Microbial diversity and putative diazotrophy in high and low microbial abundance Mediterranean sponges

Marta Ribes a#, Claudia Dziallas b#, Rafel Coma c, Lasse Riemann b

Institut de Ciències del Mar (ICM-CSIC), Barcelona, Catalunya, Spain a; Marine Biological Section, University of Copenhagen, Helsingør, Denmark b; Centre Estudis Avançats de Blanes (CEAB-CSIC), Blanes, Girona, Catalunya, Spain c.

Running Head: Prokaryotes and nifH genes in Mediterranean sponges

#Address correspondence to Marta Ribes and Claudia Dziallas: mrubes@icm.csic.es, cdziallas@mail.de.

M.R. and C.D. contributed equally to this work.
Microbial communities associated with marine sponges carry out nutrient transformations essential for the benthic-pelagic coupling however, knowledge about their composition and function is still sparse. We evaluated the richness and diversity of prokaryotic assemblages associated with three high-microbial abundance and three low-microbial abundance sympatric Mediterranean sponges to address their stability and uniqueness. Moreover, to examine functionality and because an imbalance between nitrogen ingestion and excretion has been observed on some of these species, we sequenced nitrogenase genes (nifH) and measured N₂ fixation. The prokaryotic communities in the two sponge types did not differ in terms of richness, but highest diversity was found in HMA sponges. Moreover, the discrete composition relative to surrounding seawater indicated that horizontal and vertical transmission affect the microbiomes associated with both sponge categories. nifH genes were found in all LMA and sporadically in one HMA species, and about half of these sequences were common between the different sponge species and also, they were found in the surrounding water suggesting horizontal transmission. N₂ fixation was measurable in the water but not associated with the sponges. Also, nitrogen isotopic ratios of sponge tissue indicated that N₂ fixation was not an important source of nitrogen in these Mediterranean sponges. Overall, our results suggest that compositional and functional features differ between prokaryotic communities associated with HMA and LMA sponges, which may affect sponge ecology.
INTRODUCTION

Marine sponges are sessile filter-feeder metazoans that dominate many hard-bottom benthic substrates around the world (1, 2). They filter large volumes of water and efficiently clear particles down to less than 2 μm in size (3, and references therein). Sponges host complex microbial communities that are generally specific, stable, and distinct in composition compared with the microbes in the surrounding seawater (4-7). Their acquisition can be through horizontal transmission from the surrounding water or by vertical transmission where microbes are transferred from parental sponges to progeny (6, 8).

Sponge-associated microbes can supply limiting nutrients and process metabolic waste to and from the host, respectively (4, 9, 10), but the rich taxonomy and functional genetic potential of the sponge-associated microbes suggest that they carry out additional functions (11). Sponges have been categorized as high-microbial abundance (HMA) or low-microbial abundance (LMA) sponges (12) where HMA species harbor dense, diverse microbial communities and LMA species contain fewer and less diverse microorganisms (6, 13-15). Despite of their distinct microbial compositions, HMA and LMA sponges show a similar functional gene repertoire, which has been interpreted as functional convergence between both microbiome types (10, 11, 16).

A previous study on sympatric LMA and HMA Mediterranean sponges showed a lack of balance between particulate organic nitrogen (PON) ingestion and dissolved inorganic nitrogen (DIN) excretion that could only be explained by the activity of sponge-associated microbes (17). Interestingly, it was shown that even though the amoA...
gene was present in both HMA and LMA species, a net excretion of nitrite and nitrate was only measurable in HMA. In order to account for imbalances between PON and DIN inventories, N\textsubscript{2} fixation has been proposed as an important nitrogen source (17, 18). Reports on sponge-associated N\textsubscript{2} fixation are, however, scarce, and to our knowledge constrained to a few studies on tropical sponges (19-21). In addition, diverse \textit{nifH} genes, encoding nitrogenase reductase, of putative N\textsubscript{2}-fixing bacteria (diazotrophs) were found associated with two sponges from the Florida Keys, indicating that N\textsubscript{2} fixation by sponge symbionts could be an important but neglected source of nitrogen in coral reefs (22, 23). The distribution and importance of N\textsubscript{2} fixation in marine sponges is, however, far from resolved, and in particular, information from temperate marine regions is lacking.

In the present study we expanded previous studies on the diversity and modes of acquisition of the associated microbes in HMA and LMA sponges by adding a temporal scale and asked the questions: 1) are \textit{nifH} genes present in HMA and LMA sponges and, if so, is putative diazotrophy mediated by the same microbial groups in different sponge species? 2) is N\textsubscript{2} fixation a significant source of nitrogen for marine sponges? To address these questions we combined 16S rRNA and \textit{nifH} gene sequencing with measurements of N\textsubscript{2} fixation in sponges and surrounding water from the temperate Mediterranean Sea.

**MATERIAL AND METHODS**

**Sample collection.** In May, August, and November 2009 and in February 2010 scuba divers collected five specimens of each of six sympatric sponge species from off the Montgri Coast (NW Mediterranean Sea, 42°3′N, 3°13′E) at 15 m depth. Three species were previously classified as HMA (\textit{Agelas oroides}, \textit{Chondrosia reniformis} and \textit{Petrosia}})
ficiformis) and three were classified as LMA species (Dysidea avara, Axinella damicornis, Spirastrella cunctatrix) (3, 10, 24). To remove food microbes or loosely associated microbes from the sponges, they were maintained in 0.2 µm - filtered water in the dark for 4-5 h before being frozen in liquid nitrogen. Parallel to sponge collection, 2 water samples (4 l) were collected 2-m above the sponge community. Subsamples of 300–500 mL from each 4 l sample, were filtered in the laboratory through 0.2 µm polycarbonate filters (Nucleopore), and the filters were then frozen in liquid nitrogen. Sponges and filters were stored at -80°C until DNA extraction.

