FEMS Microbiology Ecology, 91, 2015, fiv067

doi: 10.1093/femsec/fiv067
Advance Access Publication Date: 22 June 2015
Research Article

RESEARCH ARTICLE

Hologenome theory supported by cooccurrence networks of species-specific bacterial communities in siphonous algae (*Caulerpa*)

Tania Aires[1,†], Yann Moalic[2,3,†], Ester A. Serrao[1] and Sophie Arnaud-Haond[1,4,*]

[1]CCMAR, Centre of Marine Sciences, University of Algarve, Gambelas, 8005-139 Faro, Portugal, [2]IFREMER-Technopole de Brest-Iroise, BP 70, 29280 Plouzané, France, [3]UMR 6197-Laboratoire de Microbiologie des Environnements Extrêmes, Université de Bretagne Occidentale (UBO) Institut Universitaire Européen de la Mer (IUEM), CNRS, Plouzané, France and [4]UMR MARBEC (Marine Biodiversity, Exploitation and Conservation) Bd Jean Monnet, BP 171, 34203 Sète Cedex - France

*Corresponding author: UMR MARBEC (Marine Biodiversity, Exploitation and Conservation) Bd Jean Monnet, BP 171, 34203 Sète Cedex - France. Tel: +33 4 99 57 32 61: E-mail: sarnaud@ifremer.fr
†These authors contributed equally to this work.
One sentence summary: Species-specific bacterial community inside *Caulerpa* genus form very specific clusters revealed in networks of cooccurrence, suggesting that these communities play an important role in the metabolism of their host, particularly in its invasive potential.
Editor: Julian Marchesi

ABSTRACT

The siphonous algae of the *Caulerpa* genus harbor internal microbial communities hypothesized to play important roles in development, defense and metabolic activities of the host. Here, we characterize the endophytic bacterial community of four *Caulerpa* taxa in the Mediterranean Sea, through 16S rRNA amplicon sequencing. Results reveal a striking alpha diversity of the bacterial communities, similar to levels found in sponges and coral holobionts. These comprise (1) a very small core community shared across all hosts (< 1% of the total community), (2) a variable portion (ca. 25%) shared by some *Caulerpa* taxa but not by all, which might represent environmentally acquired bacteria and (3) a large (>70%) species-specific fraction of the community, forming very specific clusters revealed by modularity in networks of cooccurrence, even in areas where distinct *Caulerpa* taxa occurred in sympatry. Indirect inferences based on sequence homology suggest that these communities may play an important role in the metabolism of their host, in particular on their ability to grow on anoxic sediment. These findings support the hologenome theory and the need for a holistic framework in ecological and evolutionary studies of these holobionts that frequently become invasive.

Keywords: endophytic bacteria; bacterial community; cooccurrence network; modularity; coevolution

INTRODUCTION

Bacteria are ubiquitous worldwide and, as pathogens, commensals or symbionts, they play an often cryptic but major role in the evolution of eukaryotes through coevolution and as a factor of host speciation (Janson *et al.* 2008). The genomic richness of those microbial communities can thus play a determinant role both in adaptation and evolution of higher organisms (Zilber-Rosenberg and Rosenberg 2008; Tonon *et al.* 2011; Dittami *et al.*

Received: 16 April 2015; **Accepted:** 15 June 2015

2014). The hologenome theory (Reshef *et al.* 2006; Rosenberg *et al.* 2007) proposes the explicit consideration of the role played by the cooperation between different genomic compartments in evolution and adaptation of the holobiont (host and bacterial microbiota).

Last decade's advances on massive sequencing technologies have revolutionized our access to the microbial compartment of biodiversity, improving the understanding of its influence on geochemical cycles (Myshrall *et al.* 2010) and on the metabolism, ecology and evolution of eukaryotes (Chandler *et al.* 2011; Moore *et al.* 2011; Sachs, Skophammer and Regus 2011). Deep amplicon sequencing of 16S ribosomal RNA (rRNA) gene showed that microbial communities in many environments including marine habitats, soil, plants and animals are much more diverse than previously assumed (Huber *et al.* 2007; Costello, Lauber and Hamady 2009; Hollister *et al.* 2010; Turnbaugh *et al.* 2010).

Bacterial assemblages associated with marine organisms like sponges, corals, algae and seagrasses include groups involved in important metabolic processes such as nitrification, nitrogen fixation, sulfate reduction, photosynthesis, plant growth enhancement, morphogenesis induction or chemical defense (Chisholm *et al.* 1996; Nakanishi and Nishijima 1996; Crump and Koch 2008; Lee *et al.* 2009; Barott *et al.* 2011; Burke *et al.* 2011; Orole and Adejumo 2011). Recent molecular studies have explored the epiphytic bacterial diversity on some macroalgal species (Lachnit *et al.* 2009; Tujula *et al.* 2010; Barott *et al.* 2011; Burke *et al.* 2011). These include taxa associated with functions such as inducing the release and settlement of spores (Joint *et al.* 2007; Weinberger *et al.* 2007), or positively influencing algal growth and development (Matsuo *et al.* 2005) and providing essential nutrients (Keshtacher-Liebson, Hadar and Chen 1995; Croft *et al.* 2006). Some of these studies compared bacterial communities of different algal species in sympatry and allopatry and evidence suggested the presence of species-specific bacterial assemblages (Lachnit *et al.* 2009; Barott *et al.* 2011), supporting an important role on the metabolism of their host. However, these few studies on algae and marine plants mostly focus on epiphytic communities (Crump and Koch 2008; Tujula *et al.* 2010; Burke *et al.* 2011). There is evidence that epiphytic bacteria associated with algae are unique and different from the surrounding water (Lachnit *et al.* 2011), thus endophytic communities can be expected to be even more tightly associated with their host. Also, the endophytic communities of siphonous green algae, such as *Caulerpa* sp., have been shown to be stable over time (Hollants *et al.* 2011a) and distinct from the epiphytic bacterial assemblage of the same alga (Hollants *et al.* 2011b).

The coenocytic green algae of the genus *Caulerpa*, distributed mostly in tropical and subtropical marine waters around the world, are known to be associated with diverse endosymbiotic bacteria (Dawes and Lohr 1978; Chisholm *et al.* 1996; Meusnier *et al.* 2001; Delbridge *et al.* 2004; Aires *et al.* 2012, 2013). Several species of *Caulerpa* have become invasive in non-native regions, and they can exhibit enhanced performance in low nutrient and anoxic sediment (Chisholm *et al.* 1996; Chisholm and Moulin 2003), conferring competitive advantage over native species. Three species or species complexes occur in the Mediterranean Sea; the native *Caulerpa prolifera* (Ollivier 1929; Rayss 1941), and two non-native species, *C. racemosa* (including three varieties, two of which at least recently introduced) and *C. taxifolia* (Panayotidis 2006).

The bacterial communities of some *Caulerpa* species have been previously studied. A pioneer study on *C. taxifolia* based on DGGE (denaturing gradient gel electrophoresis) screening of clones from 16S libraries identified five major lineages among which three appeared dominant (Meusnier *et al.* 2001). These studies did not however allow the comprehensive characterization of its highly diverse bacterial community. Although NGS (next generation sequencing) of 16S allows more exhaustive inventories of bacterial communities, the coamplification of dominant chloroplastidial DNA in photosynthetic organisms prevented efficient sequencing of bacterial DNA (Jiao *et al.* 2006). By eliminating a great part of the chloroplasts through bleaching, this technical limitation has been circumvented for *Caulerpa* (Aires *et al.* 2012), allowing the study of endophytic bacterial communities associated with *C. racemosa* var. *cylindracea* that invaded the Mediterranean Sea (Aires *et al.* 2013).

The aim of the present study is the characterization of endophytic bacterial communities associated with four *Caulerpa* taxa (two species and two varieties of a third one) occurring in the Mediterranean Sea to (1) assess the diversity of endophytic communities in these congeneric species with contrasting histories in the Mediterranean; (2) discriminate species-specific or habitat-specific operational taxonomic units (OTUs) and bacterial assemblages, in order to distinguish the influence of phylogenetic factors linked to the identity of the host, versus ecological or geographical factors, in the partitioning of those communities; and (3) highlight potential coevolving bacterial lineages or clusters of OTUs embedded in the complex structure of microbial communities by using cooccurrence network analysis. The combination of these results is expected to provide information as to the species-specific nature of endophytic communities and the fitting of the hologenome concept to *Caulerpa* sp., in order to formulate hypothesis that may guide future evolutionary and functional studies on algal holobionts.

MATERIAL AND METHODS

Summary of the history of *Caulerpa* taxa in the Mediterranean

The invasive *C. taxifolia* appeared in the Mediterranean in 1984 after accidental release from Monaco Oceanographic Museum (Meinesz and Hesse 1991) and since then has invaded most of the Western Mediterranean, spreading from Cote D'Azur (where it was first reported) to Tunisia (Langar *et al.* 2000). The other Mediterranean non-native, *C. racemosa*, was recorded for the first time in the beginning of 20th century (Hamel 1926) in Sousse Harbour, Tunisia, now known to be var. *turbinata-uvifera*. Only in the early 1990s, the disastrous effects of one of the varieties (currently named var. *cylindracea*) started to affect the Mediterranean at a global scale (Nizamuddin 1991; Verlaque *et al.* 2000) by threatening/outcompeting important native species as *Posidonia oceanica*. Nowadays, three *C. racemosa* varieties are recognized in Mediterranean Sea. Both *C. racemosa* var. *lamouroxii* and *C. racemosa* var. *turbinata-uvifera* exhibit no invasive profile. Their origin, whether Red Sea or elsewhere, is still a matter of debate. The invasive variety, *C. racemosa* var. *cylindracea*, is suspected to have come from Western Australia (Panayotidis 2006; Aires *et al.* 2013).

