Habitat- and host-related variation in sponge bacterial symbiont communities in Indonesian waters

Daniel F.R. Cleary , Leontine E. Becking , Nicole J. de Voogd , Ana C.C. Pires , Ana R.M. Polónia , Conceição Egas , Newton C.M. Gomes

RESEARCH ARTICLE

Habitat- and host-related variation in sponge bacterial symbiont communities in Indonesian waters

Daniel F.R. Cleary[1], Leontine E. Becking[2,3], Nicole J. de Voogd[2], Ana C.C. Pires[1], Ana R.M. Polónia[1], Conceição Egas[4] & Newton C.M. Gomes[1]

[1]Departamento de Biologia, CESAM – Centro de Estudos do Ambiente e do Mar, Universidade de Aveiro, Aveiro, Portugal; [2]Naturalis Biodiversity Center, Leiden, The Netherlands; [3]Institute for Marine Resources and Ecosystem Studies (IMARES), Wageningen UR, Den Helder, The Netherlands; and [4]Biocant-Biotechnology Innovation Center, Cantanhede, Portugal

Correspondence: Daniel F.R. Cleary, CESAM and Department of Biology, Universidade de Aveiro, Campus Universitário de Santiago, 3810-193 Aveiro, Portugal. Tel.: +351 234 370990; fax: +351 234 370004; e-mails: cleary@ua.pt or dfrcleary@gmail.com

Received 14 December 2013; revised 10 April 2013; accepted 16 April 2013.
Final version published online 30 May 2013.

DOI: 10.1111/1574-6941.12135

Editor: Julian Marchesi

Keywords
Borneo; community composition; mangroves; marine lakes; pyrosequencing; *Porifera*.

Abstract

Marine lakes are unique ecosystems that contain isolated populations of marine organisms. Isolated from the surrounding marine habitat, many lakes house numerous endemic species. In this study, microbial communities of sponges inhabiting these lakes were investigated for the first time using barcoded pyrosequencing of 16S rRNA gene amplicons. Our main goals were to compare the bacterial richness and composition of two sponge species (*Suberites diversicolor* and *Cinachyrella australiensis*) inhabiting both marine lakes and adjacent open coastal systems. Host species and habitat explained almost 59% of the variation in bacterial composition. There was a significant difference in composition between both host species. Within *S. diversicolor*, there was little discernible difference between bacterial communities inside and outside lakes. The bacterial community of this species was, furthermore, dominated (63% of all sequences) by three very closely related alphaproteobacterial taxa identified as belonging to the recently described order *Kiloniellales*. *Cinachyrella australiensis*, in contrast, hosted markedly different bacterial communities inside and outside lakes with very few shared abundant taxa. *Cinachyrella australiensis* in open habitat only shared 9.4% of OTUs with *C. australiensis* in lake habitat. *Bacteria* were thus both highly species specific and, in the case of *C. australiensis*, habitat specific.

Introduction

Marine lakes, a very rare and unique habitat, are anchialine systems, which are small bodies of landlocked sea water isolated to varying degrees from the surrounding marine environment (Holthuis, 1973; Hamner & Hamner, 1998; Colin, 2009; Becking *et al.*, 2011); they contain brackish to almost fully marine waters. The marine character of these systems is maintained by subterranean tunnels, fissures or small dissolution channels in the surrounding rock, connecting the lakes to the adjacent sea, and as such display a wide variety in the degree of connection to the sea and environmental regimes within the lakes (Hamner & Hamner, 1998; Cerrano *et al.*, 2006; Azzini *et al.*, 2007; Becking *et al.*, 2011). These landlocked pools of water are subjected to a tidal regime, which is typically delayed and dampened compared with the adjacent sea (Hamner & Hamner, 1998; Becking *et al.*, 2011). The number of marine lakes worldwide is estimated at only 200 with clusters of ten or more lakes occurring in areas with a karstic limestone landscape in Croatia, Bermuda, Vietnam, Palau and Indonesia (Dawson *et al.*, 2009). This enclosed environment has set the stage for small, isolated, rapidly evolving populations and endemic (sub)species (Holthuis, 1973; Maciolek, 1983; Tomascik & Mah, 1994; Massin & Tomascik, 1996; Dawson & Hamner, 2005). Kakaban lake, for example, one of the largest marine lakes presently known to science and highly isolated from the adjacent sea (Tomascik & Mah, 1994; Becking *et al.*, 2011), contains many rare and endemic species across a variety of taxa including a crab (*Orcovita saltatrix* Ng & Tomascik, 1994), two holothurians (*Holothuria* (*Lessonothuria*) *cavans* Massin & Tomascik, 1996 and *Synaptula spinifera* Massin & Tomascik,

1996) and an ascidian (*Styela complexa* Kott, 1995). Surveys of marine lakes in Indonesia, Vietnam and Palau have shown sponges to be one of the most dominant taxa in terms of biomass and diversity (Azzini *et al.*, 2007; Colin, 2009; Becking *et al.*, 2011, 2013). However, to the best of our knowledge, no studies have been conducted on the diversity and composition of symbiont microbial communities associated with these unique and unexploited environments.

Sponges (phylum *Porifera*) are exclusively aquatic animals with currently 8553 extant species and with an estimated 25 000 species (Appeltans *et al.*, 2012; Van Soest *et al.*, 2012). They are successful colonisers of a wide range of habitats, from tropical to polar seas and shallow to deep waters, and are found in marine and freshwater habitats, where they are involved in a host of ecological processes (Rützler, 2004). Sponges have been shown to be unique and highly selective environments for bacteria; the bacterial assemblages they host are, furthermore, of substantial ecological, biotechnological and pharmaceutical importance (Hentschel *et al.*, 2002, 2003, 2006; Taylor *et al.*, 2007; Webster *et al.*, 2010; Jackson *et al.*, 2012; Webster & Taylor, 2012). In many cases, the bacterial symbionts either are the source or contribute significantly to the production of bio-active secondary metabolites found in sponges (Lee *et al.*, 2001; Piel, 2004; Erpenbeck & van Soest, 2007; Taylor *et al.*, 2007). As a result, there is heightened interest in these bacteria. Different sponge species cohabiting the same habitat can greatly differ in the abundance of their associated microorganisms. High-microbial-abundance sponges can contain around 10^{10} bacterial cells g^{-1} wet weight of sponge (orders of magnitude higher than concentrations in sea water); low-microbial abundance sponges contain densities of around 10^6 cells g^{-1} (similar to densities in sea water) (Hentschel *et al.*, 2006; Kamke *et al.*, 2010). Only a minute percentage of bacteria found in sponges are cultivable (Hentschel *et al.*, 2003; Jackson *et al.*, 2012). Recent advances, however, in molecular techniques, such as pyrosequencing, now enable us to assess bacterial communities at an unprecedented level of detail (Webster *et al.*, 2010; Lee *et al.*, 2011; Jackson *et al.*, 2012; Schmitt *et al.*, 2012b; White *et al.*, 2012).

In the present study, we compare the richness and composition of bacteria in two sponge species inhabiting enclosed marine lakes and surrounding open coastal habitat in the islands of Kakaban and Maratua in the Berau Delta barrier reef system. The species selected were *Suberites diversicolor* (Becking & Lim, 2009) *(Demospongiae: Hadromerida: Suberitidae)* and *Cinachyrella australiensis* (Carter, 1886) *(Demospongiae: Spirophorida: Tetillidae)*. These sponges were selected because they were relatively abundant inside and outside marine lakes. Specimens of both species were sampled in marine lakes located in the islands of Kakaban and Maratua in the Berau Delta barrier reef system, East Kalimantan, Indonesia (Fig. 1). In addition to this, we also sampled sponges from marine shallow open water habitat surrounding both islands.

Our specific goals were to (1) compare bacterial richness between host species and among habitats; (2) identify the most abundant higher bacterial taxa; (3) assess to what extent host species and habitat (marine lake vs. open water) structure sponge bacterial composition; and (4) identify dominant ($\geq$ 500 sequences) bacterial OTUs and their closest known relatives using BLAST and assess the relationships among taxa using phylogenetic analysis.

Materials and methods

Sampling

Sampling was performed by snorkelling and SCUBA diving from 10 August to 10 September 2008 inside and outside (Opn) the marine lakes of Berau, East Kalimantan Province, Indonesia (Fig. 1). The Berau Delta and barrier reef system in East Kalimantan, Indonesia, is an intricate coastal system with a variety of coastal landforms and associated ecosystems such as coral reefs and mangroves. The offshore islands of the barrier reef system, Kakaban and Maratua, contain marine lakes. Kakaban lake (Kak), a large marine lake of *c.* 4 km^2 fringed by mangroves, is located in Kakaban island (N02°08′57.3″ E118°31′26.4″). Kakaban lake is one of the largest marine lakes presently known and strongly disconnected from the adjacent sea (Tomascik & Mah, 1994; Becking *et al.*, 2011). Haji Buang Lake (Mbu), a smaller marine lake of *c.* 0.14 km^2, is located in Maratua Island (N02°12′31.2″ E118°35′46.8″). A detailed description of the lakes of Kakaban and Maratua is provided by Tomascik & Mah (1994), Tomascik *et al.* (1997) and Becking *et al.* (2011).