DNA extraction. Eight filters (4 months x 2 replicates) and 48 sponge tissue samples (~ 2 mm³; 4 months x 6 species x 2 replicates) were dissected into small sections using a sterile scalpel. DNA was extracted using the DNeasy Tissue Kit (Qiagen, Valencia, CA) with incubation with lysis buffer overnight and a 100 µl elution volume (10).

Prokaryotic composition in sponges. To gain insight into the community composition of prokaryotes associated with sponge tissue 16S rRNA genes were PCR amplified using the primers 515f and 806r (25) complemented with sample specific barcodes for paired-end Illumina sequencing and MyTaqRed chemicals (Bioline). All samples were amplified in triplicate, quality checked (1% agarose), pooled, and purified with Agencourt AMPure XP (Beckman Coulter). The DNA concentrations were then quantified (PicoGreen, Invitrogen), the samples were pooled in equimolar amounts, and sequenced on a MiSeq Illumina platform at the Berlin Genome Center. The sequences were analyzed using mothur v.1.32.0 (26) according to the online protocol
Singletons were removed from the dataset and sequences were clustered with a 97% threshold into operational taxonomic units (OTUs) using the alignment tool from ARB Silva (www.arb-silva.de), which was also used for the taxonomic assignments. Taxonomies were then verified and corrected (if required) using the Ribosomal Database Project (RDP) (http://rdp.cme.msu.edu). The normalized OTU table resulting from the analysis in mothur was used in Primer6 to calculate similarity matrices, for an analysis of similarity (ANOSIM), and similarity percentage analysis (SIMPER). Data were then visualized by non-metric multidimensional scaling (MDS). A phylogenetic tree was constructed for the OTUs containing at least 50 sequences using FastTree (http://www.microbesonline.org/fasttree/) and iTOL (http://itol.embl.de/).

**Analysis of community richness and diversity.** The indices were normalized on data sets subsampled to the least number of sequences obtained per sample (621 sequences). Richness, referring to the number of OTUs (species), was estimated by the Chao1 index and the estimated number of OTUs to yield a coverage of 100%. Diversity, referring to species richness and evenness (the relative species abundances), was estimated by the number of OTUs needed to achieve an average similarity of 50% and by the Shannon and Simpson indices.

**Composition of putative diazotrophs.** *NifH* genes were amplified using a nested PCR approach with the primers *nifH3* and *nifH4* (27) followed by the primers *nifH1* and *nifH2* (28). The PCRs were conducted with Pure Taq Ready-To-Go PCR Beads (GE Healthcare) and negative controls with pure water (Sigma). PCR reagents were mixed in an ultraviolet (UV)-treated bench (VWR), and DNA template was added in a separate
UV-treated bench. For the initial PCR, 1 µl of a 1:10 dilution of the extracted DNA was used as template, and 1 µl of the PCR product was transferred to the next PCR. PCR products were gel purified (E.Z.N.A. Gel Extraction Kit), cloned (TOPO TA Cloning Kit, Invitrogen), and sequenced commercially (GATC). A total of 451 \( nifH \) sequences were obtained with an average of 14 sequences per sample and a range of 2 – 31 sequences per sample. For the negative control samples, no bands were visible. Nevertheless, a gel slice of the expected product size (~359 bp) was excised, purified, and cloned. No \( nifH \) sequences were obtained from these negative controls.

For the obtained sequences, vector sequences were removed, and the sequences were loaded with the ARB Project (www.arb-home.de) software into the Zehr lab \( nifH \) database (www.es.ucsc.edu/~wwwzehr/research/database/) updated in April 2014. The sequences were translated and aligned in ARB and added to the existing \( nifH \) phylogenetic tree from the database using the Quick-Add method with the maximum parsimony algorithm. Closely related sequences and sequences from this study were exported, and a phylogenetic tree was calculated using FastTree (29). The abundance patterns of OTUs (sequences with ≥ 97% similarity) were used in Primer6 (Primer-E Ltd) to calculate the similarity matrices and ANOSIM. The data were then visualized by MDS.