Sampling strategy

Each sample collected at each site was composed of several (≥3) sampling units (SUs) or ramets; a piece of thallus from a single individual alga. In order to include all the different types of tissues of the alga, each single piece of thallus collected (SU) was composed of a set of interconnected fronds, stolons and rhizoids. A total of 43 SUs were analyzed for this study (Table 1),

Table 1. Number of sample units processed per locality (**ND**—non-disinfected).

Sampling locality		No. of samples sequenced
France- Villefranche-Sur-Mer	*C. taxifolia*	3 disinfected/ 3 ND
	C. racemosa var. *turbinata-uvifera*	6 disinfected
	C. prolifera	7 disinfected/ 3 ND
Tunisia- Tunis- Sidi Daoud	*C. taxifolia*	3 disinfected/ 3 ND
	C. racemosa var. *turbinata-uvifera*	3 disinfected/ 3 ND
	C. prolifera	3 disinfected/ 3 ND
Greece - Crete	*C. racemosa* var. *turbinata-uvifera*	3 disinfected/ 3 ND
	C. prolifera	3 disinfected/ 3 ND
Spain- Mallorca	*C. racemosa* var. *cylindracea*	6 disinfected/ 3 ND
	C. prolifera	6 disinfected/ 3 ND

from four taxa: three *Caulerpa* species including two different varieties from the *C. racemosa* complex (*C. racemosa* var. *turbinata-uvifera* and *C. racemosa* var. *cylindracea*). SUs were subdivided for the disinfected and non-disinfected samples, so that these treatments were applied to the same individual algae. Sediment samples were also analyzed as a control to test if environmental bacteria were present in the community associated with disinfected seaweed samples. Seaweeds were collected by scuba diving in four Mediterranean sites at depths ranging from 11 to 16 m during late spring and summer (Villefranche, Tunis, Crete and Mallorca). For each species, at least three SUs were analyzed per site, collected more than 2 m apart. *Caulerpa racemosa* and *C. prolifera* were sampled in all four Mediterranean sites whereas *C. taxifolia* was only found in two of those sites (Villefranche-Sur-Mer and Tunis). The two varieties of *C. racemosa* were not found within the same locality (Table 1) and their morphological identification was confirmed using molecular markers (Aires *et al.* 2013). As a result, two sites hosted simultaneously *C. racemosa* var. *turbinata uvifera*, *C. prolifera* and *C. taxifolia* (Tunis and Villefranche), one site hosted *C. prolifera* and *C. racemosa* var. *cylindracea* (Mallorca) and the last one hosted *C. prolifera* and *C. racemosa* var. *turbinate uvifera* (Crete). The seaweeds were kept at 4°C during transport. Within one to four hours after sampling, they were rapidly surface cleaned for sediment particles and visible epiphytes with a toothbrush, and frozen at −80° C.

Algal material processing and DNA extraction

Samples were analyzed for endophytic bacterial communities following a pilot study indicating that these were more associated with the host, whereas epiphytic bacteria were more similar to those in environmental (sediment) samples (Aires *et al.* 2012). In order to study endophytic communities, seaweeds were pretreated using a surface disinfection protocol that increases the yield of endophytic OTU sequences (Aires *et al.* 2012). The surface disinfection protocol consisted in sequential washes with ethanol and bleach and is detailed in Aires *et al.* (2012).

Bacterial DNA extraction was performed using FastDNA® SPIN Kit for Soil (MP biomedicals LLC). All the different types of *Caulerpa* structures (fronds, stolons and rhizoids) were included in each extraction. When the samples were too big to fit in the vials, samples were cut in smaller pieces making sure that every different structure was included. Non-disinfected SUs were used as control for comparative purposes, by splitting in two some of the SUs prior to disinfection, and using one of the parts as a non-disinfected control (Table 1 and Table S1, Supporting Information).

Bacterial tag-encoded 16S amplicon pyrosequencing

Extracted DNA was submitted to Biocant (Cantanhede, Portugal) to be analyzed through tag pyrosequencing (GS FLX Titanium, 454-Life Sciences-Roche) after 16S rRNA gene amplification with modified primers for region V4 (Wang *et al.* 2007). PCR amplification of the hypervariable V4 region of the 16S rRNA gene was performed using an 8 bp key tag.

Bacterial community analysis and characterization

Retrieved sequences were initially trimmed for quality with the standard SFF software tools from Roche454. All the downstream analyses were then performed using the program QIIME: Quantitative Insights Into Microbial Ecology (Caporaso *et al.* 2010). Sequences were screened for a minimum read length of 120 bp and less than 2 or more undetermined nucleotides. The filtered dataset containing only high-quality sequence was then submitted to a conservative chimera detection filter using the ChimeraSlayer method (Haas *et al.* 2011). Selected high-quality chimera-free sequences were clustered into OTUs within reads using UCLUST module from QIIME and a pairwise identity threshold of 0.97.

Representative sequences for each OTU were chosen using the 'most abundant' method on QIIME, which consists in selecting the most abundant sequence as the representative one. OTU sequence alignment was performed with Pynast using the Greengenes core set as template to align against. Taxonomic assignment used the Ribosomal Database Project classifier using an 80% confidence threshold. Sequences having the best match with eukaryotes (i.e. chloroplasts and mitochondria) were excluded from the OTU table and downstream analyses. To assign each OTU to the closest matching described taxon, BLASTN searches were performed against the SILVA database and sequences were putatively assigned to a described taxon if the e-value exceeded a minimum threshold of 0.001 (default value). The degree of relatedness between the subsets of the most common sequences was inferred using the phylogenetic reconstruction with QIIME's script make_phylogeny.py and using, by default, FastTree (Price, Dehal and Arkin 2010); root was chosen by the tree method default from QIIME.

The OTU table was normalized through rarefaction using the minimum number of sequences as the upper limit of rarefaction depths and all the analysis were performed using the corrected table. After quality control and filtration of chimeras and of organelle sequences, the normalized OTU table was used to calculate Chao I richness estimates (Chao 1984) of diversity within

samples (α-diversity) (which was also calculated prior to normalization). This measure was also calculated after pooling SUs according to sampling site. A dot plot was constructed using the Chao value found for the standardized minimum number of sequences (for both individual SUs and pooled SUs), in order to compare diversities between sites. All analyses only used disinfected SUs with coverage close to 50% and/or 1000 sequences. That procedure was not used for some of the non-disinfected SUs and sediment samples that didn't fit the criteria, either due to high yield of chloroplastidial DNA or to a high richness resulting in coverage below 50%. Those disinfected SUs that didn't fit these criteria are not represented in Table S1 (Supporting Information) and were not considered for any analysis in this study.

Pairwise diversity between samples (β-diversity) was estimated by clustering samples using the unweighted UniFrac algorithm (Lozupone and Knight 2005), on normalized OTU table, and principal coordinates analysis (PCoA) 3D plots were constructed to visualize the data. Based on the phylogenetic structure, the unweighted UniFrac algorithm is a qualitative measure which allows SUs clustering according to taxonomic composition, minimizing the bias caused by different sequencing efforts (Krych *et al.* 2013).

Statistical differences between OTU hits of different replicates of disinfected vs non-disinfected vs sediment were assessed by one-way ANOSIM Bray–Curtis distances using PAST (Ver. 2.16). (Hammer, Harper and Ryan 2001) and tested using 9999 permutations. Permutational multivariate analysis of variance (PERMANOVA) analyses were performed to compare the three different sources of variation (disinfection, *Caulerpa* taxa and location) and test for the homogeneity of bacterial community structure among those factors. Both, Bray–Curtis and unweighted Unifrac distances were used. If no significant differences were found, pairwise PERMANOVA was performed to detect the possible responsible factor among location, taxa and disinfection treatment. PERMANOVA was performed using the Vegan package in R (Oksanen *et al.* 2011).

The four species and varieties were pooled in an OTU table used to determine rarefaction curves. Proportions of each bacterial community section, core (OTUs present in all *Caulerpa* species/varieties), variable (OTUs present in less than 70% of hosts but in at least two species, as in Schmitt *et al.* 2012) and species-specific (OTUs exclusive of one of the host species/varieties) were accessed, and a Venn diagram was constructed using Venny (Oliveros 2007) in order to illustrate the core bacterial community shared among the four *Caulerpa* taxonomic entities (three species, one encompassing two genetically distinct varieties of *C. racemosa*) and the species-specific communities. Core and species-specific communities were isolated and the proportion of each order was assessed and represented graphically showing the most representative orders. Metadata was submitted to Sequence Read Archive (SRA, accession number: ERP002593 [http://www.ebi.ac.uk/ena/data/view/ERP002593]).

Haplotype network

In order to investigate possible coevolution between associated bacteria and hosts, sequences corresponding to the most ubiquitous OTU were used to build a haplotype network. These are useful to discriminate the lineages tightly associated with hosts rather than habitat dependent, and to provide a first step towards the identification of putative symbiotic lineages on the basis of their blast identification and consequent taxonomic assignment. The choice of a network of haplotypes is related to the short length of sequences obtained through pyrosequencing (usually <250 bp) which do not allow robust phylogenetic reconstruction. Haplotype reconstruction followed the protocol described in Aires *et al.* (2013). The preprocessing option implemented in Network (Foster *et al.* 2001) was used for Star contraction, using a radius with maximum size of 5 and the median joining procedure was implemented and followed by an MP procedure to remove the unnecessary median vectors and links (Polzin and Daneschmand 2003) and reduce the complexity of the network to improve its visualization.