Cores of the sponge species *C. australiensis (Demospongiae: Spirophorida: Tetillidae)* and *S. diversicolor (Demospongiae: Hadromerida: Suberitidae)* were sampled including segments of surface and interior to sample, as much as possible, the whole bacterial community; these were stored in 96% EtOH for microbial analysis. All specimens were collected from shallow water habitat (< 5 m depth). Specimens were identified to species by LE Becking and NJ de Voogd. Voucher specimens were preserved in 70% EtOH and deposited in the sponge collection of the Naturalis Biodiversity Center (RMNH Porifera). In the Berau region, *S. diversicolor* occurs predominantly in marine lakes, but can also be found in sheltered areas outside the marine lakes. The species has also been found in Singapore and Northern Australia in sheltered habitats. The external coloration of *S. diversicol-*

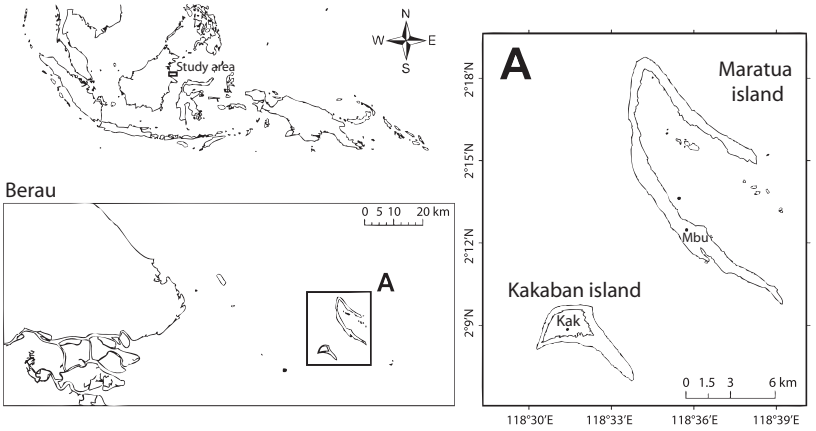

Fig. 1. Map of study area showing Indonesia in the upper left inset. The bottom left and right insets show the location of the marine lakes sampled during this study: Kakaban lake (Kak) and Haji Buang lake (Mbu).

or varies greatly between and within localities from green to red, while being bright to dark yellow internally. The variable external coloration is thought to be due to the presence of photosynthesising symbionts (Becking & Lim, 2009). *Cinachyrella australiensis* is a globular yellow sponge that occurs in a great variety of habitats across the Indo-Pacific. Both species tolerate and thrive in extreme or perturbed environments including marine lakes and intertidal areas subject to fluctuations in salinity, sediment loads and exposure to air during low tide (McDonald *et al.*, 2002; Becking & Lim, 2009; de Voogd *et al.*, 2009; Becking *et al.*, 2013). Four specimens were collected per location; both species were collected from the marine lakes in Kakaban and Haji Buang, and four additional samples were collected from open waters surrounding the islands of Kakaban and Maratua (Fig. 1).

DNA extraction and pyrosequencing

Based on the work of Hardoim *et al.* (2009), genomic DNA was extracted using 0.5 g of tissue from each individual. After cutting the tissue samples into small slices, DNA was extracted with the FastDNA® Spin Kit for Soil (MP Biomedicals). To optimise DNA extraction, we followed the standard MP Biomedicals protocol. After DNA extraction, the community 16S rRNA gene was amplified for the V3V4 hypervariable region with barcoded fusion primers containing the Roche-454 A and B Titanium sequencing adapters, an eight-base barcode sequence in adaptor A and specific sequences for the ribosomal region.

Two replicate PCRs were performed for each sample using the primer pair V3F (5′– ACTCCTACGGGAGGC AG-3′) and V4R (5′– TACNVRRGTHTCTAATYC-3′) (Wang & Qian, 2009), 1X Advantage 2 Polymerase Mix (Clontech, Mountain View, CA), 1X Advantage 2 PCR Buffer, 0.2 μM of each PCR primer, 0.2 mM dNTPs (Bioron, Ludwigshafen am Rhein, Germany), 5% DMSO (Roche Diagnostics GmbH, Mannheim, Germany) and 2 μL of genomic DNA template in a total volume of 25 μL. The PCR conditions involved a 4-min denaturation at 94°C, followed by 30 cycles of 94 °C for 30 s, 44 °C for 45 s and 68 °C for 60 s and a final extension at 68 °C for 10 min. Negative controls were included for all amplification reactions. Electrophoresis of duplicate PCR products was undertaken on a 1% (w/v) agarose gel, and the 470-bp amplified fragments were purified using AMPure XP beads (Agencourt, Beckman Coulter) or, if more than the expected fragment was amplified, gel-purified using High-Pure PCR Product Purification Kit (Roche Diagnostics GmbH), according to manufacturer's instructions. The amplicons were quantified by fluorimetry with PicoGreen dsDNA quantitation kit (Invitrogen, Life Technologies, Carlsbad, CA), pooled at equimolar concentrations and sequenced in the A direction with GS 454 FLX Titanium chemistry, according to manufacturer's instructions (Roche, 454 Life Sciences, Branford, CT) at Biocant (Cantanhede, Portugal). The sequences generated in this study can be downloaded from the NCBI Short Read Archive, accession number: SRA049887.1.

Sequence analyses of 16S rRNA gene fragments

In this study, the barcoded pyrosequencing libraries were analysed using the Quantitative Insights Into Microbial Ecology (QIIME; Caporaso *et al.*, 2010) software package (http://www.qiime.org/, last accessed 19 November 2012) on a computer running the BioLinux operating system (http://nebc.nerc.ac.uk/tools/bio-linux/bio-linux-6.0, last accessed 19 November 2012). In QIIME, fasta and qual files were used as input for the split_libraries.py script. Default arguments were used except for the minimum sequence length, which was set at 218 bp after removal of forward primers and barcodes, backward primers were removed using the 'truncate only' argument, and a sliding

window test of quality scores was enabled with a value of 50 as suggested in the QIIME description for the script. In addition to user-defined cut-offs, the split_libraries.py script performs several quality-filtering steps (http://qiime.org/scripts/split_libraries.html). OTUs were selected using the pick_otus.py script in QIIME with the usearch_ref method, default sequence similarity threshold of 0.97 and minimum cluster size of 1, and OTUs were selected using the most recent Greengenes release (Greengenes 12_10; http://qiime.wordpress.com/2012/10/16/greengenes-12_10-is-released/) as reference database. Reference-based OTU picking using the 12_10 release led to a large increase in the number of reads assigned to the reference database of soil and human microbiome data when compared to an earlier Greengenes release (release 4feb2011; http://qiime.wordpress.com/, last accessed 19 November 2012). The usearch sequence analysis tool (Edgar, 2010) implemented in QIIME provides clustering, chimera checking and quality filtering on demultiplexed sequences. Chimera checking was performed using the UCHIME algorithm, which is the fastest and most sensitive chimera-checking algorithm currently available (Edgar *et al.*, 2011). In the present study, we used de novo checking and reference-based chimera checking using a reference fasta file ('99_otus.fasta') from the Greengenes 12_10 release. The quality filtering as implemented in usearch filters noisy reads and preliminary results suggest it gives results comparable to other denoisers such as AmpliconNoise, but is much less computationally expensive (http://drive5.com/usearch/features.html, last accessed 19 November 2012). Representative sequences were selected using the pick_rep_set.py script in QIIME using the 'most_abundant' method. Reference sequences of OTUs were assigned taxonomies using default arguments in the assign_taxonomy.py script in QIIME with the rdp method (Wang *et al.*, 2007). In the assign_taxonomy.py function, we used a fasta file containing reference sequences from the Greengenes 12_10 release as training sequences for the rdp classifier. We used a modified version of the taxonomy file supplied with the Greengenes 12_10 release to map sequences onto the assigned taxonomy. Finally, we used the make_otu_table.py script in QIIME to generate a square matrix of OTUs by samples. This was subsequently used as input for further analyses using the R package (http://www.r-project.org/, last accessed 19 November 2012).

Statistical analysis

A square matrix containing the abundance of all OTUs per sample was imported into R using the read.table() function. Plant organelles, mitochondria and sequences not classified as Bacteria (e.g. *Archaea*) were removed prior to statistical analysis. Samples with < 100 sequences were also removed prior to analysis. After importing into R, we used a self-written function (Gomes *et al.*, 2010) to estimate total rarefied OTU richness for pooled samples belonging to each sponge host in each habitat.

The OTU abundance matrix was $\log_{10}(x + 1)$-transformed (to normalise the distribution of the data), and a distance matrix was constructed using the Bray–Curtis index with the vegdist() function in the VEGAN package (Oksanen *et al.*, 2009) in R. The Bray–Curtis index is one of the most frequently applied (dis)similarity indices used in ecology (Legendre & Gallagher, 2001; Cleary, 2003; Cleary & Genner, 2004; Cleary *et al.*, 2004; Becking *et al.*, 2006; de Voogd *et al.*, 2009). Variation in sponge composition among habitats (Opn, Kak and Mbu) and sponge hosts (*C. australiensis* and *S. diversicolor*) was assessed with principal coordinates analysis (PCO) using the cmdscale() function in R with the Bray–Curtis distance matrix as input. Variation among habitats and sponge hosts was tested for significance using the adonis() function in vegan. In the adonis analysis, the Bray–Curtis distance matrix of species composition was the response variable with habitat and sponge host as independent variables; the strata argument was set to habitat so that randomisations were constrained to occur within each habitat and not across all habitats. The number of permutations was set at 999; all other arguments used the default values set in the function. Weighted averages scores were computed for OTUs on the first two PCO axes using the wascores() function in the VEGAN package.

Description of phylogenetic analysis

Sequence identifiers of closely related taxa of numerically dominant OTUs ($\geq$ 500 sequences) were downloaded using the NCBI Basic Local Alignment Search Tool (BLAST) command line 'BLASTN' tool with the -db argument set to nt (Zhang *et al.*, 2000). BLAST identifies locally similar regions between sequences, compares sequences to extant databases and assesses the significance of matches; functional and evolutionary relationships can subsequently be inferred. Each run produces a list of hits based on significant similarity between pairs of sequences, that is, the target sequence and taxa present in the database (or no hits if no significantly similar sequences are found). A discussion of how significance is determined can be found at http://www.ncbi.nlm.nih.gov/BLAST/tutorial/Altschul-1.html. We used the BLASTN command line tool in a Linux environment to query representative sequences of selected taxa including all the dominant OTUs ($\geq$ 500 sequences) against the online NCBI nucleotide database. We then generated a vector containing

sequence identifiers (GIs) of the 10 top hits of all representative sequences and used the Entrez.efetch function in BioPython (Cock *et al.*, 2009) with the rettype argument set to 'gb' to download GenBank information of aforementioned top hits including the isolation source of the organism and the host. From the list of hits, we selected taxa with the highest maximum sequence identity score (and total score) to our target representative sequences and included these in a phylogenetic analysis of the dominant OTUs. Fasta files of the closely related organisms identified with BLAST were downloaded using the Entrez.efetch function with the rettype argument set to 'fasta'.