**N₂ fixation associated with sponges.** Six specimens of each of 3 HMA species (\( Agelas oroides, Chondrosia reniformis, \) and \( Petrosia ficiformis \)), and 3 LMA species (\( Dysidea avara, Axinella damicornis, \) and \( Spirastrella cunctatrix \)) were collected by scuba diving from the location and depth given above, and transported to the laboratory. A piece from each sponge was immediately frozen, and the sponges were acclimatized in
125 L aquaria with running seawater for 24 h. Sponges and seawater were then incubated in 250 mL Pyrex bottles submerged in flowing seawater under *in situ* temperature and day-night light cycle for 24 h. One set of 24 bottles [(6 sponge species + seawater + filtered seawater blank) x 3 replicates] was incubated with *in situ* water amended with 10 mL $^{15}$N$_2$–enriched artificial seawater (see below). Incubations with seawater only were used to measure the planktonic N$_2$ fixation rates. Incubations with 0.2 µm - filtered seawater served as water blanks. As controls we used the same setup of 24 bottles but incubated with air-enriched seawater instead of $^{15}$N$_2$–enriched seawater (Table S1). To measure an initial $^{15}$N$_2$ value of the plankton, duplicate 200 mL samples were obtained from the mixture of water and $^{15}$N$_2$/air –enriched seawater prior to incubation, filtered through a 0.2 µm nucleopore filter, and frozen at -80ºC. $^{15}$N$_2$–enriched water was made from artificial seawater (Cold Spring Harbor protocols) where 10 mL $^{15}$N$_2$ gas (Aldrich lot # MBBB0968V) or air was added per 1 L through the septum using a gas-tight syringe. Bottles were inverted 100 times and maintained for 24 h at which time the bubble was no longer visible. This should ensure 90–100% tracer equilibration (30). Ten mL enriched artificial seawater was then added to each 250 mL Pyrex bottle containing sponge material and incubated for 24 h.

To verify that sponges were in mint condition, picoplankton (mainly *Synechococcus* sp.) abundance and inorganic nutrients were measured at beginning and end of the incubation. For picoplankton counts, triplicate 2 mL water samples were fixed with 1% paraformaldehyde and 0.05% glutaraldehyde (final concentration), frozen in liquid nitrogen, and stored at -80ºC. Flow cytometry (FACSCalibur, Becton Dickinson, 488 nm excitation laser) was used to enumerate the *Synechococcus* sp. according to Gasol and Moran (31). Gating was done using the Becton-Dickinson Cell Quest software.
Samples for nutrient analysis were collected in acid-rinsed 50-mL plastic bottles and frozen. The samples were measured for NH₄⁺ (ammonium), NO₃⁻+NO₂⁻ (nitrate+nitrite, NOx⁻), and PO₄³⁻ (phosphate) with an Alliance autoanalyzer according to Grasshoff et al. (32). The remaining ca. 230 mL of the incubation bottles was filtered through a pre-combusted Whatman GF/F filter. The filters and the sponges were kept frozen at -80ºC until isotopic analysis (Table S1).

**Isotopic analysis.** The ratio of $^{15}$N to $^{14}$N (named $d^{15}$N) was measured for water samples (filters) and sponge tissue (Table S1). The sponges and filters were dried in a lyophilizer, and the sponge tissue was then ground with a ceramic mortar and pestle until homogenous. Ca. 2 mg tissue samples and filters were then analyzed in a Flash EA 1112 elemental analyzer (ThermoFisher Scientific) coupled to a Delta C isotope ratio mass spectrometer through a CONFLO III interface. Acetanilide of known isotopic composition was used as reference.

16S rRNA gene sequences are available from the EMBL database, BioProject PRJNA269111 and BioSamples SAMN03252029 - SAMN03252084. *nifH* gene sequences are available from the GenBank database under accession numbers KP259999 - KP20449.

**RESULTS**

Prokaryotic community composition based on 16S rRNA gene sequencing. Quality filtering of 134,258 paired-end Illumina reads retained 131,555 high-quality sequences from the sponges and seawater samples. The normalization to 621
sequences, resulted in an average richness coverage of 87% and 90-98% in seawater and sponges, respectively. The estimated number of OTUs, yielding a coverage of 100%, did not differ significantly between the HMA and LMA sponges (Table 1). All sponge samples showed a lower species richness than the water samples according to the Chao index. There was no difference in richness between HMA and LMA species (Welch’s t-test, \( p = 0.536 \); Table 1). As for diversity, the number of OTUs needed to achieve an average similarity of 50% per sponge species differed between the HMA and LMA sponges (Welch’s t-test, \( p = 0.015 \), Fig. 1). Likewise, the alpha diversity differed between the HMA and LMA species as calculated by inverse Simpson (Welch’s t-test, \( p = 0.049 \); Fig. 1) and Shannon (Welch’s test, \( p < 0.001 \); Table 1) indexes, showing that the HMA species and water hosted consistently higher diversity than the LMA species.

Prokaryotic community composition was compared along a temporal scale by analyzing May, August, November and February samples. An ANOSIM between the water and sponge types (HMA and LMA) showed significant differences between the sponges (HMA+LMA) and seawater, between HMA and LMA, and between each group and seawater (Table 2). Considering all samples, month did not affect community composition (ANOSIM, \( R = 0.032 \); \( p = 0.845 \); Table 2). Thus, the same sponge species sampled in different months clustered more closely than did samples from different sponge species; the single sponge species differed significantly from water (Table 3). Despite the lack of temporal effect, the MDS showed that the sequences of the HMA species from a particular month clustered separately from other months; i.e., May for \( P. \) ficiformis, February for \( A. \) oroides, and May and August for \( C. \) reniformis (Fig. 2).
Most of the dominant OTUs, containing at least 50 sequences, represented Proteobacteria, primarily Alpha- and Gammaproteobacteria, followed by Bacteroidetes and Actinobacteria (Figs. 3 and S1). Also Archaea, Chloroflexi and Cyanobacteria were present (Fig. S1). Only one dominant OTU was represented in all sponge species but not in the water, and none was exclusively found in water. Twenty-four % of the dominant OTUs where found only in one sponge species and 8% were shared with water (Figs. 3 and S2). Most (70%) of the dominant OTUs were found in 2-5 species and 40% of these were exclusive for sponges (Figs. 3 and S2).