Cooccurrence network analysis

Cooccurrence networks depicting correlations between OTUs across SUs were investigated (using non-parametric Spearman). Nodes represent the OTUs and the links are the positive correlations between these OTUs (as in Barberán *et al.* 2012). Starting with a matrix of cooccurrences (OTUs vs SUs), we computed adjacency matrix of correlation for OTUs observed at least in two different SUs (1386 OTUs for the overall dataset). Then, our methodology differs from the one detailed in Barberán *et al.* (2012) by two stages. First, rather than defining a threshold based on an arbitrary value of correlation, networks were analyzed at their percolation threshold as described by Moalic *et al.* (2012) and Kivelä, Arnaud-Haond and Saramäki (2015). This inner property of the network allows pinpointing the value (here correlation) threshold above which the network overall organization into a giant cluster collapses and cohesive subclusters emerge. Second, community/modularity detection algorithms were used to explore patterns of OTUs cooccurrence, by detecting modules of OTUs and assessing whether these modules of cooccurrence, possibly revealing interdependency, are species specific (Faust and Raes 2012). Modularity measures whether communities are organized as modules with dense connections among internal nodes and sparse connections among nodes belonging to distinct modules (Lancichinetti and Fortunato 2009). In other terms, the number of edges between modules is smaller than expected in a random association and different algorithms have been developed to maximize the modularity (Fortunato 2010).

Here, four different community detection algorithms were compared: Leading Eigenvector, Multi-Level, Label propagation and InfoMAP (Newman 2006; Raghavan *et al.* 2007; Blondel *et al.* 2008; Rosvall and Bergstrom 2008) by using igraph (Csárdi and Nepusz 2006). While these algorithms are implanted through different methods (see Fortunato 2010 for review), the results were similar in number and composition of modules. We chose to present the results obtained with the 'community leading eigenvector' algorithm that include a higher number of OTUs in modules. Briefly, the leading eigenvector method separates vertices (nodes, here OTUs) into two communities if their corresponding elements in the eigenvector of the modularity matrix harbor different signs. In a biological perspective, this reflects mutual exclusion of two OTUs. Classical measures of network topology such as node degree, diameter or clustering coefficient were also calculated (Newman 2003). Statistical analyses were performed in R and network visualization with Pajek (Batagelj and Mrvar 2002) and Gephi (Bastian, Heymann and Jacomy 2009).

RESULTS

Bacterial communities' diversity

A total of 181 673 high-quality sequences (after removal of chimera and chloroplast sequences) were analyzed

(SRA, accession number: ERP002593 [http://www.ebi.ac.uk/ena/data/view/ERP002593]); 460–5137 for each of the disinfected replicates (three to seven per locality) and 393–13 195 for non-disinfected samples (Table S1, Supporting Information). Sequences were clustered into OTU using a 97% identity threshold, on an average length exceeding 200 bp.

Higher diversities were found in non-disinfected samples (therefore including both endophytic communities and external communities forming the biofilm), with a minimum expected OTU of 519.05 for *C. taxifolia* from Tunis (although with weak coverage) to a maximum of 6366.73 in the communities of *C. racemosa* var. *cylindracea* from Mallorca (Table S1, Supporting Information). Analyzing each replicate individually, results show a highly variable diversity of OTUs ranging, for endophytic bacteria, from 60 OTU in a *C. taxifolia* SU to 577 in a *C. racemosa* var. *turbinata-uvifera* SU (Table S1, Supporting Information). In non-disinfected samples (endophytic and epiphytic), the number of OTU ranged from 168 in a SU of *C. taxifolia* to 2391 in sample of *C. racemosa* var. *turbinata-uvifera* (Table S1, Supporting Information). For further analyses of the endophytic communities, only the ones exhibiting coverage close to 50% and/or more than 1000 sequences were retained. Chao values standardized for the minimum number of sequences (Fig. S1, Supporting Information) show that diversity is highest in *C. racemosa* from Tunis (Chao value = 1247.13) and Villefranche Chao value = 2541.51) and lowest in *C. taxifolia* and *C. prolifera* from Villefranche (Chao value = 315.06, Chao value = 752.91, respectively).

Rarefaction curves, for pooled *Caulerpa* taxa, based on number of bacterial OTUs, show that all the species are approaching a plateau, except *C. taxifolia* (Fig. S3, Supporting Information).

β-diversity measures/statistical tests

Non-disinfected samples (total bacterial diversity) clustered, regardless of species or sites, apart from those *disinfected hosts (i.e. endophytic bacteria) and together with sediment samples, showing that global communities (both endophytic and external, including biofilm) are dominated by the 'external' bacteria and are similar to those found in the sediment (Fig. 1). Groups of SUs also differ significantly when comparing the two treatments (endophytic versus global communities) and sediment (Table S2, Supporting Information).

Results of the global PERMANOVA showed no significant interaction among the three factors (taxon, location, disinfection) together (Table S3, , Supporting Information; $P = 0.236$) although all factors interact when analyzed per pair ($P < 0.05$). The pairwise tests showed however that the interactions might be strongly impinged by the general similarity of non-disinfected samples, for which no significant differences were found in both factors 'taxa' and 'location' (Table S4 A and B, Supporting Information). For the endophytic (disinfected) samples, in contrast, communities differed significantly among different *Caulerpa* species in Villefranche and Mallorca (Table S4 A, Supporting Information), yet there were no significant differences in taxa between samples obtained from Tunis and Greece (Tunis—*C. prolifera* vs *C. racemosa* var. *turbinata-uvifera* $P = 0.089$, *C. prolifera* vs *C. taxifolia* $P = 0.094$, *C. racemosa* var. *turbinata-uvifera* vs *C. taxifolia* $P = 0.099$; Greece—*C. prolifera* vs *C. racemosa* var. *turbinata-uvifera* $P = 0.916$; pairwise PERMANOVA results—Table S4 A, Supporting information—in agreement with PCoA results, Fig. 2A). That lack of pattern/clustering is more evident in Greece. Tunis shows some partitioning, although different taxa communities are sparser when compared to Villefranche (Fig. 2A). The PCoA representations show that different *Caulerpa* species indeed exhibit clearly distinct bacterial communities regardless of site (Fig. 2A). The only exception is Greece, where bacterial communities of *C. racemosa* var. *turbinata-uvifera* and *C. prolifera* are variable but do not cluster apart.

In PCoA plots displaying bacterial communities per taxon, some clustering per locality sometimes appeared marginally (Fig. 2B). This is supported by pairwise PERMANOVA results for the 'location' factor (within the disinfected level) showing that endophytic communities predominantly reflect their host species, but are sometimes secondarily also depending on their environment (Tables S4 B, Supporting Information), or on the genetic background of their host if algal beds are genetically differentiated in space. PERMANOVA tests results based on Bray–Curtis and unweighted Unifrac distance represented the same conclusions; therefore, only unweighted Unifrac distances-based tests were represented in Tables S3 and S4 (Supporting Information).

Endophytic bacterial community characterization (core, species specific and variable)

Comparing the relative abundance of core, variable and species-specific endophytic bacterial communities shows that most (75.8%, data not shown) of the bacterial community is species specific. The variable part adds up to 23.5% and the remaining

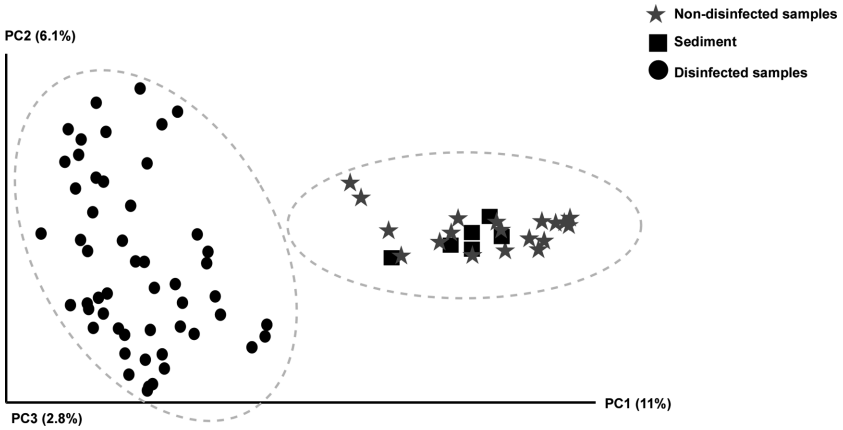

Figure 1. PCoA based on unweighted pairwise phylogenetic distances (UniFrac algorithm) between the bacterial communities of all SUs and locations and both, disinfected and non-disinfected, treatments.