We constructed a phylogenetic tree including all dominant taxa ($\geq$ 500 sequences), some additional taxa, for example the most dominant poribacterial OTU, and their closest relatives identified using BLAST as previously described. The phylogenetic tree was built using the MEGA5 program (http://www.megasoftware.net/, last accessed 20 November 2012; Tamura *et al.*, 2011) with the neighbour-joining method (Saitou & Nei, 1987) based on the Kimura 2-parameter method (Kimura, 1980). In the results, we present a bootstrap consensus tree based on 1000 replicates (Felsenstein, 1985). Branches reproduced in < 50% of the bootstrap replicates are collapsed. The bootstrap value is shown next to each branch when this exceeds 50%. This value represents the percentage of replicate trees in which the associated taxa clustered together. In the results, the tree presented is drawn to scale; branch lengths are measured in the number of substitutions per site. Fifty-nine nucleotide sequences were involved in the analysis. Codon positions included were 1st + 2nd + 3rd + Noncoding. Positions with gaps and missing data were eliminated. There were a total of 1569 positions in the final data set.

Results

The sequencing effort yielded 53 683 sequences, which were assigned to 3064 OTUs after quality control, OTU picking and removal of chimera (Tables S1 and S2, Supporting Information). The assign taxonomy script, however, failed to assign 890 OTUs to the kingdom '*Bacteria*'; 47 OTUs were identified as either chloroplasts or mitochondria. After removal of these OTUs, chloroplasts and mitochondria, our sequencing effort yielded 50 892 sequence reads and 2127 OTUs. OTUs were assigned to a total of 29 phyla. These included *Proteobacteria* (975 OTUs), *Bacteroidetes* (104), *Actinobacteria* (71), *Firmicutes* (60), *Chloroflexi* (52), *Cyanobacteria* (37), *Acidobacteria* (26), *Gemmatimonadetes* (17), *Spirochaetes* (13), *Nitrospirae* (8), *ZB3* (7), *Verrucomicrobia* (7), *Chlamydiae* (5) and *Poribacteria* (4). 716 OTUs remained unclassified at the phylum level. 773 OTUs from 14 150 sequence reads were identified from *S. diversicolor* hosts, while 1504 OTUs from 36 742 sequence reads were identified from *C. australiensis* hosts. A total of 22 OTUs were dominant, that is, were represented by more than $\geq$ 500 sequences. The most dominant OTU overall was OTU-1733, an *Alphabacterium* only found in *S. diversicolor* hosts and represented by 5631 sequences. As can be seen in the heat map (Fig. 2), the distribution of dominant OTUs suggests pronounced clustering. Taxa were largely restricted to *S. diversicolor* hosts and *C. australiensis* hosts in open vs. lake habitat. Only one OTU was found in every sponge, OTU-218.

OTU richness

Rarefied bacterial OTU richness was highest in *C. australiensis* and *S. diversicolor* hosts in Haji Buang lake, Maratua, intermediate in open water habitat and lowest in Kakaban lake. In each habitat, *C. australiensis* harboured more diverse bacterial assemblages than *S. diversicolor* (Fig. 3). Despite this, habitat appears to be a better predictor of bacterial richness than host. OTU richness approached 800 OTUs when more than 20 000 sequences were sampled for the most abundant host–habitat combination, namely *C. australiensis* hosts in open habitat. There was no evidence of an asymptote for any host–habitat combination, which indicates that true richness is higher than that reported here.

Higher taxon abundance

There were marked differences in the abundance of higher bacterial taxa between host species and among habitats (Fig. 4). The *Acidimicrobidae*, *Gammaproteobacteria* and *Deltaproteobacteria* were markedly more abundant in *C. australiensis* than in *S. diversicolor* hosts. In contrast, *Alphaproteobacteria* were more abundant in *S. diversicolor* than in *C. australiensis* hosts. This was, however, largely due to the dominance of taxa belonging to the *Kiloniellales* order in *S. diversicolor* hosts. *Alphaproteobacteria* in *C. australiensis* hosts were more abundant in lake habitat than in open habitat. Taxa belonging to the *Anaerolineae* and *Methylococcales* were largely restricted to *C. australiensis* hosts in open habitat. Of the 1061 sequences identified as *Anaerolineae*, only three were not found in *C. australiensis* hosts in open habitat. Of the 6402 sequences identified as *Methylococcales*, only two (both from a single specimen of *C. australiensis* in Haji Buang Lake) were not found in *C. australiensis* hosts in open habitat. *Betaproteobacteria* were largely restricted to *C. australiensis* hosts in lake habitat. Of the 1149 sequences identified as *Betaproteobacteria*, only two (both

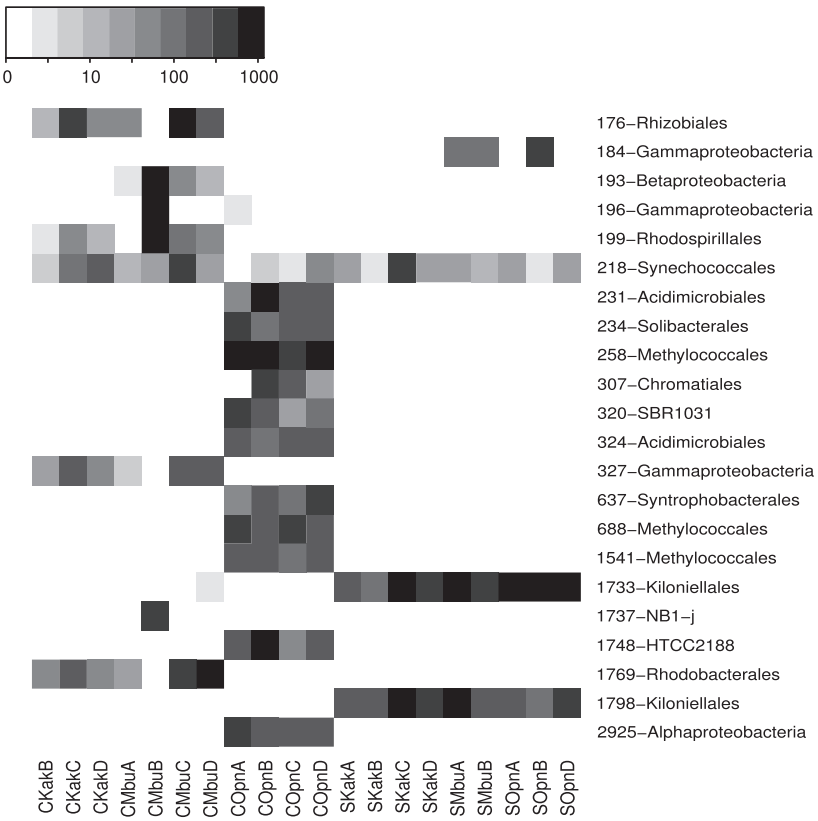

Fig. 2. Heat map showing the abundance of 16S rRNA gene sequence reads of bacterial OTUs with $\geq$ 500 sequence reads. All samples are indicated along the *x*-axis: CKak: *Cinachyrella* in Kakaban, CMbu: *Cinachyrella* in Haji Buang lake, Maratua, COpn: *Cinachyrella* in open habitat, SKak: *Suberites diversicolor* in lake Kakaban, SMbu: *S. diversicolor* in Haji Buang lake, Maratua, SOpn: *S. diversicolor* in open habitat. OTUs are indicated along the *y*-axis by numbers and the lowest taxonomic classification by QIIME. The abundance of each OTU is indicated by colours ranging from black (low abundance or absent) to white (high abundance). The scale bar indicates abundance using a logarithmic base-10 scale.

from a single specimen of *S. diversicolor* in open habitat) were not found in *C. australiensis* hosts in lake habitat. Likewise, taxa belonging to the *Rhizobiales* and *Rhodobacterales* were much more abundant in *C. australiensis* hosts in lake habitat than in open habitat or *S. diversicolor* hosts in all habitats. Taxa belonging to the *Synechococcophycideae* were found in all host–habitat combinations, but reached their greatest abundance in both sponge hosts from Kakaban lake. Finally, dominance as indicated by the relative abundance of the most dominant OTU in each sponge was higher in *S. diversicolor* than in *C. australiensis* hosts (Fig. 4). In samples from *C. australiensis* hosts, the average abundance of the most dominant OTU varied from 12.5% to 34.6% in all habitats; these OTUs included OTU-258 in open habitat and OTUs 176, 1769, 193 and 218 in lake habitats. In samples from *S. diversicolor* hosts, average abundance of the dominant OTU varied from 22.4% to 56.1% in all habitats with the dominant OTU invariably OTU-1733 or OTU-1798.

Importance of host and habitat in structuring composition

There was a highly significant difference in bacterial composition among sponge hosts ($F_{1,14} = 7.66$, $P < 0.001$, $R^2 = 0.225$), habitat ($F_{2,14} = 3.51$, $P < 0.001$, $R^2 = 0.206$) and a significant interaction between host species and habitat ($F_{2,14} = 2.70$, $P < 0.001$, $R^2 = 0.158$). Together, host and habitat explained almost 59% of the variation in bacterial composition. A PCO ordination of the first two axes is presented in Fig. 5. There are three distinct clusters: (1) samples from *C. australiensis* hosts in open water habitat; (2) samples from *S. diversicolor* hosts in open water and lake habitats; and (3) samples from *C. australiensis* hosts in lake habitats. Axis 1 separates samples from *C. australiensis* hosts in open water habitat from *S. diversicolor* hosts in open water and lake habitats and *C. australiensis* hosts in lake habitats. Axis 2 separates samples from *C. australiensis* hosts and *S. diversicolor* hosts. 80.6–86.6% of OTUs were restricted to these clusters (Fig. 6; *S. diversicolor*: 80.6%, i.e. 623 of 773 OTUs; *C. australiensis* lakes: 81.3%; *C. australiensis* open: 86.6%). Less than 2% of OTUs were found in all three clusters (37 of 2127). *Suberites diversicolor* shared 14.2% of OTUs (110 of 773) with *C. australiensis* in lake habitat and 10.0% (77 of 773) with *C. australiensis* in open habitat. *Cinachyrella australiensis* in open habitat shared 9.4% of OTUs (74 of 791) with *C. australiensis* in lake habitat.