The HMA sponges had four dominant OTUs in common, which were not found in water (OTUs 24, 82, 111, and 200 representing unclassified Gammaproteobacteria, Alteromonadaceae of the Gammaproteobacteria, unclassified Gammaproteobacteria, and Cyanobacteria, respectively). Three OTUs were found only in two HMA species but not in water (OTUs 67, 75, and 161 representing unclassified Bacteroidetes, unclassified Acidobacteria, and Planctomycetes, respectively; Fig. 3). Each of the HMA species hosted 2 or more species-specific OTUs not found in water.

The LMA sponges had five dominant OTUs (27, 47, 65, 83, and 546 representing Alpha- and Betaproteobacteria; Fig. 3) in common but they were also found in the water samples. Two (S. cunctatrix and A. damicornis) of the LMA species hosted two or more specific OTUs not found in the water. Non species-specific OTUs were found in any LMA species.

Putative diazotrophs associated with sponges and water. The composition and temporal persistence of nifH genes were analyzed in all of the sponge species and in
the surrounding water (Table 4). Only the LMA species *D. avara* and *A. damicornis*
contained consistently amplifiable *nifH* genes; 168 and 160 *nifH* gene sequences were
obtained, respectively. For *S. cunctatrix*, *nifH* genes could be amplified only from August
samples; 14 sequences obtained. Of the HMA species, only *P. ficiformis* showed a weak
PCR product in the August samples, resulting in 3 *nifH* sequences. A total of 106 *nifH*
sequences were obtained from the water samples. Whereas rarefaction curves for the
LMA samples were saturated, the curves for water samples were not (data not shown).
However, the duplicate sponge samples showed some variation (Table 4).

The obtained *nifH* sequences fell into the canonical Clusters I, III, and IV including
Proteobacteria and Cyanobacteria, anaerobic bacteria, and diverse *nifH* homologs found
in methanogens, respectively (33, 34). Most of the sponge samples contained *nifH*
sequences in more than one sub-cluster (Fig. 4). Only sub-cluster 3G contained
sequences from the three LMA species and water. Certain sub-clusters included
sequences from water and some sponge species: Sub-clusters 1G, 1K, 3E, and 3S
contained sequences from the water, *D. avara*, and *A. damicornis*. Sub-cluster 1A
contained water and *D. avara* sequences, and sub-clusters 3L, 3P, 4A, and 4F contained
water and *A. damicornis* sequences (Table 4). Only four of the 15 sub-clusters were
specific for sponges: Sub-clusters 1E and 3J were specific for *D. avara* and sub-clusters
3I and 3Q for *D. avara* and *A. damicornis*. Subcluster 3H contained only sequences from
water (Table 4).

The ANOSIM of sequence composition of sponge species and water showed
significant differences (R = 0.21, p = 0.005) because of the water column sequences (Fig.
5). However, there were no significant differences between the sponges if water was
excluded. *D. avara* and *A. damicornis* showed significant differences with a two-way ANOSIM using the sponge species and sampling months as parameters (R = 0.531, p = 0.037). Overall, significant differences were not observed between months; however, there was a difference between May and August (R = 0.185, p = 0.05) and May and February (R = 0.22, p = 0.04) when all sponge samples were considered.

**Nitrogen fixation.** Sponge feeding was confirmed throughout the incubation by a depletion of *Synechococcus* cells in sponge bottles compared to the controls (Welch’s t-test, all p < 0.05; Fig. S3) as well as by excretion of inorganic nutrients (Welch’s t-test, all p < 0.05; Fig. S4). The d\textsuperscript{15}N signal in the plankton (-4.6‰) before \textsuperscript{15}N\textsubscript{2} addition (SW-INI) did not differ significantly from filtered water (-4‰, F-SW-INI; Fig. 6a). Addition of \textsuperscript{15}N\textsubscript{2} (SW*-INI) relative to the blank with filtered seawater (F-SW*-INI; Fig. 6a). After 24 h of incubation, the d\textsuperscript{15}N of plankton from the bottles amended with \textsuperscript{15}N\textsubscript{2} (SW*, Fig. 6a) increased more than 10-fold (65‰; Welsh’s t-test, p <0.0001), whereas no significant changes were detected in the plankton controls (SW) and blanks (F-SW, F-SW*). The plankton that survived the sponge incubation with \textsuperscript{15}N\textsubscript{2} enriched water showed an increase in d\textsuperscript{15}N to 7.6‰ in *D. avara*, which was not significantly different from the initial value of 4.3‰ SW*-INI; however, there was a significant difference in *A. damicornis* (Welch’s t-test, p <0.0001) and *S. cunctatrix* (Welch’s t-test, p =0.006; 55 and 32‰, respectively; Fig. 6a).

The initial background level of d\textsuperscript{15}N in the sponge tissue was 5 - 8‰ in all the species but in *C. reniformis*, which showed a significantly (Welch’s t-test, p <0.05) lower value of 2‰ (Fig. 6b). *A. oroides* was the only species showing increased d\textsuperscript{15}N when...
incubated with $^{15}\text{N}_2$, reaching a mean value of 6‰, which was not significantly different from that of the control (Fig. 6b).