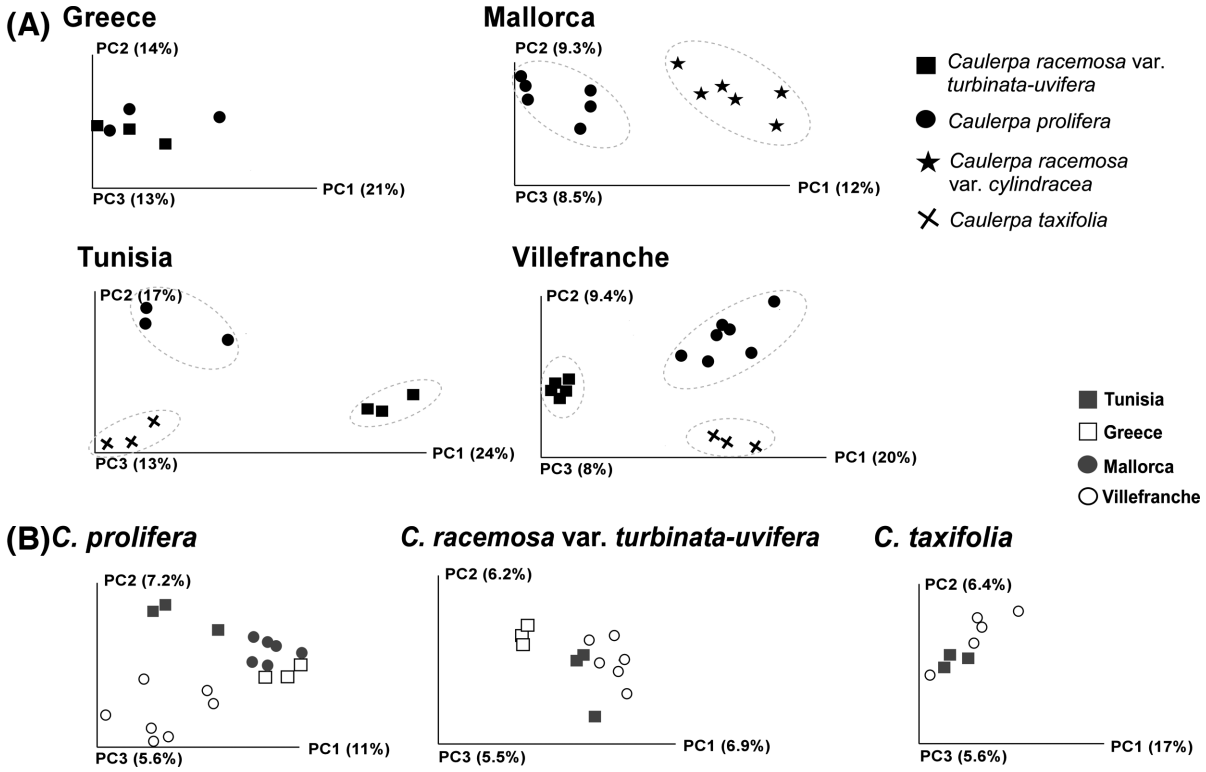

Figure 2. PCoA based on unweighted pairwise phylogenetic distances (UniFrac algorithm) of *Caulerpa* samples from Mediterranean Sea represented **A**–by site; **B**—by species.

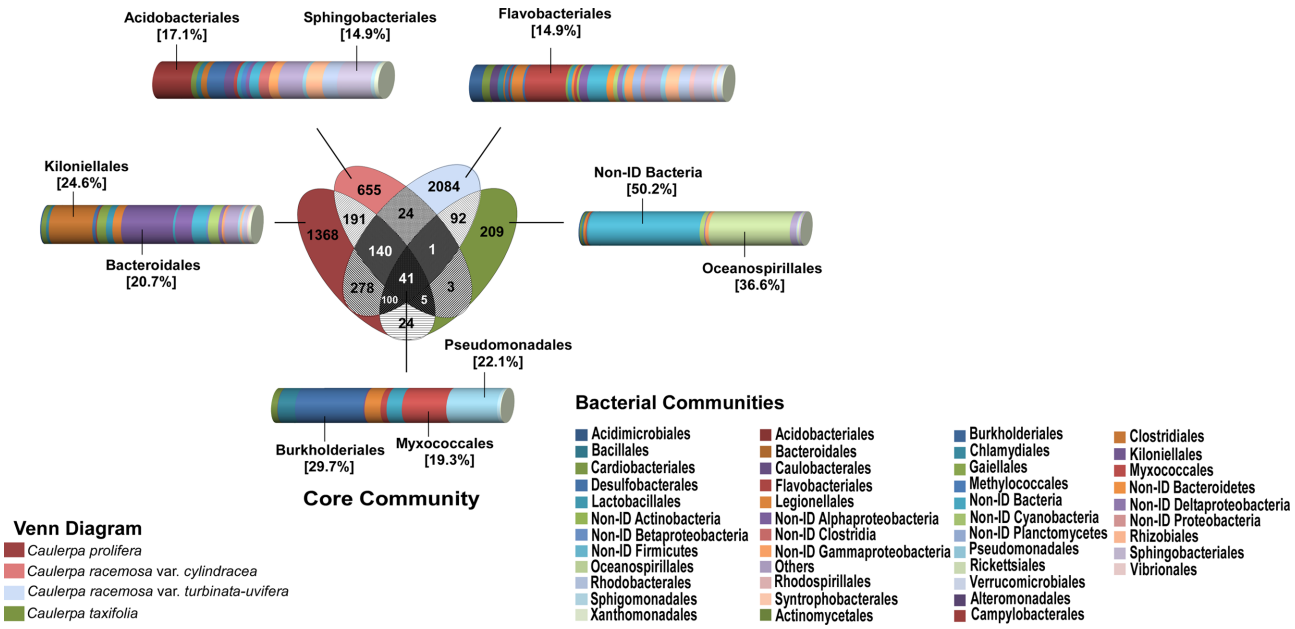

Figure 3. Venn diagram representing bacterial communities shared within the four Mediterranean *Caulerpa* taxonomic entities. The main orders of species-specific and core communities are represented as bar graphs and the percentages of the most conspicuous are indicated in the respective graph.

small part (0.7%) represents the OTUs shared among all the host species (data not shown).

Core and species-specific communities also differ in the relative proportions of bacterial groups (Fig. 3). The most represented classes in all SUs were Alpha-, Beta- and Gammaproteobacteria. The core community was composed of 41 different OTUs shared among all *Caulerpa* taxa where the most conspicuous orders, showing in at least 70% of the replicates, were Burkholderiales (29.7%), Pseudomonadales (22.1%) and Myxococcales (19.3%) (Fig. 3). Residual orders found to be shared among taxa were not represented in a majority of the SUs and cannot be considered really 'core OTUs'. *Caulerpa prolifera* and *C. racemosa* var. *turbinata-uvifera* shared the greatest number of OTUs (278, 7.5%), whereas *C. taxifolia* and

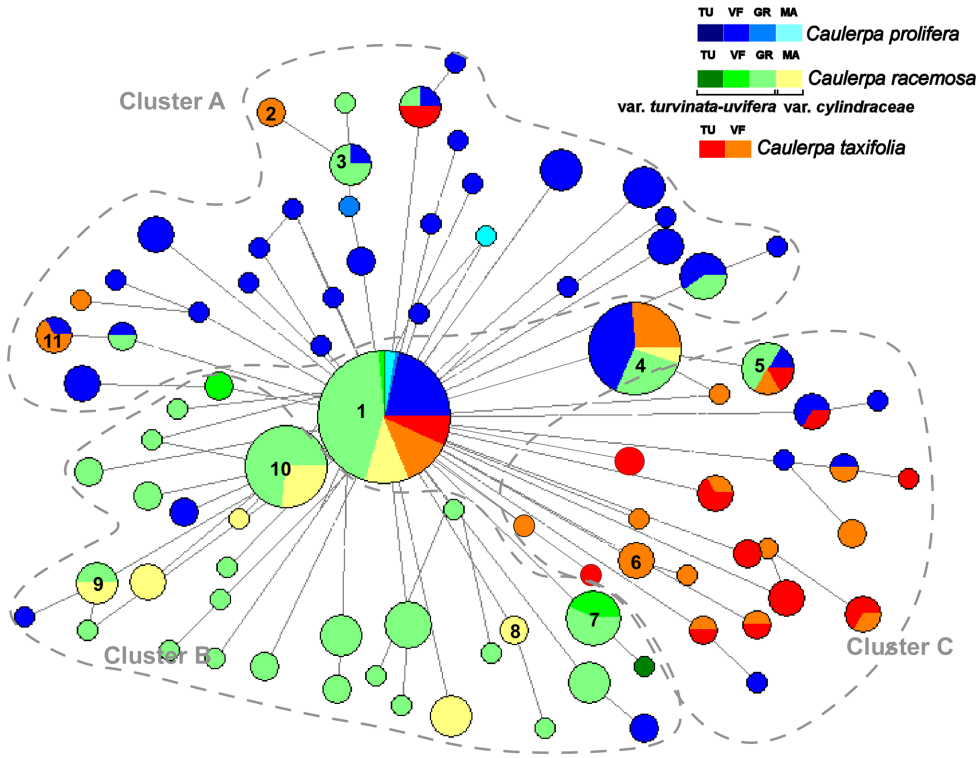

Figure 4. Network of haplotypes from the most ubiquitous OTU (5647) in the set of OTUs present in the *Comamonadaceae* family. Pie charts have size illustrating the proportion of each haplotype and the different slice colors represent the proportion observed in each *Caulerpa* taxa. The network was drawn without keeping distance of links proportional to the number of mutations, in order to illustrate the clustering rather than the divergence. Nodes 1–11 result of collapse during star contraction; each number has the correspondence in Table S5 (Supporting Information). Dashed lines are delineating haplotypes according to taxa for an easier visualization (Clusters A–C). Legend: **TU**—Tunisia, **VF**—Villefranche, **GR**—Greece, **MA**—Mallorca; **1**—*Caulerpa racemosa* var. *turbinata-uvifera*, **2**—*Caulerpa racemosa* var. *cylindracea*.

C. racemosa var. *cylindracea* exhibited the smallest amount of shared OTUs (three OTUs, 0.3%). From the 278 OTUs shared between the native *C. prolifera* and the introduced *C. racemosa* var. *turbinata-uvifera*, the most common were Actinobacteria (17.2%), Betaproteobacteria (18.7%) and Gammaproteobacteria (21.9%). The last two classes were composed of 80.4% Burkholderiales and 35% Pseudomonadales, respectively. As for the species with the least shared OTUs (*C. taxifolia* and *C. racemosa* var. *cylindracea*), 75% of the rare shared ones were Burkholderiales and 25% Pseudomonadales.