There does not appear to be a pronounced difference in composition among samples from *S. diversicolor* in

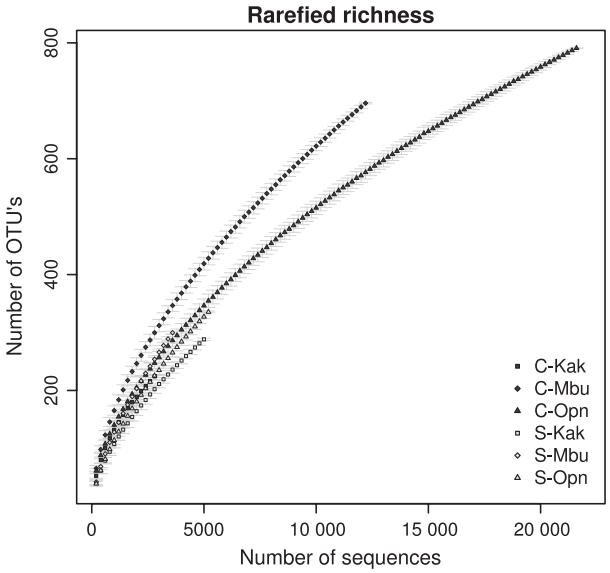

Fig. 3. Species accumulation curves as a function of the number of sequences using resampling of bacterial 16S rRNA gene sequences from *Suberites diversicolor* and *Cinachyrella australiensis* hosts in lake Kakaban (Kak), Haji Buang lake Maratua (Mbu) and open habitat (Opn). Samples are pooled per treatment.

lake or open water habitats. *Suberites diversicolor* hosted three dominant OTUs (184, 1733 and 1798). OTU-184 was identified as a *Gammaproteobacterium,* and its closest relatives included an organism isolated from an oceanic dead zone environment and mussel gill tissue, but with maximum identities of only *c.* 93% (Table 1). OTUs 1733 and 1798 were both identified as belonging to the recently described alphaproteobacterial order *Kiloniellales*. *Suberites diversicolor* in fact hosted three closely related and abundant alphaproteobacterial taxa (OTUs 1733, 1798 and 1850). Together, these three OTUs made up 63.4% of the total bacterial community. All three of these taxa clustered together with organisms isolated from sandy reef sediment and two sponges (*Hymeniacidon heliophila* and *Halichondria* sp.; Fig. 7). They form a strongly supported (bootstrap value = 100) cluster and are distinct from other (putative) members of the *Kiloniellales* including the type species *Kiloniella laminariae* and may represent a novel family (or families) within the order.

In contrast to samples obtained from *S. diversicolor* hosts, there was a pronounced difference in bacterial composition of samples obtained from *C. australiensis* hosts in open vs. lake habitat, which would explain the significant interaction in the adonis analysis. *Cinachyrella australiensis* sponges in lake habitat thus host unique bacterial assemblages compared with sponges in open habitat. Most of the dominant *Alpha-* and *Betaproteobacteria* were restricted to lake habitat. OTUs 199 and 1769 are related to organisms recently isolated from sponges in the Great Barrier Reef. The closest known relative of OTU-193, a *Betaproteobacterium*, was isolated from the sponge *Xestospongia muta* in Florida, but the sequence identity was only 91% (Table 1). The only dominant *Alphaproteobacterium* isolated from *C. australiensis* sponges in open habitat was OTU-2925, whose closest relative was isolated from carbonate sediments in the SouthWest Indian Ridge. Dominant OTUs from *C. australiensis* identified as belonging to the *Gamma-* or *Deltaproteobacteria* were all closely related to organisms isolated from sponges. *Methylococcales* taxa were very closely related (sequence identities > 99%) to organisms isolated from sponges in the Great Barrier Reef and Florida. Interestingly, two OTUs from lake habitat (196 and 327) form a distinct cluster, but this cluster clusters together with that of the *Methylococcales* taxa and is distinct from other *Gammaproteobacteria*. All other OTUs identified as belonging to other phyla were closely related to taxa isolated from sponges with the exception of OTU-218 (Fig. 7). OTU-218 was the only OTU found in all sponge specimens and was identified as belonging to the family *Synechococcaceae* (Table 1). It was very closely related (maximum identity = 100%) to an organism isolated from sea water in the Mediterranean Sea. Only four OTUs were identified as belonging to the proposed phylum *Poribacteria,* and these were all only found in *C. australiensis* hosts in open habitat. The closest relative (sequence identity = 93.72%) of the most abundant of these (OTU-226) was an organism isolated from the sponge *Ircinia variabilis* in the Mediterranean sea.

Discussion

In our study, *Proteobacteria* dominated both sponge species. In contrast to Schmitt *et al.* (2012b), where *Poribacteria* were relatively abundant and present in a variety of sponge species, they were only a minor component of the bacterial flora in this study and were, furthermore, restricted to *C. australiensis* hosts in open habitat. A large number of OTUs remained unassigned at phylum level in the present study. This high number of unassigned sequences is similar to those reported by White *et al.* (2012), a pyrosequencing study of bacterial symbionts within *Axinella corrugata* sponges; in their study, 36% of 16S rRNA gene fragments were unassigned OTUs. In addition to the above, we recorded a large number of OTUs that were unassigned to kingdom, even after quality control and removal of chimera. These included some moderately abundant OTUs. We did not, however, include these OTUs in subsequent analyses, but their presence is noteworthy.

With respect to bacterial OTU richness, there appeared to be both a habitat and species effect. OTU richness was highest in Haji Buang lake, Maratua, intermediate in

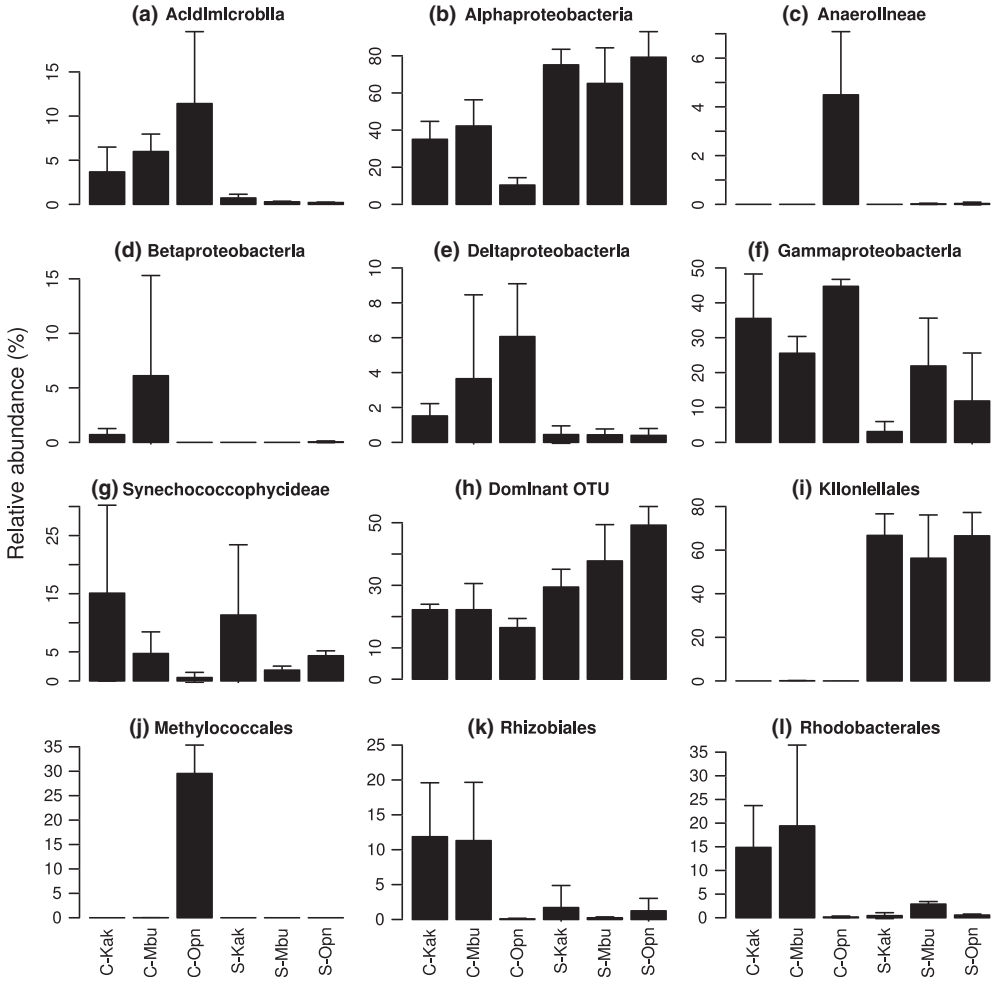

Fig. 4. Relative abundance of the most abundant bacterial classes and the dominant OTU for samples from *Suberites diversicolor* and *Cinachyrella australiensis* hosts in lake Kakaban (Kak), Haji Buang lake Maratua (Mbu) and open habitat (Opn). Error bars represent a single standard deviation. The dominant OTU represents the mean abundance for the single most abundant OTU in each sample, thus not necessarily the same OTU.

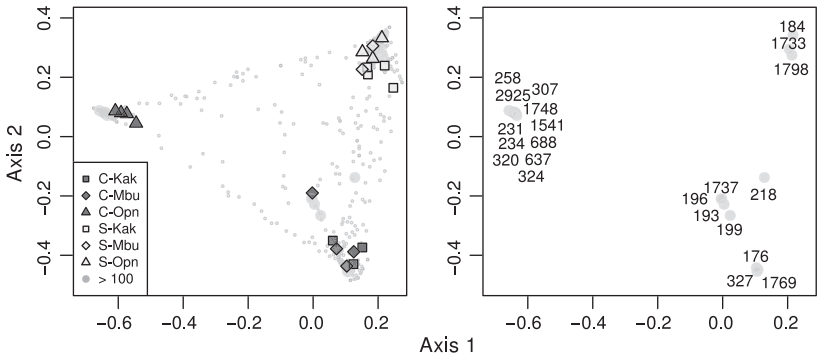

Fig. 5. Ordination showing the first two axes of the PCO analysis. Symbols represent host–habitat combinations for *Suberites diversicolor and Cinachyrella australiensis* hosts in lake Kakaban (Kak), Haji Buang lake Maratua (Mbu) and open habitat. Numbers represent dominant ($\geq$ 500 sequence reads) OTUs. Very small circles represent OTUs < 500 sequence reads.

open habitat and lowest in Kakaban. In each habitat, however, *C. australiensis* hosted more OTUs than *S. diversicolor*. None of the rarefaction curves, though, approached an asymptote indicating that true richness is higher. Other studies of microbial communities of sponges also failed to sample till saturation (Webster *et al.*, 2010; Lee *et al.*, 2011; Jackson *et al.*, 2012; Schmitt *et al.*, 2012b; White *et al.*, 2012).

