenk, 2001; Eisen, 2002). All these newly discovered phyla in sponges were recorded at very low abundance in this study (Supplementary Table S5), indicating that the tag pyrosequencing technology could successfully reveal rare microbial groups in sponges that might have otherwise been undetected by other molecular techniques. Future study using specific primers for PCR-mediated extension and subsequent phylogenetic analysis can further confirm the presence, quantity and novelty of these groups in sponges. Although rarefaction analysis indicated that further sequencing effort might be needed for certain samples to avoid underestimation of microbial diversity (Figures 2 and 3), our study has already revealed highly diverse microbial communities in sponges and supports the notion that diverse communities are largely composed of a wide variety of organisms in low abundance as shown by other studies (Hughes *et al.*, 2001).

The 16S amplicons generated in this study covered not only the V6 hypervariable region but also the V5 region, with an average amplicon length of up to 300 bp. When compared with the study on sponge-associated bacterial communities using the same technology in which 50–60 bp amplicons covering the V6 region were studied (Webster *et al.*, 2009), our study provided longer reads and allowed better taxonomic assignment and diversity analysis. In addition, our study was the first to use this high-throughput pyrosequencing technology to reveal the archaeal community in sponges. There have been limited studies on sponge-associated archaeal communities (Webster *et al.*, 2001, 2004; Lee *et al.*, 2003) and only 44 unique archaeal sequences have been recorded for sponges up to year 2007 (Taylor *et al.*, 2007). All these archaeal sequences, except one being closely related to Euryarchaeota, were affiliated with Crenarchaeota. Our study revealed up to 100 archaeal OTUs with an estimate of up to 300 archaeal species from a single sponge species (Figure 4; Supplementary Table S4). This is by far the highest diversity ever recorded for sponges. Similar to previous studies, most of our sponge-derived archaeal tags were affiliated with Crenarchaeota and a small proportion of them could be assigned to the sponge-specific cluster (Figures 6 and 7). Our study contributed to enrich a substantial number of Euryarchaeaote sequences for sponges. Furthermore, our study reported a non-negligible proportion of unclassified reads in sponges. They may represent novel species that have never been reported from other sponges or elsewhere and certainly deserve further examination.

Acknowledgements

We thank HC Chung, YH Wong, M Li and JP Ren from Hong Kong University of Science and Technology (HKUST), D Lau from Baptist University of Hong Kong and the technical team from King Abdullah University of Science and Technology (KAUST) for technical help during sample collection. We also thank Professor Rob von Soest, Zoological Museum, University of Amsterdam, for identification of sponges, and HC Chung, K Cheng and YK Tam, HKUST, for assistance in nutrient analysis. We are grateful to Professor MW Taylor, University of Auckland, for providing an ARB database containing the alignment of all sponge-derived sequences for comparison. This publication is based on the work supported by an award (SA-C0040/UK-C0016) made by KAUST to P-Y Qian.

References

Al-Farawati R, Al-Maradni A, Niaz R. (2008). Chemical characteristics (nutrients, fecal sterols and polyaromatic hydrocarbons) of the surface waters for Sharm Obhur, Jeddah, Eastern Coast of the Red Sea. *JKAU: Mar Sci* **19**: 95–119.

Al-Harbi SM, Khomayis HS. (2005). Seasonal variations of chlorophyll a and micro-nutrients in Obhur Creek, Red Sea. *JKAU: Mar Sci* **16**: 91–104.

Amann RI, Ludwig W, Schleifer KH. (1995). Phylogenetic identification and *in situ* detection of individual microbial cells without cultivation. *Microbiol Rev* **59**: 143–169.

Baker GC, Smith JJ, Cowan DA. (2003). Review and re-analysis of domain-specific 16S primers. *J Microb Methods* **55**: 541–555.

Binladen J, Gilbert MTP, Bollback JP, Panitz F, Bendixen C, Nielsen R *et al.* (2007). The use of coded PCR primers enables high-throughput sequencing of multiple homolog amplification products by 454 parallel sequencing. *PLoS One* **2**: e197.

Borchiellini C, Chombard C, Manuel M, Alivon E, Vacelet J, Boury-Esnault N. (2004). Molecular phylogeny of Demospongiae: implications for classification and scenarios of character evolution. *Mol Phyl Evol* **32**: 823–837.

Boury-Esnault N. (2006). Systematics and evolution of Demospongiae. *Can J Zool* **84**: 205–224.

Claesson MJ, O'Sullivan O, Wang Q, Nikkilä J, Marchesi JR, Smidt H *et al.* (2009). Comparative analysis of pyrosequencing and a phylogenetic microarray for exploring microbial community structure in the human distal intestine. *PLos One* **4**: e6669.