Despite the large number of exclusive OTUs (1368), the native *C. prolifera* presents only two conspicuous species-specific orders. Kiloniellales is the leading order (24.6%) followed by Bacteroidales with 20.7% (Fig. 3). The invasive *C. taxifolia* is the host species that shows the least variety of OTUs within its specific bacterial community portion (209 OTUs). Most of them could not be assigned beyond domain (50.2% of non-ID bacteria) but the others were mainly identified as belonging to the order Oceanospirillales (36.6%) (Fig. 3). As for the two varieties of *C. racemosa*, the *C. racemosa* var. *turbinata-uvifera* shows the highest diversity of species-specific OTUs (2084 OTUs) with Flavobacteriales ranking as the most abundant order (14.9%) (Fig. 3). The other variety, *C. racemosa* var. *cylindracea* exhibits lower richness (655 OTUs) and Actinobacteriales (17.1%) and Sphingobacteriales (14.9%) rank as its main species-specific orders (Fig. 3).

Haplotype network

Within the most abundant order in the core community, Burkholderiales, the most represented family was Comamonadaceae, which include nitrogen-fixing and nitrate-reducing bacteria (Gihring *et al.* 2011). Within this family, the most ubiquitous OTU (OTU5647, ~206 base pairs—KR110204–KR111423) had highest similarity (99%) with *Alicycliphilus denitrificans* (GenBank CP002657.1) on the whole 210-bp fragment. The network of haplotypes for this most ubiquitous OTU (OTU5647) showed a huge node common to all species and sites and some minor nodes linking haplotypes within each species regardless the sampling site (Fig. 4). These independent nodes cluster according to species (Clusters A–C).

The few nodes of the haplotype network appearing as shared among different species, mostly encompass species-specific OTUs (1–11 in Fig. 4 and Table S5, Supporting Information) but were collapsed during the network contraction needed to reduce the number of nodes and depicting a 'readable' network. In reality, they represent closely related, yet distinct, haplotypes among species (Table S5, Supporting Information), with the exception of three nodes showing identical haplotypes shared among taxa (node 1, 4 and 5) but sampled in distinct localities of the Mediterranean.

Cooccurrence tracked by community network analysis

As the aim was to compare the module compositions among the *Caulerpa* taxa, four 'taxa-specific' networks were built and community structure was investigated at their respective percolation threshold (Fig S4 and Table S6, Supporting Information). Then, based on the overall dataset, the global network encompassing 1386 OTUs was analyzed to be compared the taxa-specific networks (Fig. 5). Its percolation threshold is reached at the correlation value of 0.76, with an overall network topology also divided in three subclusters (1386 nodes and 6207

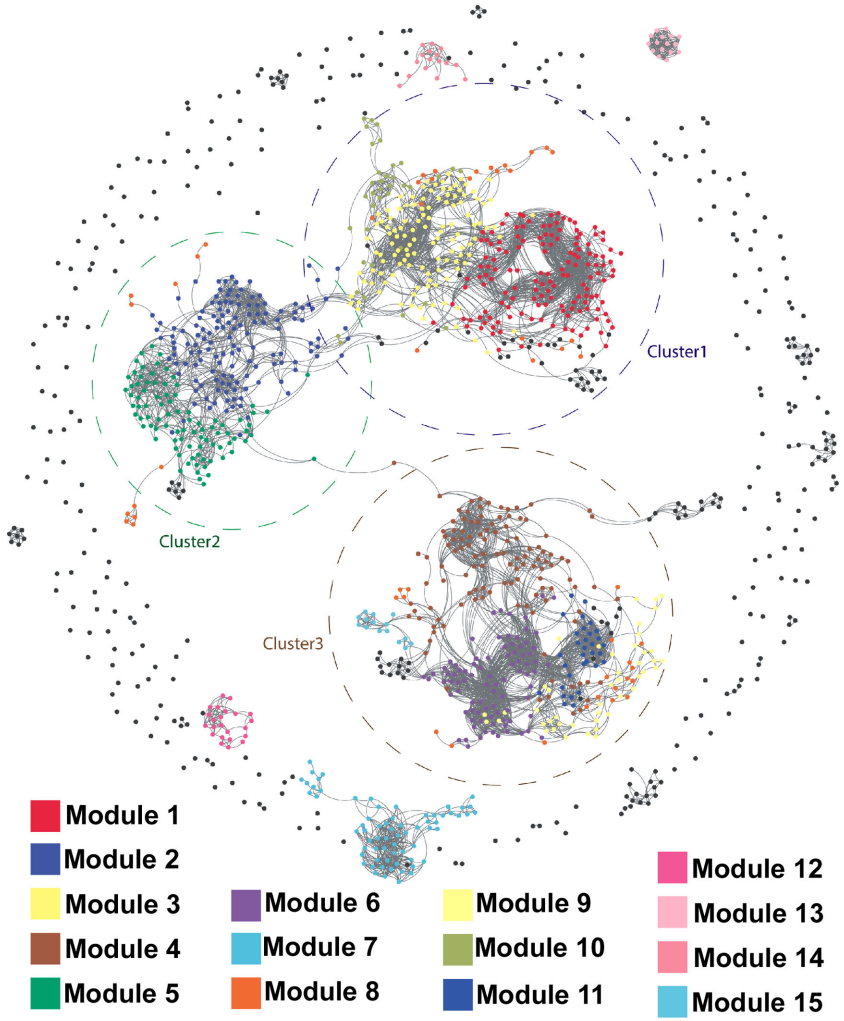

Figure 5. Global network of cooccurring bacterial lineages in *Caulerpa* based on correlation analysis. Each node is an OTU and each link has a minimal spearman correlation value of 0.76 which is the percolation threshold of the network. Nodes are colored according to their module association (numbered from 1 to 15)

links). The average clustering coefficient was 0.45, which reveals a system made of distinct clusters with nodes (OTUs) showing a significant excess of node's neighborhood connectivity, compared to random expectations (<CCo> = 3.3E-03 after 10 000 random simulations). Then, network community algorithm was computed and highlighted through colors on the network (Fig. 5). Fifteen modules are colored as in Fig. S4 (Supporting Information). The size of the modules ranges from 145 to 15 nodes (OTUs).

The level of host taxa specificity of OTU modules emerges from Fig. 6 where nodes of the global network are colored according to their host taxa (Fig. 6A–D). These figures illustrate the cooccurrence of 252, 231, 96 and 21 OTUs, respectively for *C. prolifera*, *C. racemosa* var. *turbinata-uvifera*, *C. racemosa* var. *cylindracea* and *C. taxifolia* .These four networks are superimposed in Fig. 6E to illustrate the inter genera sharing of OTUs and modules. The analysis is strongly impinged by the sampling effort and the alpha diversity within taxa as the two most sampled taxa (*C. prolifera*, *C. racemosa* var. *turbinata-uvifera*), yet the same major modules are emerging when the networks are built per taxa (Fig. S4, Supporting Information) or overall (Fig. 5), and the global figure highlights a strong taxa-specific modularity. Indeed, the majority of OTUs organized in modules are specific to only one taxa. Major modules show however some sharing among *Caulerpa* genera, but limited to very few OTUs. Both recently invasive species show the most limited number of modules, in particular *C. taxifolia*. A single module, one of the smallest, is shared by the four taxa (module 14, in purple and red on Fig. 6E).

Finally, to see if the modularity structure reflects some discrepancies in the taxonomic compositions of the modules, the OTUs diversity was investigated for each of the modules. First, at the phylum level, one can see that the main phyla are all represented inside each of the 15 modules (Fig. 7). With the exception of module 12, Proteobacteria are the most abundant in all the modules. Among the other phyla, Firmicutes, Bacteroidetes and Actinobacteria are among the best represented. At the order level, Burkholderiales has the highest proportion of OTUs in all modules except the 2, 7 and 9 (dominated by Lactobacillales, Alphaproteobacteria and Actinomycetales) where they rank second or third (Table S7, Supporting Information).

DISCUSSION

The results presented here reveal an extreme diversity and a striking specificity of endophytic bacteria associated with *Caulerpa* species within the Mediterranean Sea. The two