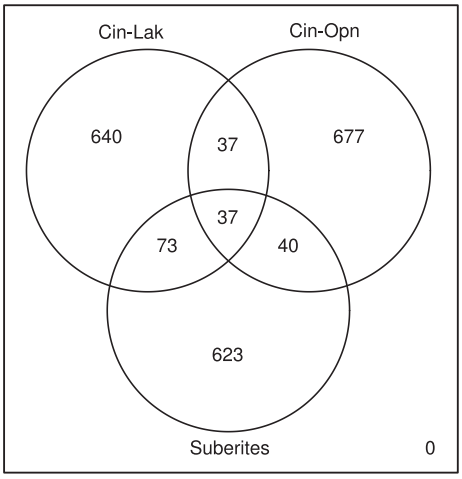

Fig. 6. Venn diagram showing the number of OTUs restricted to given host–habitat combinations, namely *Suberites diversicolor* hosts in open and lake habitat, *Cinachyrella australiensis* hosts in lake habitat (Cin-Lak) and *C. australiensis* hosts in open habitat (Cin-Opn). Overlapping circles indicate shared OTUs.

There were marked differences in the abundance of bacterial classes among sponge hosts and habitats. These results are in accordance with Lee *et al.* (2011). They reported pronounced differences in the abundance of bacterial phyla in different sponge species (*Hyrtios erectus*, *Stylissa carteri*, *Xestospongia testudinaria*) from different locations. Likewise, Erwin *et al.* (2011) reported differences in the relative abundances and presence/absence of OTUs between the sponges *Hymeniacidon heliophila* and *Haliclona tubifera*.

Both sponge species exhibited pronounced dominance with respect to the relative abundance of single taxa in a given sponge individual. This was most pronounced with *S. diversicolor* in open habitats where, on average, close to 50% of the sponge community was dominated by a single OTU. In comparison, the mean relative abundance of the dominant OTU in mangrove rhizosphere and sediment in Brazil was < 2% for all microhabitats (Gomes *et al.*, 2010). The only study we know of where dominance approached the level reported in this study was another sponge study by Webster *et al.* (2010). In most of their samples, the bulk of bacterial diversity consisted of rare OTUs with only one or a few tags with a relatively low proportion of highly abundant taxa. The sponge *Lanthella basta*, however, exhibited pronounced dominance.

Most of the OTUs identified in this study were restricted to a single host (*S. diversicolor* vs. *C. australiensis*) or a single habitat (lake vs. surrounding water). Only 37 OTUs (of 2127, thus, 1.7% of OTUs) were found in both host species in open and lake habitat. The only OTU found in all samples was assigned to the family *Synechococcaceae*. This sequence had 100% sequence similarity to a number of organisms isolated from sea water and various sponge species including *Crella cyathophora* and *S. carteri* (Giles *et al.*, 2012) in Saudi Arabian waters and a large number of cultured organism identified as *Synechococcus* spp. including an organism isolated from the Sargasso Sea (Ahlgren & Rocap, 2012). *Synechococcus* is a widespread *Cyanobacterium* that can be very abundant in the marine euphotic zone. It is also an important component of the autotrophic plankton community (Waterbury *et al.*, 1979).

Most of the dominant OTUs identified during this study were closely related to organisms isolated from sponges, also called Plus-OTUs by other authors as opposed to Minus-OTUs, which are OTUs assigned to a non-sponge-derived sequence (Schmitt *et al.*, 2012a, b). Minus-OTUs in the present study included OTU-1733, whose closest relative was isolated from marine reef sandy sediment, OTU-2925, whose closest relative was isolated from carbonate sediment, OTU-184, whose closest relative was isolated from an oceanic dead zone, and OTU-637, whose closest relative was isolated from an oil field. Given the relatively low sequence similarities (Table 1), however, it is likely that closer, possibly sponge-derived, relatives will be found in the future.

The pronounced dominance of *Kiloniellales* OTUs in *S. diversicolor* hosts is intriguing. Three OTUs identified as belonging to the *Kiloniellales* made up more than 63% of the total bacterial community in *S. diversicolor*. Previous studies have shown that *Suberites* species and their endosymbiotic bacteria produce strong antimicrobial compounds suggesting that the sponge host is a strongly selective environment for bacteria (Thakur *et al.*, 2003; Wiens *et al.*, 2011; Flemer *et al.*, 2012). In addition to this, the type species of the order *Kiloniellales*, *K. laminariae*, was first isolated by selecting active antibiotic producers on agar plates (Wiese *et al.*, 2009).

Kiloniella laminariae is a mesophilic, chemoheterotrophic aerobe with the potential for denitrification and exhibits a typical marine growth response. In its natural environment, it was found in association with the brown alga *Laminaria saccharina*. Although the dominant OTUs inhabiting *S. diversicolor* were identified as *Kiloniellales*, sequence similarity values were below 90% to other (putative) members of the *Kiloniellales* including the type species *K. laminariae*. They did, however, cluster together with these species forming a well-supported cluster in our phylogenetic analysis. The results of the BLAST analysis and our phylogenetic tree, however, indicate the existence of a distinct cluster within the *Kiloniellales*, possibly a new family (or families). This cluster includes organisms isolated from the sponges *H. heliophila* and *Halichondria* sp. in Alabama and an organism isolated from sandy reef sediment in Hawaii (Erwin *et al.*, 2011; Gao *et al.*, 2011). The sheer dominance of the three *Kiloniellales* OTUs in

Table 1. List of abundant OTUs and closely related organisms identified using BLAST search.

OTU	Sum	Host–habitat	Class	Order	Family	GI	Sq ident	Sq len	Source	Location	References
1733	5631	*Suberites*	*Alphaproteobacteria*	*Kiloniellales*	Unclassified	226880021	96.31	405	Marine reef sandy sediment	Hawaii: Oahu, Kilo Nalu reef	Gao *et al.* (2011)
1733	5631					335060676	94.33		Sponge *Hymeniacidon heliophila*	USA: Alabama, Dauphin Isl	Erwin *et al.* (2011)
258	3285	*Cinachyrella* open	*Gammaproteobacteria*	*Methylococcales*	*Methylococcaceae*	345330439	99.53	430	Sponge *Rhopaloeides odorabile*	Australia: Rib Reef, GBR	Webster *et al.* (2011)
1798	2895	*Suberites*	*Alphaproteobacteria*	*Kiloniellales*	Unclassified	226880021	96.31	405	Marine reef sandy sediment	Hawaii: Oahu, Kilo Nalu reef	Gao *et al.* (2011)
176	1528	*Cinachyrella* lakes	*Alphaproteobacteria*	*Rhizobiales*	*Hyphomicrobiaceae*	148732264	98.77	405	Site S25 near Coco's Island	Costa Rica	(K. Rojas-Jimenez, C. Del Valle, B. Leon, W. Ulate, F. Albertazzi, K. Heildelberg and G. Tamayo-Castillo, Unpublished data)
1769	1496	*Cinachyrella* lakes	*Alphaproteobacteria*	*Rhodobacterales*	*Rhodobacteraceae*	400269119	95.54	404	Sponge *Coelocarteria singaporensis*	Australia: Orpheus Island, GBR	(N.S. Webster, H.M. Luter, R.M. Soo, E.S. Botte, C.N. Battershill, D. Abdo and S. Whalan, unpubished data)
231	1476	*Cinachyrella* open	*Acidimicrobidae*	*Acidimicrobiales*	*wb1_P06*	82470225	98.52	406	Sponge *Corticium candelabrum*	Palau	Sharp *et al.* (2007)
688	1246	*Cinachyrella* open	*Gammaproteobacteria*	*Methylococcales*	*Methylococcaceae*	126033048	99.07	430	Sponge *Agelas dilatata*	Bahamas: Little San Salvador Isl	Taylor *et al.* (2007)
218	1192	Ubiquitous	*Synechococcophycideae*	*Synechococcales*	*Synechococcaceae*	407728972	100	407	Seawater	Spain: Catalunya	Erwin *et al.* (2012)

Table 1. Continued

OTU	Sum	Host–habitat	Class	Order	Family	GI	Sq ident	Sq len	Source	Location	References
2925	1086	*Cinachyrella* open	*Alphaproteobacteria*	Unclassified	Unclassified	364524658	97.78	405	Carbonate sediments	South West Indian Ridge	(J. Li, H. Zhou, X. Peng and J. Li, unpublished data)
193	1052	*Cinachyrella* lakes	*Betaproteobacteria*	Unclassified	Unclassified	134290589	90.99	431	Sponge *Xestospongia muta*	USA: Key Largo, FL	Schmitt *et al.* (2008)
1748	1017	*Cinachyrella* open	*Gammaproteobacteria*	*HTCC2188*	*HTCC2089*	110265023	96.28	430	Sponge larva *Ircinia felix*	USA: Key Largo, FL	Schmitt *et al.* (2007)
234	830	*Cinachyrella* open	*Solibacteres*	*Solibacterales*	*Solibacteraceae*	400269018	99.26	405	Sponge *Cymbastela coralliophila*	Australia: Orpheus Island, GBR	(N.S. Webster, H.M. Luter, R.M. Soo, E.S. Botte, C.N. Battershill, D. Abdo and S. Whalan, unpubished data)
199	821	*Cinachyrella* lakes	*Alphaproteobacteria*	*Rhodospirillales*	*Rhodospirillaceae*	400269153	99.26	405	Sponge *Cinachyra* sp.	Australia: Orpheus Island, GBR	(N.S. Webster, H.M. Luter, R.M. Soo, E.S. Botte, C.N. Battershill, D. Abdo and S. Whalan, unpubished data)
1541	786	*Cinachyrella* open	*Gammaproteobacteria*	*Methylococcales*	*Methylococcaceae*	126033057	100	245	Sponge *Agelas dilatata*	Bahamas: Little San Salvador Isl	Taylor *et al.* (2007)
327	668	*Cinachyrella* lakes	*Gammaproteobacteria*	Unclassified	Unclassified	110265053	96.77	433	Sponge larva *Ircinia felix*	USA: Key Largo, FL	Schmitt *et al.* (2007)
324	652	*Cinachyrella* open	*Acidimicrobidae*	*Acidimicrobiales*	*wb1_P06*	158342512	98.28	406	Sponge *Rhopaloeides odorabile*	Australia: Pelorus island	Webster *et al.* (2008)
196	636	*Cinachyrella* lakes	*Gammaproteobacteria*	Unclassified	Unclassified	110265053	95.58	432	Sponge larva *Ircinia felix*	USA: Key Largo, FL	Schmitt *et al.* (2007)
320	602	*Cinachyrella* open	*Anaerolineae*	*SBR1031*	*A4b*	126033013	97.54	406	Sponge *Agelas dilatata*	Bahamas: Little San Salvador Isl	Taylor *et al.* (2007)