DISCUSSION

Diversity, richness and composition of sponge-associated prokaryotes. The use of Illumina sequencing revealed that microbial communities associated with HMA sponges were more diverse than those associated with LMA sponges as indicated by the Shannon and Simpson diversity indexes, emphasizing the rare and dominant OTUs, respectively. This is consistent with reports using other techniques (6, 13, 35). The Shannon diversity indexes were similar to those calculated for other LMA and HMA species from the Red Sea (36) but differed from a previous study with the same Mediterranean sponge species (24), possibly reflecting difference in sequencing depth. Against expectations, we found no difference in richness between HMA and LMA sponges. While the exact reason for this is unclear, we suggest that our integration of samples over time, which included transient associated microbes may have contributed to the similar richness observed for the HMA and LMA sponge groups.

Our analysis of prokaryotes in sponges and surrounding seawater showed as common features in HMA and LMA species the existence of a specific microbial mixture composed of 1) microbes found in one, few or most of sponge species, which were shared with the water, and 2) “sponge-specific” microbes, which were not present in the water but found in single, few or many sponge species. Overlap between OTUs in the surrounding water and inside a sponge has been interpreted as sign of horizontal transmission (37) whereas “sponge-specific” microbes are likely an outcome of vertical
Interestingly, we found the highest proportion of OTUs potentially transmitted vertically in HMA sponges. In contrast, a higher proportion of the OTUs where acquired from the environment in the LMA sponges. These findings are consistent with a recent suggestion that the two acquisition mechanisms co-occur in sponges (39), but also documents that the relative importance of the mechanisms differs between sponge types.

Diverse phyla dominate microbial communities associated with LMA sponges from different ocean provinces (13, 36, 40). Indeed, our results support the idea that LMA sponges host diverse prokaryotic communities without general common dominant groups at a geographical scale. The HMA sponge communities were dominated by Proteobacteria, Cyanobacteria, Bacteroidetes, Acidobacteria and Planctomycetes. However, phyla and sub-phyla such as Chloroflexi and Gamma- and Alphaproteobacteria, which have been reported to characterize HMA species (36, 40), were found but were not abundant in our samples.

Based on the examined sponge species and water samples, no temporal differences were found in microbial community composition; however, the MDS representation indicated monthly changes for the three HMA species reflecting transient predominance of diverse phylogenetic groups. Based on an earlier study (38) and also on our findings of a higher percentage of OTUs shared with the surrounding seawater in LMA sponges, we anticipated a higher proportion and abundance of temporary prokaryotic groups in LMA relative to in HMA species, but this was not observed. In general, temporal variability in sponge-associated prokaryotic communities has been attributed to OTUs that are low in abundance or transient over time, whereas dominant
OTUs in sponge communities are assumed to be stable through seasons (40, 41, 42), suggesting that sponge hosts buffer against environmental fluctuations. Our finding, however, of abundant transient groups causing monthly differences in the microbial composition in HMA species contrasts with this generalization, and highlights the importance of transient horizontally transmitted microbes in HMA sponges.

**Putative diazotrophs associated with sponges and water.** Symbiosis with $N_2$ fixing prokaryotes often confers a significant ecological advantage to a host that may allow for survival in otherwise marginal habitats (43). For instance, $N_2$ fixation, conceivably carried out in anoxic zones of sponge tissue (44), may maintain nitrogen balance in marine sponges (17, 18). Indeed, $N_2$ fixation has been detected in tropical marine sponges (19, 20, 45, 46, 47). Since planktonic production is at times nitrogen limited in the Mediterranean Sea, especially during summer (48, 49), $N_2$ fixation would presumably be advantageous, and consequently prevalent, in microbial communities hosted by the indigenous sponges. Surprisingly, we found $nifH$ genes in all of the LMA but only in one HMA species, although a general presence of this functional gene was expected in LMA and HMA species based on data from tropical regions (23, 47, 50). Our $nifH$ sequences fell in 15 sub-clusters of which only 6 contained sequences previously found in marine sponges; in most sub-clusters the related sequences originated from microbial mats, lake sediments or termites. The limited overlap with published $nifH$ sequences likely reflects that many of these originate from tropical and not temperate sponges; e.g. *Ircinia strobilina* and *Mycale laxissima* (23, 47), and that only few $nifH$ sequences are available from marine sponge overall.
The fact that most *nifH* sub-clusters contained sequences from sponges and water samples may indicate that those microbes are transient and non-food related members of the sponge microbiota acquired from the water. This is based on the premise that all the sponges cleaned-up their filtration systems for food-microbes because they were kept alive and active in filtered seawater for 4-5 hours before tissue was fixed for analysis. The presence of *nifH* sequences from different sponge species in common sub-clusters was in contrast with the functional convergence (different microbial groups with the same function in each sponge species) found in previous analyses of sponge-associated *amoA* genes (10, 16). This points to the existence of potential differences in the coevolution between sponges and microbes mediating different functions. Despite overlaps between *nifH* genes in the water and sponges, the diazotrophic communities differed significantly due to variability in the relative abundance of *nifH* phylotypes and clusters in water versus sponges. To what extent this was caused by the poor coverage of *nifH* gene diversity in the water samples is, however, unknown. Nevertheless, the consistent presence of *nifH* genes in some LMA species, and the lack in most HMA species, may indicate that LMA species actively select for particular microbial phenotypes, as was suggested for microbial phylotypes above. The fact that *nifH* could only be amplified from some sponge species in summer time may be indicative of a seasonal pattern driven by planktonic nitrogen exhaustion in summer (51). Though, additional data are needed to test this hypothesis.