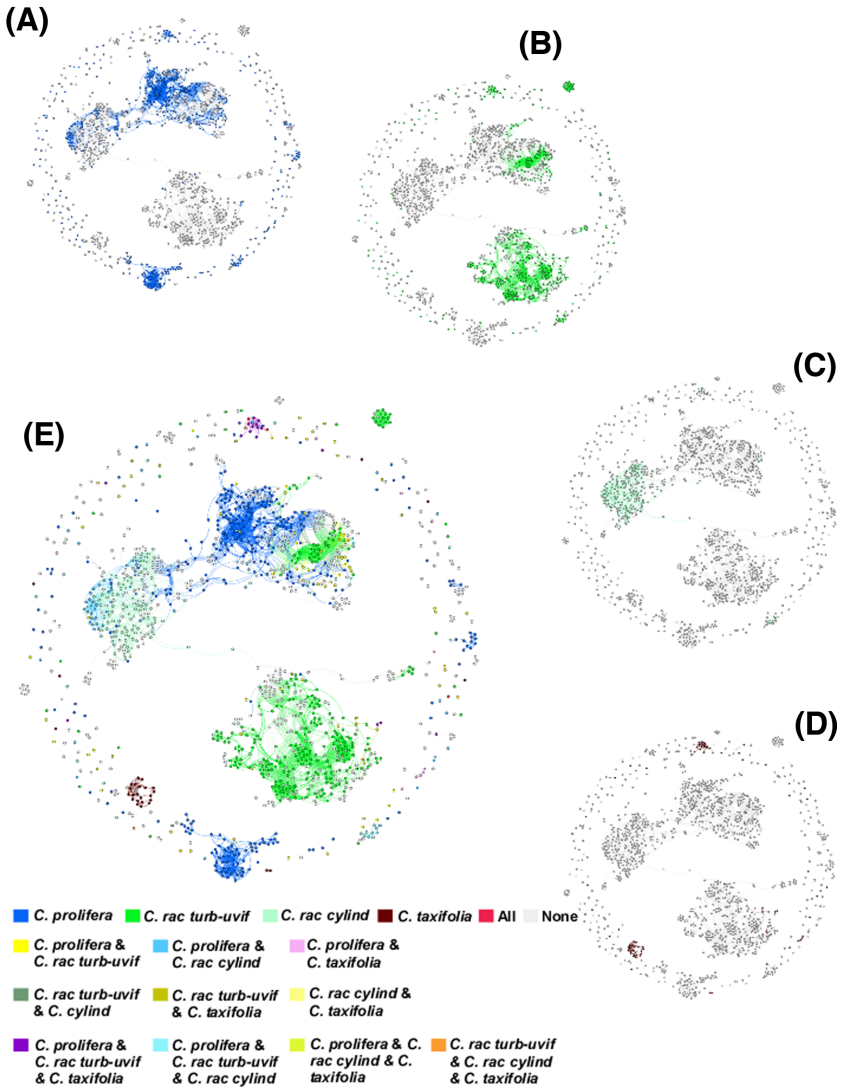

Figure 6. Overlay of global networks. Conserving the same topology of global network in Fig. 5, nodes are colored in **A**, **B**, **C** and **D** according to their belonging to modules observed in each of the four taxa, respectively *C. prolifera*, *C. racemosa* var. *turbinata-uvifera*, *C. racemosa* var. *cylindracea* and *C. taxifolia*. These four networks are superimposed in **E** and nodes shared among taxa are colored according to as given in the legend.

richest taxa, *C. prolifera* and *C racemosa* var. *turbinata*, exhibit an alpha diversity that exceeds the one revealed in sponges or corals (Kvennefors *et al.* 2010; Schmitt *et al.* 2012). The analyses also confirm the systematic differentiation between endophytic and epiphytic bacterial communities (Fig. 1), as recently reported for *C. racemosa* (Aires *et al.* 2013). The distinction between epiphytic and endophytic communities exceeds even the already high differentiation of endophytic communities among taxa. Endophytic communities are significantly (Table S3, Supporting Information) differentiated among, and stable within taxa, providing a diagnostic signature of their host species (Figs 2 and 5). This contrasts with epiphytic communities, the most studied algal bacteria thus far, as they include the biofilm community. In this study, they indeed appear undifferentiated among taxa and largely influenced by the environment, as underlined by their similarity to sediment communities (Fig. 1). Additionally, the large dominance of endophytic communities of taxon-specific OTUs (>70%, Fig. 3) (partly organized in spatially stable modules of cooccurrence; Fig. 5) implies the existence of strong biological filters shaping specific bacterial–host associations resulting in the existence of stable, rich and complex holobionts.

Core community

The existence of a core bacterial community, shared by all four *Caulerpa* taxa, could be explained by lateral transfer of endosymbionts among host species mediated by vectors like grazers (Handeler *et al.* 2010), combined with a selective filter of the algae to maintain or not transferred strains. They may also be explained through coevolution and the vertical inheritance of bacteria during sexual reproduction and asexual proliferation by fragmentation (Hollants *et al.* 2013b). From a similar study on the specificity of Flavobacteriaceae among *Bryopsis*, Hollants et al. (2013b) suggested that species-specific community could have partly evolved along with the host during speciation in which similar traits that select for the uptake of specific bacteria were retained among species (Hollants *et al.* 2013b). Following this rationale, specific bacteria may thus have been retained or discarded during *Caulerpa*'s evolution depending on

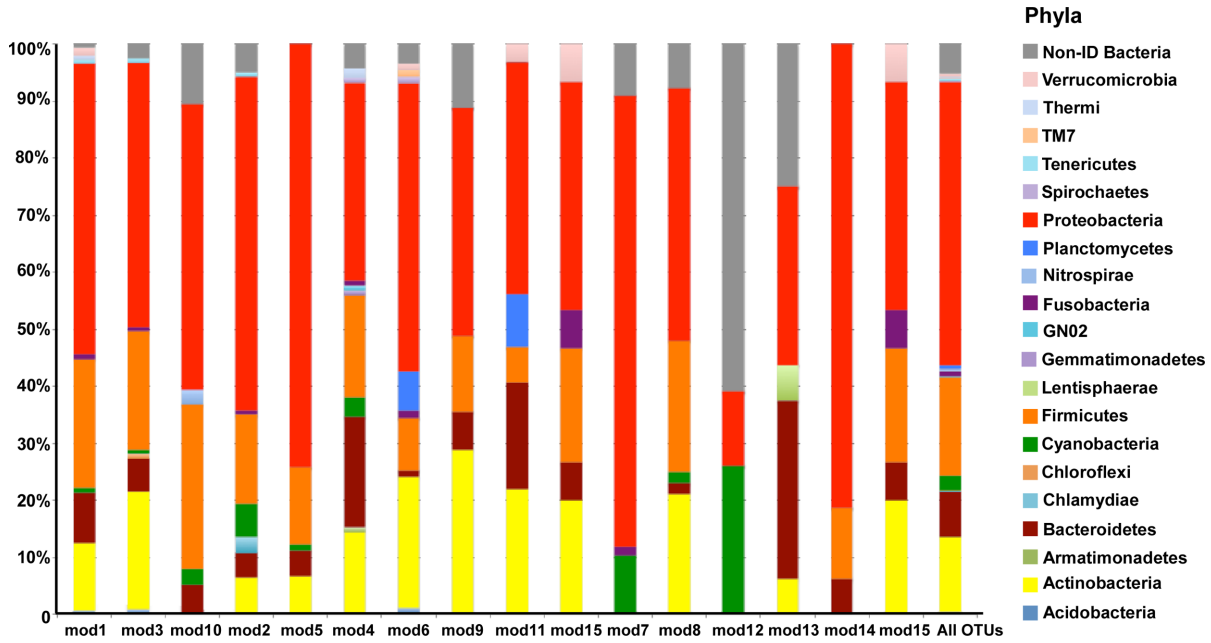

Figure 7. Bacterial diversity contained in modules of the global network: proportion of the main phyla.

physiological needs faced during niche evolution and divergence (Margulis and Fester 1992).

The majority of retrieved sequences showed best hits with the orders Burkholderiales and Pseudomonadales that were consistently represented in all species and locations, suggesting that these phylogenetic groups encompass key members of the endophytic bacterial community in *Caulerpa* sp. (Fig. 3). When sampled in their native range in Australia, *C. racemosa* var. *cylindracea* and *C. taxifolia* also showed a high prevalence of those same OTUs (Aires *et al.* 2013). The highly represented *Burkholderia* has been recognized as endosymbionts in several plant species and mostly nitrogen fixing (Elbeltagy *et al.* 2001; Coenye and Vandamme 2003; Masson-Boivin *et al.* 2009) or enhancers of plant growth (Elbeltagy *et al.* 2001). A bacterial core/shared community consisting of Gammaproteobacteria, CFB group, Alphaproteobacteria, Firmicutes and Actinobacteria was recently defined for all algal groups by screening 161 macroalgal–bacterial studies from the last 55 years Hollants *et al.* (2013a). Contrastingly, the main group found in *the Caulerpa* core community (Burkholderiales) belongs to the non-generalist Betaproteobacteria class. This is in agreement with the hypothesis that these OTUs might be tightly associated with *Caulerpa*, whereas the other conspicuous group, Pseudomonadales (Class Gammaproteobacteria), might be more generalist symbionts found associated with most algal taxa (Hollants *et al.* 2013a).

One of the most ubiquitous OTUs within *Caulerpa* core community was chosen to illustrate the segregation of haplotypes among host and/or habitat. This OTU (table not shown), persistent among all taxa and habitats, showed high similarity (>99%) with a strain identified as *A. denitrificans*. This strain, a nitrate-reducing betaproteobacterium, is described as fully functional in anaerobic conditions (Mechichi *et al.* 2003) often characterizing sediments colonized by *Caulerpa* sp. (Holmer *et al.* 2009). Such high homology leads to speculate that the metabolic function ensured by these bacteria in *Caulerpa* species may increase their performances in eutrophicated/polluted sediments. Moreover, the network of OTU haplotypes (Fig. 4) that are consistently undetected in sediment or epiphytic communities shows a tendency for sequences to display a species-specific pattern (cluster A—*C. prolifera*, B—*C. racemosa*, C—*C. taxifolia*). Besides, haplotypes of this OTU show a closer relation between both *C. racemosa* varieties compared to the other species. Considering that *Caulerpa* taxa have been sampled in sympatry in sites distant by hundreds of kilometers, and this OTU is shared among *C. racemosa* var. *cylindracea* in antipodal Australian and Mediterranean Seas (Aires *et al.* 2013), this haplotype network might thus reflect an early stage of ongoing coevolution between host and part of their endophytic bacterial community.

Species-specific communities and cooccurrence modules

Studies in other host species suggested that the core bacterial community may be important in maintaining essential functions required for the symbiosis to persist (Loudon *et al.* 2014). Yet, depending on the level of divergence of host species, long-term coevolved bacteria such as those putatively present in the common ancestor of those taxa are, depending on selective pressures acting on them, likely to have diverged beyond the clustering threshold of similarity (97%) applied here. Besides, *Caulerpa* sp. may also have developed tight relationship with other species-specific strains essential for their distinct metabolisms. Species-specific bacterial communities are often observed in algae (e.g. Lachnit *et al.* 2009, Barott *et al.* 2011). Here, specific OTUs represent the largest fraction (Fig. 3) of endophytic bacterial communities and cooccurrence modules cluster according to species in almost all locations (Fig. 2A, Fig. 5, Table S3, Supporting Information), showing a limited geographic influence (Fig. 2B, Table S4, Supporting Information), and strikingly different dominant classes of bacteria among the four studied taxa (Fig. 3). Speculation on the putative function of these bacteria based on their BLAST assignment is however suggesting their involvement in similar functions in line with the host environment, including nitrogen and carbon cycling in the

presence of excess of nutrient, possible eutrophication and anoxic conditions (Ceccherelli and Cinelli 1997).