Table 1. Continued

OTU	Sum	Host–habitat	Class	Order	Family	GI	Sq ident	Sq len	Source	Location	References
307	563	*Cinachyrella* open	*Gammaproteobacteria*	*Chromatiales*	Unclassified	110265084	96.28	430	Sponge larva *Ircinia felix*	USA: Key Largo, FL	Schmitt *et al.* (2007)
184	536	*Suberites*	*Gammaproteobacteria*	Unclassified	Unclassified	260069446	93.02	430	Saanich Inlet, 120 m depth	48.5883 N 123.5037 W	Walsh *et al.* (2009)
184	536					187473670	92.79		*Bathymodiolus* sp. mussel gill tissue	Papua New Guinea: Manus basin	Won *et al.* (2008)
1737	512	*Cinachyrella* lakes	*Deltaproteobacteria*	*NB1-j*	Unclassified	400269180	99.07	431	Sponge *Cinachyra* sp.	Australia: Orpheus Island, GBR	(N.S. Webster, H.M. Luter, R.M. Soo, E.S. Botte, C.N. Battershill, D. Abdo and S. Whalan, unpubished data)
637	666	*Cinachyrella* open	*Deltaproteobacteria*	*Syntrophobacterales*	*Syntrophobacteraceae*	326372184	97.91	431	Zhongyuan oil field	China	(Z. Wu, D. Lu and Z. Liu, unpublished data)
637	666					379771488	96.75		Sponge *Geodia barretti*	Norway	Radax *et al.* (2012)
203	257	*Cinachyrella* open	*Alphaproteobacteria*	*Rhodospirillales*	*Rhodospirillaceae*	22797718	98.52	405	Sponge *Aplysina aerophoba*	Mediterranean: Banyuls sur Mer	Hentschel *et al.* (2002)
174	225	*Cinachyrella* open	*Nitrospira*	*Nitrospirales*	*Nitrospiraceae*	210161954	99.05	420	Sponge *Corallistes* sp.	?	(J.V. Lopez, C.L. Peterson, A. Ledger, K. Sfanos, S.A. Pomponi, B. Schoch and P.J. McCarthy, unpublished data)
187	158	*Cinachyrella* lakes	*Nitrospira*	*Nitrospirales*	*Nitrospiraceae*	62944578	99.05	420	Sponge *Cymbastela concentrica*	Australia	Taylor *et al.* (2004a)
226	87	*Cinachyrella* open	Unclassified	Unclassified	Unclassified	407728880	93.72	430	Sponge *Ircinia variabilis*	Spain: Catalunya	Erwin *et al.* (2012)
1850	456	*Suberites*	*Alphaproteobacteria*	*Kiloniellales*	Unclassified	226880021	98.28	405	Marine reef sandy sediment	Hawaii: Oahu, Kilo Nalu reef	Gao *et al.* (2011)

Table 1. Continued

OTU	Sum	Host–habitat	Class	Order	Family	GI	Sq ident	Sq len	Source	Location	References
1850	456					335060688	95.07		Sponge *Halichondria* sp.	USA: Alabama, Dauphin Isl	Erwin *et al.* (2011)
1850	456					343202360	89		Marine alga *Saccharina latissima*	Germany:Baltic Sea	Wiese *et al.* (2009)
1850	456					74136944	89		Mature marine biofilm	?	(K.K. Kwon, S.J. Lee, Y.K. Lee, K.H. Cho and H.K. Lee, unpublished data)
1850	456					158537152	86		Oil-polluted saline soil	China: Xianhe, Shangdong	Zhao *et al.* (2010)

OTU, OTU number of dominant OTUs and selected other OTUs; GI, GenInfo sequence identifier of closest related organisms identified using BLAST; Sq ident, sequence identity; Sq len, length of representative sequence used to construct phylogenetic tree; Source, isolation source; Location, location where organism was sampled;?, No location.

S. diversicolor thus indicates that they are well adapted to any antimicrobial substances produced by the host sponge and/or produce strong antimicrobial substances themselves. Either way, the sponge *S. diversicolor* is probably an interesting candidate to explore for novel bioactive compounds, particularly in relation to antimicrobial activity.

In *C. australiensis* hosts, we recorded almost twice as many OTUs (1504 vs. 773 in *S. diversicolor*) and less pronounced dominance. The most abundant OTU (258) was assigned to the family *Methylococcaceae*. Despite being assigned to the *Methylococcaceae*, its closest relative (99% sequence similarity) was identified as a *Nitrosococcus* species (order: *Chromatiales*) isolated from the sponge *Rhopaloeides odorabile* in the Great Barrier Reef (Webster *et al.*, 2011). In a phylogenetic study of the *Gammaproteobacteria*, Gao *et al.* (2009) noted that species from the *Thiotrichales*, *Cardiobacteriales*, *Legionellales*, *Chromatiales*, *Methylococcales* and *Xanthomonadales* revealed deeper branching in trees and unresolved relative branching positions. In the tree, they present in Fig. 1 of their article, *Chromatiales* and *Methylococcales* taxa clustered together. Likewise, Cutiño-Jiménez *et al.* (2010) noted that several distinctive insertions found in most gammaproteobacterial orders were absent from groups including the *Xanthomonadales*, *Legionellales*, *Chromatiales*, *Methylococcales*, *Thiotrichales* and *Cardiobacteriales*.

Other dominant taxa in *C. australiensis* hosts included OTU-176 (family: *Hyphomicrobiaceae*) and OTU-1769 (family: *Rhodobacteraceae*), both restricted to lake habitat. The *Hyphomicrobiaceae* are ubiquitous in terrestrial and freshwater habitats, but only a few have been recorded from marine environments (Huo *et al.*, 2012). Species belonging to the *Hyphomicrobiaceae* have been identified as important methylotrophic denitrifiers (Liessens *et al.*, 1993; Osaka *et al.*, 2006). OTU-1769 was an *Alphaproteobacterium* with 93% sequence similarity to an organism identified as *Rhodovulum imhoffii* isolated from a marine aquaculture pond (Srinivas *et al.*, 2007). The dominance of OTUs related to taxa known to be involved in nutrient cycling, namely nitrogen fixation (*Synechococcus*), ammonia oxidation (*Nitrosococcus*), denitrification (*Hyphomicrobiaceae*), possible denitrification (*Kiloniellales)* and sulphur oxidation (*Rhodovulum*), suggests that sponges in this study may play an important role in reef and lake nutrient dynamics. In Floridian reefs, Southwell *et al.* (2008) demonstrated that the majority of benthic nitrification occurred within sponges and that sponge composition and abundance probably had a strong influence on the concentration and speciation of dissolved inorganic nitrogen in the reef water column. Increased nitrate concentrations near the benthos may in turn affect coral reef community structure. Sponges are found in much greater densities in marine lakes than in either adjacent coral