**N₂ fixation in sponges.** The presence of *nifH* genes in the LMA species was not accompanied by any measurable N₂ fixation. The high concentrations of oxygen and NH₃ could have suppressed N₂ fixation, but our measurements of contemporary planktonic N₂ fixation suggest that the experimental conditions were suitable for N₂ fixation. It is,
however, well known that the mere presence of a functional gene does not imply that a process will be measurable at any time or under all conditions. Indeed, N₂ fixation can be ephemeral *in situ* (sensu 52) and difficult to measure even in pure cultures of diazotrophs (e.g., 53), and published rates from tropical sponges have been very low (21). Hence, we cannot out rule some potential N₂ fixation in the examined Mediterranean sponges, but our data do indicate that N₂ fixation is not a major source of nitrogen. Consequently, imbalances between PON and DIN inventories found for these sponge species (17) are likely not driven by N₂ fixation.

**d¹⁵N as an indicator of N₂ fixation.** Stable isotope composition has been used to investigate nutritional sources in marine symbioses and it has been suggested that low d¹⁵N values of about 1.5 o/oo may be indicative of N₂ fixation (54). Therefore, sponge tissue with this level of d¹⁵N would suggest N₂ fixation by the microbial symbionts (22, 46). Accordingly, the high d¹⁵N values found for LMA sponges from Florida Keys (46) suggest none or negligible N₂ fixation in these, which falls in line with our findings of high d¹⁵N levels and no measurable N₂ fixation in the Mediterranean LMA sponges. Of the examined sponges only the HMA species *C. reniformis* showed a low d¹⁵N of 2 o/oo potentially indicating nitrogen import through N₂ fixation. The lack of *nifH* amplification and direct measurements of N₂ fixation in this species suggest, however, that stable isotope composition is not an ideal indicator for microbial N₂ fixation in sponges. Indeed, preferential uptake of isotopically light nitrate or ammonium from the ambient water or preferential remineralization and uptake of isotopically light nitrogen from DON and PON sources could generate sponge d¹⁵N value similar to those produced by microbial N₂ fixation (47).
**Conclusions.** This first in-depth investigation of combined 16S rRNA and \textit{nifH} gene richness and diversity over different months in Mediterranean sponges shows that species of HMA and LMA sponges host unique associated prokaryotic communities, respectively, that differ from those in the ambient water. These sponge types host microbiota with similar richness but with distinctly higher diversity in HMA sponges. Our findings indicate that horizontal microbial transmission co-occurs with vertical transmission in both sponge categories but the former has prevalence in LMA species. Most \textit{nifH} sub-clusters were shared between sponge-species and the water samples, which may suggest that those microbes are transient members of the sponge microbiota. This is consistent with the \text{d}^{15}\text{N} ratios and \text{\textsuperscript{15}N} incorporation assays indicating that sponge-associated \text{N}_2 fixation is an insignificant and/or ephemeral source of nitrogen in the examined sponges of the temperate Mediterranean Sea.

**ACKNOWLEDGEMENTS**

We thank Paula Lopez-Sendino, Amaia Astarloa, and Jeanett Hansen for excellent laboratory assistance, and Dr. Rafel Simó and Pilar Teixidor (CCiT-UB) for advice on isotope manipulation and analysis. This work was supported by the Spanish "Ministerio de Economía y Competitividad" through grants CGL2010-18466 and CGL2013-43106-R to R.C. and M.R., grants 09-066396 and 11-105450 from The Danish Council for Independent Research, Natural Sciences to L.R., and a Marie Curie Fellowship from EU (project PIEF-GA-2011-299810) to C.D. This is a contribution from the Marine Biogeochemistry and Global Change research group funded by the “Generalitat de Catalunya” (Catalan Government) through grant 2014SGR1029.
Supplementary information is available at the AEM Journal's website.
REFERENCES


Figure legends

Fig. 1. Number of 16S rRNA gene based OTUs contributing to 50% of the average similarity per sponge species or water as calculated by SIMPER analysis and inverse Simpson index as proxy for Alpha diversity. Error bars show standard deviations (n = 8; i.e. duplicates from 4 months), ordinate values give number of OTUs and inverse Simpson index, respectively.

Fig. 2. Non-metric multidimensional scaling (MDS) plot of the similarity of 16S rRNA gene composition in samples from sponges and water at four sampling dates. The 8 symbols per sample represent 2 replicates and 4 months.

Fig. 3. Phylogenetic tree including the 16S rRNA gene sequences representing the most abundant OTUs (>50 sequences per OTU). Presence in one or more samples of a sponge species or water is indicated by colored circles. HMA = sponges with high microbial abundance, LMA = sponges with low microbial abundance.

Fig. 4. Phylogenetic tree of obtained nifH sequences. The accession number of a sequence representing a nearest relative to our sequences is provided for each sub-cluster. The colored boxes show the presence of this sub-cluster in the sampled sponge species and in water. Green and purple arrows show sub-clusters also containing (48) and (23) sequences, respectively.
Fig. 5. Non-metric multidimensional scaling (MDS) plot based on similarities of *nifH* genes obtained from water and sponge samples. The 8 symbols per sample represent 2 replicates and 4 months.