Chao A. (1984). Non-parametric estimation of the number of classes in a population. *Scand J Stat* **11**: 265–270.

Chao A, Lee SM. (1992). Estimating the number of classes via sample coverage. *J Ame Stat Assoc* **87**: 210–217.

Clarke KR. (1993). Non-parametric multivariate analyses of changes in community structure. *Aust J Ecol* **18**: 117–143.

Clarke KR, Warwick RM. (1994). *Change in Marine Communities: An Approach to Statistical Analysis and Interpretation*. Plymouth Marine Laboratory: Plymouth, United Kingdom.

Cole JR, Wang Q, Cardenas E, Fish J, Chai B, Farris RJ *et al.* (2008). The Ribosomal Database Project: improved alignments and new tools for rRNA analysis. *Nucl Acids Res* **37**: 141–145.

Edgar RC. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucl Acids Res* **32**: 1792–1797.

Eilers H, Pernthaler J, Gloeckner FO, Amann R. (2000). Culturability and *in situ* abundance of pelagic bacteria the north sea. *Appl Environ Microbiol* **66**: 3044–3051.

Eisen JA. (2002). The complete genome sequence of the green sulfur bacterium *Chlorobium tepidum*. *Proc Natl Acad Sci USA* **99**: 9509–9514.

Friedrich AB, Fischer I, Proksch P, Hacker J, Hentschel U. (2001). Temporal variation of the microbial community associated with the Mediterranean sponge *Aplysina aerophoba*. *FEMS Microbiol Ecol* **38**: 105–113.

Garrity GM, Holt JG. (2001). Phylum BVII. Thermomicrobia phy. nov. In: Boone DR, Castenholz RW, Garrity GM (eds). *Bergey's Manual of Systematic Bacteriology*, *Vol. 1, The Archaea and the Deeply Branching and Phototrophic Bacteria*, 2nd edn. Springer: New York. pp 447.

Garson MG, Flowers AE, Webb RI, Charan RD, McCaffrey EJ. (1998). A sponge/dinoflagellate association in the haposclerid sponge *Haliclona* sp.: cellular origin of cytotoxic alkaloids by Percoll density gradient fractionation. *Cell Tissue Res* **293**: 365–373.

Ghiglione JF, Palacios C, Marty JC, Mével G, Labrune C, Conan P *et al.* (2008). Role of environmental factors for the vertical distribution (0–1000 m) of marine bacterial communities in the NW Mediterranean Sea. *Biogeosciences* **5**: 1751–1764.

Giovannoni SJ, Stingl U. (2005). Molecular diversity and ecology of microbial plankton. *Nature* **437**: 343–348.

Hardoim CC, Costa R, Araujo FV, Hajdu E, Peixoto R, Lins U *et al.* (2009). Diversity of bacteria in the marine sponge *Aplysina fulva* in Brazilian coastal waters. *Appl Environ Microbiol* **75**: 3331–3343.

Hentschel U, Fieseler L, Wehrl M, Gernert C, Steinert M, Hacker J *et al.* (2003). Microbial diversity of marine sponges. In: Müller WEG (ed). *Sponges (Porifera)*. Springer: Heidelberg, Germany. pp 59–83.

Hentschel U, Hopke J, Horn M, Friedrich AB, Wagner M, Hacker J *et al.* (2002). Molecular evidence for a uniform microbial community in sponges from different oceans. *Appl Environ Microbiol* **68**: 4431–4440.

Hentschel U, Schmid M, Wanger M, Fieseler L, Gernert C, Hacker J. (2001). Isolation and phylogenetic analysis of bacteria with antimicrobial activities from the Mediterranean sponges *Aplysina aerophoba* and *Aplysina cavernicola*. *FEMS Microb Ecol* **35**: 305–312.

Hughes JB, Hellmann JJ, Ricketts TH, Bohannan BJ. (2001). Counting the uncountable: statistical approaches to estimating microbial diversity. *Appl Environ Microbiol* **67**: 4399–4406.

Ilan M, Gugel J, Van Soest RWM. (2004). Taxonomy, reproduction and ecology of new and known Red Sea sponges. *Sarsia* **89**: 388–410.

Joachimiak MP, Weisman JL, May BCH. (2006). JColorGrid: software for the visualization of biological measurements. *BMC Bioinformatics* **7**: 225–229.

Katoh T. (2008). Recent developments in the MAFFT multiple sequence alignment program. *Briefings in Bioinformatics* **9**: 286–298.

Lee EY, Lee HK, Lee YK, Sim CK, Lee JH. (2003). Diversity of symbiotic archaeal communities in marine sponges from Korea. *Biomol Eng* **20**: 299–304.