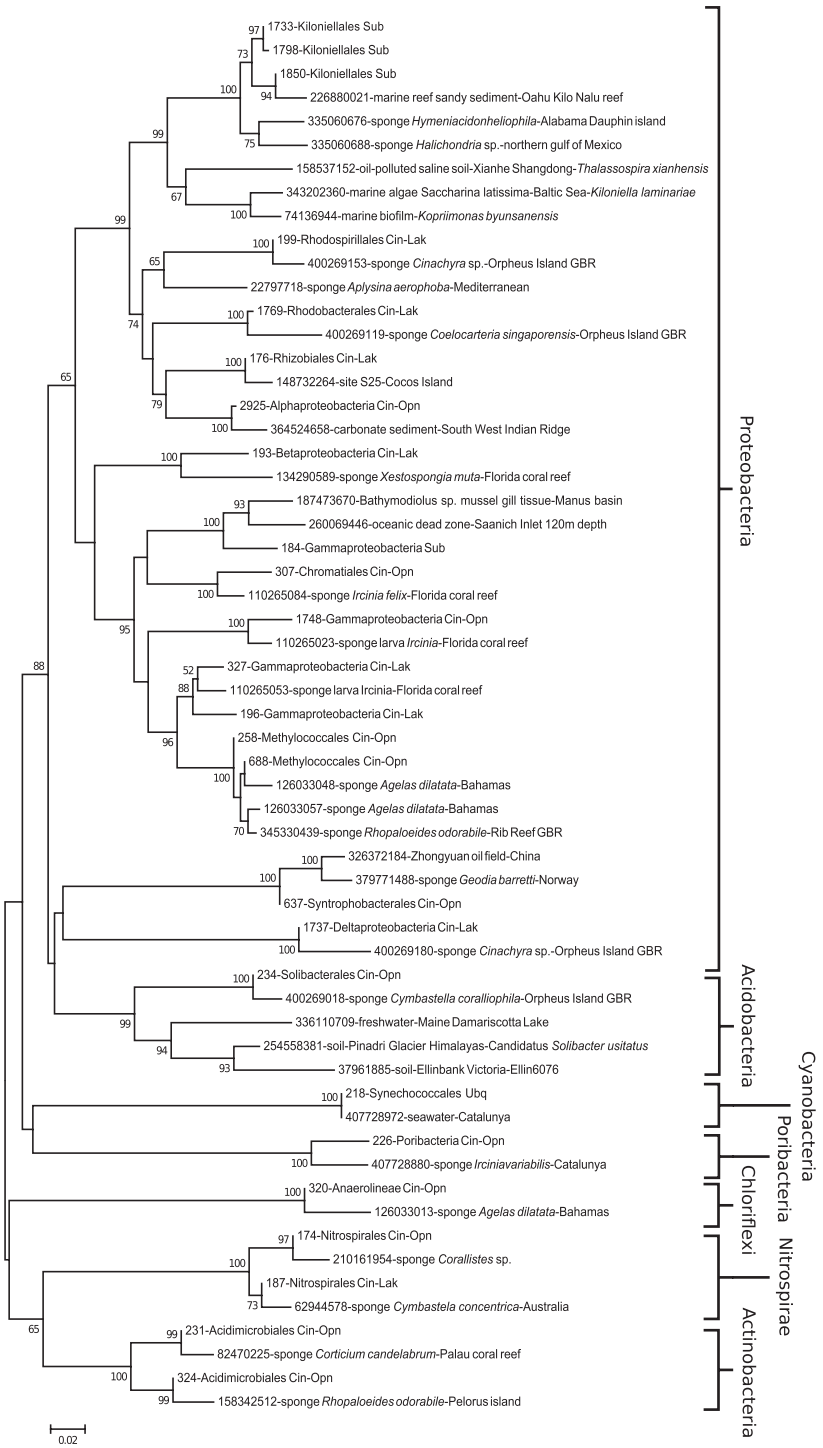

Fig. 7. Phylogenetic tree of the bacterial 16S rRNA gene sequences recovered from *Suberites diversicolor* and *Cinachyrella australiensis* hosts in lake Kakaban, Haji Buang lake, Maratua and open habitat; bootstrap values lower than 50% were omitted. The number of each OTU is indicated as are GenInfo sequence identifiers of sequences obtained using BLAST. Classes of bacteria are indicated. OTUs are assigned to the following clusters: Sub: found in *Suberites diversicolor* hosts, Cin-Opn: found in *C. australiensis* hosts in open habitat, Cin-Lak: found in *C. australiensis* hosts in lake habitat, Ubq: ubiquitous, found in all host individuals. For organisms found using BLAST, we include the host and/or habitat from which the organism was isolated as well as the geographical locality where the organism was isolated.

reefs or mangroves (Becking *et al.*, 2013). They often completely cover the roots of mangrove trees that fringe the marine lakes. Given the generally much higher densities of bacteria in sponges than in the surrounding water, sponges may play a crucial role in nutrient dynamics within the lake ecosystem.

Importance of host and habitat in structuring composition

In the present study, we have demonstrated the importance of both host species and habitat in structuring bacterial composition. Almost 59% of the variation in

composition was explained by the combination of both factors. We were able to discern three distinct clusters in the PCO ordination: (1) a cluster representing assemblages hosted in *S. diversicolor*; (2) a cluster representing assemblages hosted in *C. australiensis* sponges sampled in open water; and (3) a cluster representing assemblages hosted in *C. australiensis* sponges sampled in lake habitat. The ordination confirms the findings of the heat map and demonstrates that *C. australiensis* hosts very different bacterial communities inside and outside marine lake habitat. That this is not the case with *S. diversicolor* may be due to pronounced selective pressure of the sponge on its microbial community, possibly through the production of antimicrobial proteins or antimicrobial activity of the microorganisms themselves. In addition to this, the main populations of *S. diversicolor* are found within lake habitat in contrast to *C. australiensis,* which maintains large and extensive populations outside the lakes.

Our study provides the first assessment of bacterial communities inhabiting sponges in marine lakes. Bacterial composition also differed strongly between sponge host species. The widespread *C. australiensis* showed very pronounced variation in composition between lake and open water habitat, whereas the recently described *S. diversicolor* showed very little difference. Much yet remains to be studied in marine lake environments including bacteria and other microorganisms such as *Archaea* in other sponge species, other organisms, water and sediment. Given the unique nature of the marine lakes, it is probable that sponges and other lake organisms host unique bacteria with potentially valuable pharmaceutical and/or biotechnological properties.

Acknowledgements

Research was funded by Project LESS CORAL – FCOMP-01-0124-FEDER013994, ref[a] PTDC/AAC-AMB/115304/2009, Project PEst-C/MAR/LA0017/2011, cofunded by both the Portuguese Foundation for Science and Technology and COMPETE (POFC) and The Netherlands Organisation for Scientific Research (ALW #817.01.008, Rubicon #825.12.007). The Indonesian Institute of Sciences (PPO-LIPI) and the Indonesian State Ministry of Research and Technology (RISTEK) provided the research permits in Indonesia.

References

Ahlgren N & Rocap G (2012) Diversity and distribution of marine *Synechococcus*: multiple gene phylogenies for consensus classification and development of qPCR assays for sensitive measurement of clades in the ocean. *Front Microbiol* **3**: 213.

Appeltans W, Ahyong ST, Anderson G *et al.* (2012) The magnitude of global marine species diversity. *Curr Biol* **22**: 1–14.

Azzini F, Calcinai B, Cerrano C, Bavestrello G & Pansini M (2007) Sponges of the marine Karst lakes and of the coast of the islands of Ha Long Bay (North Vietnam). *Porifera Research: Biodiversity, Innovation and Sustainability* (Custodia MR, Lobo-Hajdu G, Hajdu E & Muricy G, eds), pp. 157–164. Museo Nacional, Rio de Janeiro.

Becking LE & Lim SC (2009) A new *Suberites* (*Hadromerida*: *Suberitidae*) species from the tropical Indo West-Pacific. *Zool Meded* **83**: 853–862.

Becking LE, Cleary DFR, de Voogd NJ, Renema W, de Beer M, van Soest RWM & Hoeksema BW (2006) Beta diversity of tropical marine benthic assemblages in the Spermonde Archipelago, Indonesia. *Mar Ecol* **27**: 76–88.

Becking LE, Renema W, Santodomingo NK, Hoeksema BW, Tuti Y & de Voogd NJ (2011) Recently discovered landlocked basins in Indonesia reveal high habitat diversity in anchialine systems. *Hydrobiologia* **677**: 89–105.

Becking LE, Cleary DFR & de Voogd NJ (2013) Sponge composition and abundance in marine lakes and coastal mangroves of Berau, Indonesia. *Mar Ecol Prog Ser* **481**: 105–120.

Caporaso JG, Kuczynski J, Stombaugh J *et al.* (2010) QIIME allows analysis of high-throughput community sequencing data. *Nat Methods* **7**: 335–336.

Cerrano C, Azzini F, Bavestrello G, Calcinai B, Pansini M, Sarti M & Thung DC (2006) Marine lakes of Karst islands in Ha Long Bay (Vietnam). *Chem Ecol* **22**: 489–500.

Cleary DFR (2003) An examination of scale of assessment, logging and ENSO-induced fires on butterfly diversity in Borneo. *Oecologia* **135**: 313–321.

Cleary DFR & Genner MJ (2004) Changes in rain forest butterfly diversity following major ENSO-induced fires in Borneo. *Glob Ecol Biogeogr* **13**: 129–140.

Cleary DFR, Mooers AØ, Eichhorn KAO & Menken SBJ (2004) Diversity and community composition of butterflies and odonates in an ENSO-induced fire affected habitat mosaic: a case study from East Kalimantan, Indonesia. *Oikos* **105**: 426–446.

Cock PJ, Antao T, Chang JT *et al.* (2009) Biopython: freely available Python tools for computational molecular biology and bioinformatics. *Bioinformatics* **25**: 1422–1423.

Colin PL (2009) *Marine Environments of Palau* (Coral Reef Research Foundation, eds), pp. 414. Indo-Pacific Press, San Diego, CA.

Cutiño-Jiménez AM, Martins-Pinheiro M, Lima WC, Martín-Tornet A, Morales OG & Menck CFM (2010) Evolutionary placement of Xanthomonadales based on conserved protein signature sequences. *Mol Phylogenet Evol* **54**: 524–534.

Dawson MN & Hamner WM (2005) Rapid evolutionary radiation of marine zooplankton in peripheral environments. *P Natl Acad Sci USA* **102**: 9235–9240.

Dawson MN, Martin LE, Bell LJ & Patris S (2009) Marine lakes. *Encyclopedia of Islands* (Gillespie R & Clague DA, eds), pp. 603–607. University of California Press, Berkeley.

Edgar RC (2010) Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* **26**: 2460–2461.

Edgar RC, Haas BJ, Clemente JC, Quince C & Knight R (2011) UCHIME improves sensitivity and speed of chimera detection. *Bioinformatics* **27**: 2194–2200.

Erpenbeck D & van Soest RWM (2007) Status and perspective of sponge chemosystematics. *Mar Biotechnol* **9**: 2–19.

Erwin P, Olson J & Thacker R (2011) Phylogenetic diversity, host-specificity and community profiling of sponge-associated bacteria in the Northern Gulf of Mexico. *PLoS ONE* **6**: e26806.

Erwin PM, Pita L, Lopez-Legentil S & Turon X (2012) Stability of sponge-associated bacteria over large seasonal shifts in temperature and irradiance. *Appl Environ Microbiol* **78**: 7358–7368.

Felsenstein J (1985) Confidence limits on phylogenies: an approach using the bootstrap. *Evolution* **39**: 783–791.

Flemer B, Kennedy J, Margassery LM, Morrissey JP, O'Gara F & Dobson AD (2012) Diversity and antimicrobial activities of microbes from two Irish marine sponges, *Suberites carnosus* and *Leucosolenia* sp. *J Appl Microbiol* **112**: 289–301.

Gao B, Mohan R & Gupta RS (2009) Phylogenomics and protein signatures elucidating the evolutionary relationships among the *Gammaproteobacteria*. *Int J Syst Evol Microbiol* **59**: 234–247.

Gao Z, Wang X, Hannides A, Sansone F & Wang G (2011) Impact of redox-stratification on the diversity and distribution of bacterial communities in sandy reef sediments in a microcosm. *Chin J Oceanol Limnol* **29**: 1209–1223.

Giles E, Kamke J, Moitinho-Silva L, Taylor M, Hentschel U, Ravasi T & Schmitt S (2012) Bacterial community profiles in low microbial abundance sponges. *FEMS Microbiol Ecol* **83**: 232–241.

Gomes NCM, Cleary DFR, Pinto FN, Egas C, Almeida A, Cunha A, Mendonca-Hagler LCS & Smalla K (2010) Taking root: enduring effect of rhizosphere bacterial colonization in mangroves. *PLoS ONE* **5**: e14065.

Hamner WM & Hamner PP (1998) Stratified marine lakes of Palau (Western Caroline Islands). *Phys Geogr* **19**: 175–220.