In all the four *Caulerpa* taxa, the main taxa-specific orders are known to include bacterial strains associated with coping mechanisms in disturbed environmental conditions (Reichenbach 1989; Egan *et al.* 2013). *Caulerpa* species are usually associated with disturbed sediments (silt sediments with high levels of nitrites, phosphates and organic matter) which can form toxic conditions for several other species (Chisholm *et al.* 1997; Pérez-Ruzafa *et al.* 2012). The bacterial orders Kiloniellales, Oceanospirillales and Flavobacteriales and Sphingobacteriales (found as the prevalent orders in *C. prolifera*, *C. taxifolia* and both *C. racemosa*, respectively) are all associated with denitrifying, N_2 fixation and nutrient recycling in the presence of high C:N ratios (Kirchman 2002; Chisholm and Moulin 2003; Goffredi *et al.* 2007; Wakelin, Colloff and Kookana 2008; Wiese *et al.* 2009; Goecke *et al.* 2010; Hollants *et al.* 2013b).

By using the formalism of network theory, assemblages of cooccurring OTUs appear as modules in the cooccurrence networks (Figs 5 and 7). That shows the need for such holistic approach to dissect host–bacterial relationships, as these may involve assemblages (modules) of cooccurring OTUs rather than sets of binary relationships between host and inter-independent strains. The interactions that need to be studied, in terms of metabolism and evolution, are thus not the mere pairwise host–bacteria relationship. But instead, the way bacterial communities assemble into consortia likely involved in complementary or synergistic interactions shaping the metabolism of the holobiont as a whole. The important *Caulerpa* taxa specificity of cooccurrence modules suggest a radical switch in the bacterial communities during the course of *Caulerpa* sp. evolution, possibly owing to distinct metabolic pathways required in the slightly different niches occupied by those four taxa. Each of the four *Caulerpa* holobionts revealed here can be considered as exhibiting its own metabolic machinery, with ancient symbiosis or rare events of leaks followed by specific divergence, resulting in the few shared OTUs (Figs 3 and 6). The hypothesis of occasional lateral transfer in sympatry followed by capture and possible coevolution has been supported in other eukaryotic lineages, such as vesicomyid clams and their bacterial chemosynthetic endosymbionts (Stewart, Young and Cavanaugh 2008; Decker *et al.* 2013). Similar events of occasional leaks are in line with the observation of a larger amount of shared OTUs between phylogenetically more distant taxa (*C. prolifera* and *C. racemosa* var. *turbinata uvifera*) cooccurring in sympatry since the longest time frame, as the former is native and the latter suspected to have been present for at least more than a century in the Mediterranean (Ollivier 1929; Rayss 1941). Contrastingly, *C. racemosa* var. *cylindracea* and *C. taxifolia* that recently invaded the Mediterranean show much lower diversities and higher specificity of their bacterial assemblages (Figs 3 and 6).

CONCLUSIONS

This study reveals the siphonophous green algae of the genus *Caulerpa* as a model of a complex and rich holobiont. Many mechanisms of acquisition and evolution of endosymbionts appear to be able to coexist in the holobiont system. Plausible scenarios include (1) long-term coevolution of some common lineages resulting in species-specific haplotypes, (2) the sporadic events of lateral transfer of bacteria possibly favored by long-term coexistence in sympatry and (3) the capture of environmental bacteria that may become integrated in the holobiont on a longer term. Besides species specificity of OTUs or haplotypes within shared OTUs, the cooccurrence of species-specific modules suggest that some bacterial strains may form metabolic consortia that would contribute to shaping the structure of bacterial communities (Barberán *et al.* 2012) and may also accelerate diversification among species. Different species-specific communities might thus have similar metabolic functions despite their different taxonomy, as suggested by Burke *et al.* (2011b). The results reported here pave the road for future research to test those hypotheses and advance our understanding of the ecology and evolution of these holobionts. Functional metagenomics and metatranscriptomics analysis of *Caulerpa* sp. may allow testing the hypothesis that, regardless of community composition, different bacterial associations would represent different 'machinery' performing similar metabolic processes in distinct species. The discrimination of bacterial assemblages associated with gametes may also allow reconsidering the ancientness and frequency of strict coevolutionary patterns as well as the relative role of vertical versus lateral acquisition in the evolution of these rich holobionts. Finally, these studies will help in understanding the role of bacterial communities on the success and invasive trajectories of their host, as during the past decades several *Caulerpa* taxa, including three of the four studied here, have extended their range through invasion in different parts of the world, including the Mediterranean.

FUNDING

This work was supported by the Portuguese Foundation for Science and Technology (FCT) and FEDER with the project IBISA (PTDC / MAR / 64749 / 2006) and a PhD fellowship (TA) [SFRH/BD/30043/2006] from FCT and FSE.

SUPPLEMENTARY DATA

Supplementary data is available at FEMSEC online.

ACKNOWLEDGEMENTS

We thank Gary Kendrick, Renae Hovey and Jeff for field work and all the logistic help; Professor Craig Atkins for lab facilities supplying; Frédéric Zuberer, David Luquet and Habib Langar (from the INSTM of Salâmbo, Tunis) for their invaluable help in sampling *Caulerpa* in Villefranche, Marseille and Tunis; Regino Martínez and Núria Marbà for help in sampling in Mallorca; R. Cunha for her valuable comments, and BIOCANT center for metagenomics analysis and support in bioinformatic issues.

Conflict of interest. None declared.

REFERENCES

Aires T, Marbà N, Serrao EA, *et al.* Selective elimination of chloroplastidial DNA for metagenomics of bacteria associated with the green alga Caulerpa Taxifolia (Bryopsidophyceae)1. *J Phycol* 2012;**48**:483–90.

Aires T, Serrão EA, Kendrick G, *et al.* Invasion is a community affair: Clandestine followers in the bacterial community associated to green algae, Caulerpa racemosa, track the invasion source. *PLoS One* 2013;**8**:e68429.

Barberán A, Bates ST, Casamayor EO, *et al.* Using network analysis to explore co-occurrence patterns in soil microbial communities. *ISME J* 2012;**6**:343–51.

Barott KL, Rodriguez-Brito B, Janouškovec J, *et al.* Microbial diversity associated with four functional groups of benthic reef algae and the reef-building coral Montastraea annularis. *Environ Microbiol* 2011;**13**:1192–204.

Bastian M, Heymann S, Jacomy M. Gephi: An open source software for exploring and manipulating networks. In: *Proceedings 3rd International AAAI Conference on Weblogs and Social Media*, San Jose, California, 2009.

Batagelj V, Mrvar A. Pajek - Analysis and visualization of large networks. *Graph Drawing* 2002;**2265**:477–8.

Blondel VD, Guillaume J-L, Lambiotte R, *et al.* Fast unfolding of communities in large networks. *J Stat Mech Theory E* 2008;**2008**:P10008.

Burke C, Thomas T, Lewis M, *et al.* Composition, uniqueness and variability of the epiphytic bacterial community of the green alga Ulva australis. *ISME J* 2011;**5**:590–600.

Caporaso JG, Kuczynski J, Stombaugh J, *et al.* Correspondence QIIME allows analysis of high- throughput community sequencing data Intensity normalization improves color calling in SOLiD sequencing. *Nat Publ Gr* 2010;**7**:335–6.

Ceccherelli G, Cinelli F. Short-term effects of nutrient enrichment of the sediment and interactions between the seagrass Cymodocea nodosa and the introduced green alga Caulerpa taxifolia in a Mediterranean bay. *J Exp Mar Bio Ecol* 1997;**217**:165–77.

Chandler JA, Morgan Lang J, Bhatnagar S, *et al.* Bacterial Communities of Diverse *Drosophila* Species: Ecological Context of a Host–Microbe Model System. *PLoS Genet* 2011;**7**:e1002272, DOI:10.1371/journal.pgen.1002272.

Chao A. Nonparametric estimation of the number of Classes in a population. *Scand J Stat* 1984;**11**:265–70.

Chisholm JRM, Dauga C, Ageron E et al.'Roots' in mixotrophic algae. *Nature* 1996;**381**:382.

Chisholm JRM, Fernex FE, Mathieu D, *et al.* Wastewater discharge, seagrass decline and algal proliferation on the Cote d'Azur. *Mar Pollut Bull* 1997;**34**:78–84.

Chisholm JRM, Moulin P. Stimulation of nitrogen fixation in refractory organic sediments by Caulerpa taxifolia (Chlorophyta). *Limnol Oceanogr* 2003;**48**:787–94.

Coenye T, Vandamme P. Diversity and significance of Burkholderia species occupying diverse ecological niches. *Environ Microbiol* 2003;**5**:719–29.

Costello E, Lauber C, Hamady M. Bacterial community variation in human body habitats across space and time. *Science (80–)* 2009;**326**:1694–7.

Croft MT, Warren MJ, Smith AG. Algae need their vitamins. *Eukaryot Cell* 2006;**5**:1175–83.

Crump BC, Koch EW. Attached bacterial populations shared by four species of aquatic angiosperms. *Appl Environ Microb* 2008;**74**:5948–57.

Csárdi G, Nepusz T. The igraph software package for complex network research. *Inter Journal Complex Syst* 2006;**1695**:1695.

Dawes CJ, Lohr CA. Cvtoplasmic organization and endosymbiotic bacteria in the growing points of Caulerpa prolifera. *Rev Algol* 1978;**13**:309–14.