Fig. 6. $^{15}$N to $^{14}$N ratios ($d^{15}$N) in parts per 1000 (‰) for plankton (a) and sponge tissue (b). In black, initial values (INI) representing natural background level. In white, after 24 h of incubation. Error bars are SE. SW: Seawater, F-SW: Filtered seawater, *D. avara* (DA), *A. damicornis* (AD), *S. cunctatrix* (SC), *A. oroides* (AO), *C. reniformis* (CR) and *P. ficiformis* (PF). Asterisks indicate incubations with $^{15}$N$_2$ enriched water.
Table 1. Estimated number of OTUs yielding a coverage of 100% and real coverage of subsampled 16S rRNA gene sequences, richness (Chao1) and diversity (H: Shannon-Wiener) indices per sponge species and seawater. Given are averages ± standard deviation (n = 8). LMA: Low microbial abundance; HMA: High microbial abundance.

<table>
<thead>
<tr>
<th>Sponge type</th>
<th>Species</th>
<th>OTUs</th>
<th>Coverage [%]</th>
<th>Chao1</th>
<th>H</th>
</tr>
</thead>
<tbody>
<tr>
<td>LMA</td>
<td><em>D. avara</em></td>
<td>586 ± 161</td>
<td>92 ± 2</td>
<td>505 ± 152</td>
<td>2.44 ± 0.29</td>
</tr>
<tr>
<td></td>
<td><em>A. damicornis</em></td>
<td>200 ± 99</td>
<td>98 ± 1</td>
<td>229 ± 85</td>
<td>1.80 ± 0.17</td>
</tr>
<tr>
<td></td>
<td><em>S. cunctatrix</em></td>
<td>679 ± 254</td>
<td>90 ± 4</td>
<td>628 ± 229</td>
<td>2.32 ± 0.92</td>
</tr>
<tr>
<td>HMA</td>
<td><em>A. oroides</em></td>
<td>521 ± 275</td>
<td>91 ± 3</td>
<td>483 ± 179</td>
<td>4.33 ± 0.40</td>
</tr>
<tr>
<td></td>
<td><em>C. reniformis</em></td>
<td>221 ± 75</td>
<td>97 ± 0.5</td>
<td>283 ± 69</td>
<td>3.69 ± 0.15</td>
</tr>
<tr>
<td></td>
<td><em>P. ficiformis</em></td>
<td>370 ± 164</td>
<td>93 ± 0.5</td>
<td>439 ± 117</td>
<td>4.62 ± 0.16</td>
</tr>
<tr>
<td></td>
<td>seawater</td>
<td>1443 ± 583</td>
<td>87 ± 3</td>
<td>1000 ± 164</td>
<td>5.10 ± 0.34</td>
</tr>
</tbody>
</table>
Table 2: ANOSIM (analysis of similarities) results per sponge type and water for 16S rRNA genes obtained by Illumina sequencing. Month = May, August, November and February, HMA = sponges with high microbial abundance, LMA = sponges with low microbial abundance.

<table>
<thead>
<tr>
<th></th>
<th>R</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Month</td>
<td>0.032</td>
<td>0.846</td>
</tr>
<tr>
<td>Water vs. sponges</td>
<td>0.141</td>
<td>0.006</td>
</tr>
<tr>
<td>Sponges (incl. water)</td>
<td>0.906</td>
<td>0.001</td>
</tr>
<tr>
<td>HMA vs. LMA</td>
<td>0.220</td>
<td>0.001</td>
</tr>
<tr>
<td>HMA vs. water</td>
<td>0.306</td>
<td>0.001</td>
</tr>
<tr>
<td>LMA vs. water</td>
<td>0.282</td>
<td>0.003</td>
</tr>
</tbody>
</table>
Table 3. Analysis of similarities (ANOSIM) of 16S rRNA gene sequences obtained from sponge species and water samples. HMA = sponges with high microbial abundance; LMA = sponges with low microbial abundance.

<table>
<thead>
<tr>
<th></th>
<th>R</th>
<th>LMA</th>
<th></th>
<th>HMA</th>
<th></th>
<th>water</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>p</td>
<td>D. avara</td>
<td>A. damicornis</td>
<td>S. cunctatrix</td>
<td>A. oroides</td>
<td>C. reniformis</td>
</tr>
<tr>
<td>LMA</td>
<td></td>
<td>1.000</td>
<td>1.000</td>
<td>0.891</td>
<td>0.694</td>
<td>0.889</td>
</tr>
<tr>
<td>A. damicornis</td>
<td>0.002</td>
<td>1.000</td>
<td>0.981</td>
<td>0.989</td>
<td>0.798</td>
<td>1.000</td>
</tr>
<tr>
<td>S. cunctatrix</td>
<td>0.001</td>
<td>0.002</td>
<td>0.990</td>
<td>0.705</td>
<td>0.897</td>
<td>1.000</td>
</tr>
<tr>
<td>HMA</td>
<td></td>
<td>0.001</td>
<td>0.001</td>
<td>0.001</td>
<td>0.844</td>
<td>0.836</td>
</tr>
<tr>
<td>A. oroides</td>
<td>0.002</td>
<td>0.002</td>
<td>0.001</td>
<td>0.002</td>
<td>0.470</td>
<td>1.000</td>
</tr>
<tr>
<td>C. reniformis</td>
<td>0.001</td>
<td>0.002</td>
<td>0.001</td>
<td>0.001</td>
<td>0.008</td>
<td>0.977</td>
</tr>
<tr>
<td>P. ficiformis</td>
<td>0.001</td>
<td>0.002</td>
<td>0.001</td>
<td>0.001</td>
<td>0.001</td>
<td>0.002</td>
</tr>
</tbody>
</table>