Hardoim CCP, Costa R, Araújo FV, Hajdu E, Peixoto R, Lins U, Rosado AS & van Elsas JD (2009) Diversity of bacteria in the marine sponge *Aplysina fulva* in Brazilian coastal waters. *Appl Environ Microbiol* **75**: 3331–3343.

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

Hentschel U, Fieseler L, Wehrl M, Gernert C, Steinert M, Hacker J & Horn M (2003) Microbial diversity of marine sponges. *Prog Mol Subcell Biol* **37**: 59–88.

Hentschel U, Usher K & Taylor M (2006) Marine sponges as microbial fermenters. *FEMS Microbiol Ecol* **55**: 167–177.

Holthuis LB (1973) Caridean shrimps found in land-locked saltwater pools at four indo-west pacific localities (Sinai Peninsula, Funafuti Atoll, Maui and Hawaii Islands), with the description of one new genus and four new species. *Zool Verhandel* **128**: 1–48.

Huo YY, Cheng H, Han XF *et al.* (2012) Complete genome sequence of *Pelagibacterium halotolerans* B2T. *J Bacteriol* **194**: 197–198.

Jackson SA, Kennedy J, Morrissey JP, O'Gara F & Dobson ADW (2012) Pyrosequencing reveals diverse and distinct sponge-specific microbial communities in sponges from a single geographical location in Irish Waters. *Microb Ecol* **64**: 105–116.

Kamke J, Taylor M & Schmitt S (2010) Activity profiles for marine sponge-associated bacteria obtained by 16S rRNA vs 16S rRNA gene comparisons. *ISME J* **4**: 498–508.

Kimura M (1980) A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. *J Mol Evol* **16**: 111–120.

Kott P (1995) A new colonial *Styela* (*Ascidiacea*: *Styelidae*) from an isolated marine habitat, Kakaban Island, East Kalimantan, Indonesia. *Raffles Bull Zool* **43**: 469–474.

Lee Y, Lee J & Lee H (2001) Microbial symbiosis in marine sponges. *J Microbiol* **39**: 254–264.

Lee OO, Wang Y, Yang J, Lafi F, Al-Suwailem A & Qian P (2011) Pyrosequencing reveals highly diverse and species-specific microbial communities in sponges from the Red Sea. *ISME J* **5**: 650–664.

Legendre P & Gallagher E (2001) Ecologically meaningful transformations for ordination of species data. *Oecologia* **129**: 271–280.

Liessens J, Germonpre R, Kersters I, Beernaert S & Verstraete W (1993) Removing nitrate with a methylotrophic fluidized bed: microbiological water quality. *J Am Water Works Assoc* **85**: 155–161.

Maciolek J (1983) Distribution and biology of Indo-Pacific insular hypogeal shrimps. *Bull Mar Sci* **33**: 606–618.

Massin C & Tomascik T (1996) Two new *Holothurians* (*Echinodermata*: *Holothuroidea*) from an anchialine lagoon of an uplifted atoll, Kakaban Island, East Kalimantan, Indonesia. *Raffles B Zool* **44**: 157–172.

McDonald JI, Hooper JNA & McGuinness KA (2002) Environmentally influenced variability in the morphology of *Cinachyrella australiensis* (Carter 1886) (Porifera: Spirophorida: Tetillidae). *Mar Freshw Res* **53**: 79–84.

Ng P & Tomascik T (1994) *Orcovita saltatrix*, a new genus and species of anchialine varunine crab (*Crustacea*: *Decapoda*: *Brachyura*: *Grapsidae*) from Kakaban Island, Indonesia. *Raffles B Zool* **42**: 937–948.

Oksanen J, Kindt R, Legendre P, O'Hara B, Simpson GL, Solymos P, Stevens MHH & Wagner H (2009) *vegan: Community Ecology Package. R package Version 1.15-4*. R Foundation for Statistical Computing, Vienna, Austria. http://CRAN.R-project.org/package=vegan.

Osaka T, Yoshie S, Tsuneda S, Hirata A, Iwami N & Inamori Y (2006) Identification of acetate- or methanol-assimilating

bacteria under nitrate-reducing conditions by stable-isotope probing. *Microb Ecol* **52**: 253–266.

Piel J (2004) Metabolites from symbiotic bacteria. *Nat Prod Rep* **21**: 519–538.

Radax R, Rattei T, Lanzen A, Bayer C, Rapp HT, Urich T & Schleper C (2012) Metatranscriptomics of the marine sponge Geodia barretti: tackling phylogeny and function of its microbial community. *Environ Microbiol* **14**: 1308–1324.

Rützler K (2004) Sponges on coral reefs: a community shaped by competitive cooperation. *Sponge Science in the New Millennium* (Pansini M, Pronzato R, Bavestrello G & Manconi R, eds), pp. 85–148. Bollettino dei Musei e degli Istituti Biologici dell'Universita di Genova, Genova.

Saitou N & Nei M (1987) The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol Biol Evol* **4**: 406–425.

Schmitt S, Weisz JB, Lindquist N & Hentschel U (2007) Vertical transmission of a phylogenetically complex microbial consortium in the viviparous sponge *Ircinia felix*. *Appl Environ Microbiol* **73**: 2067–2078.

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.

Schmitt S, Hentschel U & Taylor MW (2012a) Deep sequencing reveals diversity and community structure of complex microbiota in five Mediterranean sponges. *Hydrobiologia* **687**: 341–351.

Schmitt S, Tsai P, Bell J *et al.* (2012b) Assessing the complex sponge microbiota: core, variable and species-specific bacterial communities in marine sponges. *ISME J* **6**: 564–576.

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.

Southwell MW, Popp BN & Martens CS (2008) Nitrification controls on fluxes and isotopic composition of nitrate from Florida Keys sponges. *Mar Chem* **108**: 96–108.

Srinivas TN, Kumar PA, Sasikala C & Ramana CV (2007) *Rhodovulum imhoffii* sp. Nov. *Int J Syst Evol Microbiol* **57**: 228–232.

Tamura K, Peterson D, Peterson N, Stecher G, Nei M & Kumar S (2011) Mega5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol Biol Evol* **28**: 2731–2739.

Taylor MW, Schupp P, De Nys R, Kjelleberg S & Steinberg P (2004a) Biogeography of bacteria associated with the marine sponge *Cymbastela concentrica*. *Environ Microbiol* **7**: 419–433.

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

Thakur NL, Hentschel U, Krasko A, Pabel CT, Anil AC & Müller WEG (2003) Antibacterial activity of the sponge *Suberites domuncula* and its primmorphs: potential basis for epibacterial chemical defense. *Aquat Microb Ecol* **31**: 77–83.

Tomascik T & Mah AJ (1994) The ecology of 'Halimeda lagoon": an anchialine lagoon of a raised atoll, Kakaban Island, East Kalimantan, Indonesia. *Trop Biodiv* **2**: 385–399.

Tomascik T, Mah AJ, Nontji A & Moosa MK (1997) Part II. *The Ecology of the Indonesia Seas* (Periplus, eds), pp. 643–1388. Tuttle Publishing, Singapore.

Van Soest RWM, Boury-Esnault N, Vacelet J, Dohrmann M, Erpenbeck D, de Voogd NJ, Santodomingo N, VanHoorne B, Kelly M & Hooper JNA (2012) Global diversity of sponges (Porifera). *PLoS ONE* **7**: e35105.

de Voogd NJ, Becking LE & Cleary DFR (2009) Sponge community composition in the Derawan Islands, NE Kalimantan, Indonesia. *Mar Ecol Prog Ser* **396**: 169–180.

Walsh DA, Zaikova E, Howes CG, Song YC, Wright JJ, Tringe SG, Tortell PD & Hallam SJ (2009) Metagenome of a versatile chemolithoautotroph from expanding oceanic dead zones. *Science* **326**: 578–582.

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

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

Waterbury JB, Watson SW, Guillard RRL & Brand LE (1979) Widespread occurrence of a unicellular, marine planktonic, *Cyanobacterium*. *Nature* **277**: 293–294.

Webster N & Taylor M (2012) Marine sponges and their microbial symbionts: love and other relationships. *Environ Microbiol* **14**: 335–346.

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

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

Webster NS, Cobb RE, Soo R, Anthony SL, Battershill CN, Whalan S & Evans-Illidge E (2011) Bacterial community dynamics in the marine sponge *Rhopaloeides odorabile* under *in situ* and ex situ cultivation. *Mar Biotechnol* **13**: 296–304.

White J, Patel J, Ottesen A, Arce G, Blackwelder P & Lopez J (2012) Pyrosequencing of bacterial symbionts within *Axinella corrugata* sponges: diversity and seasonal variability. *PLoS ONE* **7**: e38204.

Wiens M, Schröder HC, Korzhev M, Wang XH, Batel R & Müller WE (2011) Inducible ASABF-type antimicrobial peptide from the sponge *Suberites domuncula*: microbicidal and hemolytic activity *in vitro* and toxic effect on molluscs *in vivo*. *Mar Drugs* **9**: 1969–1994.

Wiese J, Thiel V, Gärtner A, Schmaljohann R & Imhoff J (2009) *Kiloniella laminariae* gen. nov., sp. nov., an *Alphaproteobacterium* from the marine macroalga *Laminaria saccharina*. *Int J Syst Evol Microbiol* **59**: 350–356.

Won YJ, Jones WJ & Vrijenhoek RC (2008) Absence of cospeciation between deep-sea mytilids and their thiotrophic endosymbionts. *J Shellfish Res* **27**: 129–138.

Zhang Z, Schwartz S, Wagner L & Mille W (2000) A greedy algorithm for aligning DNA sequences. *J Comput Biol* **7**: 203–214.

Zhao B, Wang H, Li R & Mao X (2010) *Thalassospira xianhensis* sp. nov., a polycyclic aromatic hydrocarbon-degrading marine bacterium. *Int J Syst Evol Microbiol* **60**: 1125–1129.

Supporting Information

Additional Supporting Information may be found in the online version of this article:

Table S1. Total abundance (Sum) of each OTU, abundance of each OTUs per sample of all OTUs and the taxonomic assignment of each OTU.

Table S2. Total abundance (Cin Lak sqs, Cin Opn sqs, Sub Sqs) and number of OTUs (Cin Lak OTUs, Cin Opn OTUs, Sub OTUs) per cluster aggregated according to kingdom, phylum and class.