Lee OO, Chui PY, Wong YH, Pawlik JR, Qian PY. (2009a). Evidence for vertical transmission of bacterial symbionts from adult to embryo in the Caribbean sponge *Svenzea zeai*. *Appl Envioron Microbiol* **75**: 6147–6156.

Lee OO, Wong YH, Qian PY. (2009b). Inter- and intraspecific variations of bacterial communities associated with marine sponges from San Juan Island, Washington. *Appl Environ Microbiol* **75**: 3513–3521.

Lee OO, Yang LH, Li X, Pawlik JR, Qian PY. (2007). Surface bacterial community, fatty acid profile, and antifouling activity of two congeneric sponges from Hong Kong and the Bahamas. *Mar Ecol Prog Ser* **339**: 25–40.

Liu WT, Marsh TL, Cheng H, Forney LJ. (1997). Characterization of microbial diversity by determining terminal restriction fragment length polymorphisms of genes encoding 16S rRNA. *Appl Environ Microbiol* **63**: 4516–4522.

Liu Z, DeSantis TZ, Andersen GL, Knight R. (2008). Accurate taxonomy assignments from 16S rRNA sequences produced by highly parallel pyrosequencers. *Nucl Acids Res* **36**: e120.

Lozupone C, Hamady M, Knight R. (2006). UniFrac—An online tool for comparing microbial community diversity in a phylogenetic context. *BMC Bioinformatics* **7**: 371.

Lozupone C, Knight R. (2005). UniFrac: a new phylogenetic method for comparing microbial communities. *Appl Environ Microbiol* **71**: 8228–8235.

Ludwig W, Klenk HP. (2001). Overview: a phylogenetic backbone and taxonomic framework for prokaryotic systematics. In: Boone DR, Castenholz RW (eds). *Bergey's Manual of Systematic Bacteriology*. Springer-Verlag: Berlin. pp 49–65.

Macy JM, Nunan K, Hagen KD, Dixon DR, Harbour PJ, Cahill M *et al.* (1996). *Chrysiogenes arsenatis* gen. nov., sp. nov., a new arsenate-respiring bacterium isolated from gold mine wastewater. *Int J Syst Bacteriol* **46**: 1153–1157.

Maldonado M, Cortadellas N, Trillas MI, Rützler K. (2005). Endosymbiotic yeast maternally transmitted in a marine sponge. *Biol Bull* **209**: 94–106.

Margulies M, Egholm M, Altman WE, Attiya S, Bader JS, Bemben LA *et al.* (2005). Genome sequencing in microfabricated high-density picolitre reactors. *Nature* **437**: 376–380.

Martiny JB, Bohannan BJ, Brown JH, Colwell RK, Fuhrman JA, Green JL *et al.* (2006). Microbial biogeography: putting microorganisms on the map. *Nat Rev Microbiol* **4**: 102–112.

Morris RM, Rappe MS, Connon SA, Vergin KL, Siebold WA, Carlson CA *et al.* (2002). SAR11 clade dominates ocean surface bacterioplankton communities. *Nature* **420**: 806–810.

Müller WEG, Müller IM. (2003). Origin of the metazoan immune system: identification of the molecules and their functions in sponges. *Integr Comp Biol* **43**: 281–292.

Oren M, Steindler L, Ilan M. (2005). Transmission, plasticity and the molecular identification of cyanobacterial symbionts in the Red Sea sponge *Diacarnus erythraenus*. *Mar Biol* **148**: 35–41.

Pommier T, Canback B, Riemann L, Bostrom KH, Simu K, Lundberg P *et al.* (2007). Global patterns of diversity and community structure in marine bacterioplankton. *Mol Ecol* **16**: 867–880.

Preston CM, Wu KY, Molinski TF, DeLong EF. (1996). A psychrophilic crenarchaeon inhabits a marine sponge: *Cenarchaeum symbiosum* gen. nov., sp. nov. *Proc Natl Acad Sci USA* **93**: 6241–6246.

Qian PY, Dobretsov S, Dahms HU, Pawlik J. (2006). Antifouling activity and microbial diversity of two congeneric sponges, *Callyspongia* spp. from Hong

Kong and the Bahamas. *Mar Ecol Prog Ser* **324**: 151–165.

Radwan M, Hanora A, Zan J, Mohamed NM, Abo-Elmatty DM, Abou-El-Ela SH *et al.* (2010). Bacterial community analyses of two Red Sea Sponges. *Mar Biotechnol* **12**: 350–360.

Santavy DL, Willenz P, Colwell RR. (1990). Phenotypic study of bacteria associated with the Caribbean sclerosponge, *Ceratoporella nicholsoni*. *Appl Environ Microbiol* **56**: 1750–1762.