|                     | water | 0.001 | 0.001 | 0.002 | 0.001 | 0.001 | 0.002 |
Table 4. Number of sequences from sponge and waters samples within specific *nifH* gene sub-clusters as defined in Figure 4. Duplicates for each sample are shown. HMA = sponges with high microbial abundance; LMA = sponges with low microbial abundance.

<table>
<thead>
<tr>
<th></th>
<th>1A</th>
<th>1E</th>
<th>1G</th>
<th>1K</th>
<th>3E</th>
<th>3G</th>
<th>3H</th>
<th>3I</th>
<th>3J</th>
<th>3L</th>
<th>3P</th>
<th>3Q</th>
<th>3S</th>
<th>4A</th>
<th>4F</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>water</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>May 1</td>
<td>1</td>
<td>1</td>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>3</td>
<td>5</td>
</tr>
<tr>
<td>May 2</td>
<td>2</td>
<td>3</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>Aug 1</td>
<td>6</td>
<td>3</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>Aug 2</td>
<td>5</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Nov 1</td>
<td>1</td>
<td>2</td>
<td>7</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>Nov 2</td>
<td>1</td>
<td>1</td>
<td>6</td>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Feb 1</td>
<td>1</td>
<td>2</td>
<td>7</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>Feb 2</td>
<td>1</td>
<td>1</td>
<td>6</td>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td><strong>HMA:</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>P. ficiformis</em></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Aug 1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Aug 2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>LMA:</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>S. cunctatrix</em></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Aug 1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Aug 2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>D. avara</em></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>May 1</td>
<td>7</td>
<td>9</td>
<td>7</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>May 2</td>
<td>2</td>
<td>3</td>
<td>20</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Aug 1</td>
<td>24</td>
<td>1</td>
<td>14</td>
<td>13</td>
<td>6</td>
<td>1</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Aug 2</td>
<td>1</td>
<td>4</td>
<td>6</td>
<td>3</td>
<td>4</td>
<td>3</td>
<td>7</td>
<td>4</td>
<td>17</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nov 1</td>
<td>6</td>
<td>3</td>
<td>2</td>
<td>4</td>
<td>4</td>
<td>7</td>
<td>4</td>
<td>17</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nov 2</td>
<td>25</td>
<td>1</td>
<td>4</td>
<td>3</td>
<td>17</td>
<td>4</td>
<td>17</td>
<td>17</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Feb 1</td>
<td>4</td>
<td>3</td>
<td>6</td>
<td>4</td>
<td>7</td>
<td>4</td>
<td>17</td>
<td>17</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Feb 2</td>
<td>9</td>
<td>1</td>
<td>7</td>
<td>4</td>
<td>1</td>
<td>4</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>A. damicornis</em></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>May 1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>May 2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Aug 1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Aug 2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nov 1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nov 2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Feb 1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Febr 2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Fig. 1. Number of 16S rRNA gene based OTUs contributing to 50% of the average similarity per sponge species or water as calculated by SIMPER analysis and inverse Simpson index as proxy for Alpha diversity. Error bars show standard deviations (n = 8; i.e. duplicates from 4 months), ordinate values give number of OTUs and inverse Simpson index, respectively.
Fig. 2. Non-metric multidimensional scaling (MDS) plot of the similarity of 16S rRNA gene composition in samples from sponges and water at four sampling dates. The 8 symbols per sample represent 2 replicates and 4 months.
Fig. 3. Phylogenetic tree including the 16S rRNA gene sequences representing the most abundant OTUs (>50 sequences per OTU). Presence in one or more samples of a sponge species or water is indicated by colored circles. HMA = sponges with high microbial abundance, LMA = sponges with low microbial abundance.
Fig. 4. Phylogenetic tree of obtained nifH sequences. The accession number of a sequence representing a nearest relative to our sequences is provided for each sub-cluster. The colored boxes show the presence of this sub-cluster in the sampled sponge species and in water. Green and purple arrows show sub-clusters also containing (48) and (23) sequences, respectively.
Fig. 5. Non-metric multidimensional scaling (MDS) plot based on similarities of *nifH* genes obtained from water and sponge samples. The 8 symbols per sample represent 2 replicates and 4 months.
Fig. 6. $^{15}$N to $^{14}$N ratios ($d^{15}$N) in parts per 1000 (‰) for plankton (a) and sponge tissue (b). In black, initial values (Ini) representing natural background level. In white, after 24 h of incubation. Error bars are SE. SW: Seawater, F-SW: Filtered seawater, _D. avara_ (DA), _A. damicornis_ (AD), _S. cunctatrix_ (SC), _A. oroides_ (AO), _C. reniformis_ (CR) and _P. ficiformis_ (PF). Asterisks indicate incubations with $^{15}$N$_2$ enriched water.