Schloss PD, Handelsman J. (2005). Introducing DOTUR, a computer program for defining operational taxonomic units and estimating species richness. *Appl Environ Microbiol* **71**: 1501–1506.

Schmitt S, Angermeier H, Schiller R, Lindquist N, Hentschel U. (2008). Molecular microbial diversity survey of sponge reproductive stages and mechanistic insights into vertical transmission of microbial symbionts. *Appl Environ Microbiol* **74**: 7694–7708.

Sharp KH, Eam B, Faulkner DJ, Haygood MG. (2007). Vertical transmission of diverse microbes in the tropical sponge *Corticium* sp. *Appl Environ Microbiol* **73**: 622–629.

Simpson TL. (1984). *The Cell Biology of Sponges*. Springer-Verlag: New York, NY, USA.

Sogin ML, Morrison HG, Huber JA, Welch DM, Huse SM, Neal PR *et al.* (2006). Microbial diversity in the deep sea and the underexplored 'rare biosphere'. *Proc Natl Acad Sci USA* **103**: 12115–12120.

Steger D, Ettinger-Epstein P, Whalan S, Hentschel U, de Nys R, Wagner M *et al.* (2008). Diversity and mode of transmission of ammonia-oxidizing archaea in marine sponges. *Environ Microbiol* **10**: 1087–1094.

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

Taylor MW, Schupp PJ, Dahllof I, Kjelleberg S, Steinberg PD. (2004). Host specificity in marine sponge-associated bacteria, and potential implications for marine microbial diversity. *Environ Microbiol* **6**: 121–130.

Taylor MW, Schupp PJ, de Nys R, Kjelleberg S, Steinberg PD. (2005). Biogeography of bacteria associated with the marine sponge *Cymbastela concentric*. *Environ Microbiol* **7**: 419–433.

Thacker RW, Starnes S. (2003). Host specificity of the symbiotic cyanobacterium *Oscillatoria spongeliae* in marine sponge, *Dysides* spp. *Mar Biol* **142**: 643–648.

Vacelet J. (1982). Algal-sponge symbioses in the coral reef of New Caledonia: a morphological study. In: Gomez ED, Birkeland CE (eds). *Proceedings of the 4th International Coral Reef Symposium*, Vol. 2 Marine Sciences Center, University of the Philippines: Manila. pp 713–719.

Vacelet J, Donadey C. (1977). Electron microscope study of the association between some sponges and bacteria. *J Exp Mar Biol Ecol* **30**: 301–314.

Venter J, Remington K, Heidelberg J, Halpern A, Rusch D, Eisen J *et al.* (2004). Environmental genome shotgun sequencing of the Sargasso Sea. *Science* **304**: 66–74.

Vogel G. (2008). The inner lives of sponges. *Science* **320**: 1028–1030.

Wang Y, Qian PY. (2009). Conservative fragments in bacterial 16S rRNA genes and primer design for 16S ribosomal DNA amplicons in metagenomic studies. *PLoS ONE* **4**: e7401.

Webster NS, Cobb RE, Negri AP. (2008). Temperature thresholds for bacterial symbiosis with a sponge. *ISME J* **2**: 830–842.

Webster NS, Hill RT. (2001). The culturable microbial community of the Great Barrier Reef sponge *Rhopaloeides odorabile* is dominated by α-proteobacterium. *Mar Biol* **138**: 843–851.

Webster NS, Negri AP, Munro MMHG, Battershill CN. (2004). Diverse microbial communities inhabit Antarctic sponges. *Environ Microbiol* **6**: 288–300.

Webster NS, Taylor MW, Behnam F, Lücker S, Rattei T, Whalan S *et al.* (2010). Deep sequencing reveals exceptional diversity and modes of transmission for bacterial sponge symbionts. *Environ Microb* **12**: 2070–2082.

Webster NS, Watts JEM, Hill RT. (2001). Detection and phylogenetic analysis of novel Crenarchaeote and Euryarchaeote 16S ribosomal RNA gene sequences from a Great Barrier Reef sponge. *Mar Biotech* **3**: 600–608.

Wichels A, Wurtz S, Dopke H, Schutt C, Gerdts G. (2006). Bacterial diversity in the breadcrumb sponge *Halichondria panicea* (Pallas). *FEMS Microbiol Ecol* **56**: 102–118.

Wilkinson CR. (1984). Immunological evidence for the Precambrian origin of bacterial symbioses in marine sponges. *Proc R Soc Lond B* **220**: 509–517.

Zhu P, Li Q, Wang G. (2008). Unique microbial signatures of the alien Hawaiian marine sponge *Suberites zeteki*. *Microb Ecol* **55**: 406–414.

Supplementary Information accompanies the paper on The ISME Journal website (http://www.nature.com/ismej)