Decker C, Olu K, Arnaud-Haond S, *et al.* Physical proximity may promote lateral acquisition of bacterial symbionts in vesicomyid clams. *PLoS One* 2013;**8**:e64830.

Delbridge L, Coulburn J, Fagerberg W, *et al.* Community profiles of bacterial endosymbionts in four species of Caulerpa. *Symbiosis* 2004;**37**:335–44.

Dittami SM, Eveillard D, Tonon T. A metabolic approach to study algal-bacterial interactions in changing environments. *Mol Ecol* 2014;**23**:1656–60.

Egan S, Harder T, Burke C, *et al.* The seaweed holobiont: understanding seaweed-bacteria interactions. *FEMS Microbiol Rev* 2013;**37**:462–76.

Elbeltagy A, Nishioka K, Sato T, *et al.* Endophytic colonization and in planta nitrogen fixation by a Herbaspirillum sp. isolated from wild rice species. *Appl Environ Microb* 2001;**67**:5285–93.

Faust K, Raes J. Microbial interactions: from networks to models. *Nat Rev Microbiol* 2012;**10**:538–50.

Fortunato S. Community detection in graphs. *Phys Rep* 2010;**486**:75–174.

Forster P, Torroni A, Renfrew C, *et al.* Phylogenetic star contraction applied to Asian and Papuan mtDNA evolution. *Mol Biol Evol* 2001;**18**:1864–81.

Gihring TM, Zhang G, Brandt CC, *et al.* A limited microbial consortium is responsible for extended bioreduction of uranium in a contaminated aquifer. *Appl Environ Microb* 2011;**77**:5955–65.

Goecke F, Labes A, Wiese J, *et al.* Chemical interactions between marine macroalgae and bacteria. *Mar Ecol Prog Ser* 2010;**409**:267–99.

Goffredi SK, Johnson SB, Vrijenhoek RC. Genetic diversity and potential function of microbial symbionts associated with newly discovered species of Osedax polychaete worms. *Appl Environ Microb* 2007;**73**:2314–23.

Haas BJ, Gevers D, Earl AM, *et al.* Chimeric 16S rRNA sequence formation and detection in Sanger and 454-pyrosequenced PCR amplicons. *Genome Res* 2011;**21**:494–504.

Hamel H. Quelques algues rares ou nouvelles pour la flore mediterraneenne. *Bull Mus Nat Sci Nat* 1926;**32**:420.

Hammer Ø, Harper DAT, Ryan PD. Paleontological statistics software package for education and data analysis. *Palaeontol Electron* 2001;**4**:9–18.

Handeler K, Wagele H, Wahrmund U, *et al.* Slugs' last meals: molecular identification of sequestered chloroplasts from different algal origins in Sacoglossa (Opisthobranchia, Gastropoda). *Mol Ecol Resour* 2010;**10**:968–78.

Hollants J, Decleyre H, Leliaert F, *et al.* Life without a cell membrane: challenging the specificity of bacterial endophytes within Bryopsis (Bryopsidales, Chlorophyta). *BMC Microbiol* 2011a;**11**:255.

Hollants J, Leliaert F, De Clerck O, *et al.* What we can learn from sushi: a review on seaweed-bacterial associations. *FEMS Microbiol Ecol* 2013a;**83**:1–16.

Hollants J, Leliaert F, Verbruggen H, *et al.* Permanent residents or temporary lodgers: characterizing intracellular bacterial communities in the siphonous green alga Bryopsis. *Proc Biol Sci* 2013b;**280**:20122659.

Hollants J, Leroux O, Leliaert F, *et al.* Who is in there? Exploration of endophytic bacteria within the siphonous green seaweed Bryopsis (Bryopsidales, Chlorophyta). *PLoS One* 2011b;**6**:e26458.

Hollister EB, Engledow AS, Hammett AJM, *et al.* Shifts in microbial community structure along an ecological gradient of hypersaline soils and sediments. *ISME J* 2010;**4**:829–38.

Holmer M, Marbà N, Lamote M, *et al.* Deterioration of sediment quality in seagrass meadows (*Posidonia oceanica*) invaded by macroalgae (Caulerpa sp.). *Estuar Coast* 2009;**32**:456–66.

Huber JA, Mark Welch DB, Morrison HG, *et al.* Microbial population structures in the deep marine biosphere. *Science* 2007;**318**:97–100.

Janson EM, Stireman JO, Singer MS, *et al.* Phytophagous insect-microbe mutualisms and adaptive evolutionary diversification. *Evolution* 2008;**62**:997–1012.

Jiao JY, Wang HX, Zeng Y, *et al.* Enrichment for microbes living in association with plant tissues. *J Appl Microbiol* 2006;**100**: 830–7.

Joint I, Tait K, Wheeler G. Cross-kingdom signalling: exploitation of bacterial quorum sensing molecules by the green seaweed Ulva. *Philos T Roy Soc B* 2007;**362**:1223–33.

Keshtacher-Liebson E, Hadar Y, Chen Y. Oligotrophic bacteria enhance algal growth under iron-deficient conditions. *Appl Environ Microb* 1995;**61**:2439–41.

Kirchman DL. The ecology of Cytophaga-Flavobacteria in aquatic environments. *FEMS Microbiol Ecol* 2002;**39**:91–100.

Kivelä M, Arnaud-Haond S, Saramäki J. EDENetworks: A user-friendly software to build and analyse networks in biogeography, ecology and population genetics. *Mol Ecol Resour* 2015;**15**:117–22.

Krych L, Hansen CHF, Hansen AK, *et al.* Quantitatively different, yet qualitatively alike: a meta-analysis of the mouse core gut microbiome with a view towards the human gut microbiome. *PLoS ONE* 2013;**8**:e62578.

Kvennefors ECE, Sampayo E, Ridgway T, *et al.* Bacterial communities of two ubiquitous great barrier reef corals reveals both site-and species-specificity of common bacterial associates. *PLoS One* 2010;**5**:e10401, DOI:10.1371/journal.pone.0010401.

Lachnit T, Blümel M, Imhoff JJF, *et al.* Specific epibacterial communities on macroalgae: phylogeny matters more than habitat. *Aquat Biol* 2009;**5**:181–6.

Lachnit T, Meske D, Wahl M, *et al.* Epibacterial community patterns on marine macroalgae are host-specific but temporally variable. *Environ Microbiol* 2011;**13**:655–65.

Lancichinetti A, Fortunato S. Community detection algorithms: a comparative analysis. *Phys Rev E* 2009;**80**:056117.

Langar H, Djellouli A, Ben Mustapha K, *et al.* Première signalisation de Caulerpa taxifolia (Vahl) C. Agardh en Tunisie. *Bull Inst Sci Tech Mer* 2000;**27**:7–8.

Lee OO, Chui PY, Wong YH, *et al.* Evidence for vertical transmission of bacterial symbionts from adult to embryo in the Caribbean sponge Svenzea zeai. *Appl Environ Microb* 2009;**75**:6147–56.

Loudon AH, Woodhams DC, Parfrey LW, *et al.* Microbial community dynamics and effect of environmental microbial reservoirs on red-backed salamanders (Plethodon cinereus). *ISME J* 2014;**8**:830–40.

Lozupone C, Knight R. UniFrac: a new phylogenetic method for comparing microbial communities. *Appl Environ Microb* 2005;**71**:8228–35.

Margulis L, Fester R. *Symbiosis as a Source of Evolutionary Innovation*. Cambridge, MA, USA: The MIT Press, 1992.

Masson-Boivin C, Giraud E, Perret X, *et al.* Establishing nitrogen-fixing symbiosis with legumes: how many rhizobium recipes? *Trends Microbiol* 2009;**17**:458–66.

Matsuo Y, Imagawa H, Nishizawa M, *et al.* Isolation of an algal morphogenesis. 2005;**307**:2005.

Mechichi T, Stackebrandt E, Fuchs G. Alicycliphilus denitrificans gen. nov., sp. nov., a cyclohexanol-degrading, nitrate-reducing beta-proteobacterium. *Int J Syst Evol Micr* 2003;**53**:147–52.

Meinesz A, Hesse B. Introduction et invasion de l'algue tropicale Caulerpa taxifolia en Méditerranée nord-occidentale. *Oceanol Acta* 1991;**14**:415–26.

Meusnier I, Olsen JL, Stam WT, *et al.* Phylogenetic analyses of Caulerpa taxifolia (Chlorophyta) and of its associated bacterial microflora provide clues to the origin of the Mediterranean introduction. *Mol Ecol* 2001;**10**:931–46.

Moalic Y, Desbruyères D, Duarte CM, *et al.* Biogeography revisited with network theory: retracing the history of hydrothermal vent communities. *Syst Biol* 2012;**61**:127–37.

Moore AM, Munck C, Sommer MO, *et al.* Functional metagenomic investigations of the human intestinal microbiota. *Front Microbiol* 2011;**2**:1–8.

Myshrall KL, Mobberley JM, Green SJ, *et al.* Biogeochemical cycling and microbial diversity in the thrombolitic microbialites of Highborne Cay, Bahamas. *Geobiology* 2010;**8**: 337–54.

Nakanishi K, Nishijima M. Bacteria that induce morphogenesis in *Ulva pertusa* (chlorophyta). 1996;**32**:479–82.

Newman M. The structure and function of complex networks. *SIAM Rev* 2003;**45**:167–256.

Newman MEJ. Finding community structure in networks using the eigenvectors of matrices. *Phys Rev E* 2006;**74**: 036104.

Nizamuddin M. *The Green Marine Algae of Libya*. Bern: Elga, 1991.

Oksanen J, Blanchet FG, Kindt R, *et al. Vegan: Community Ecology Package*, 2011. http://cran.r-project.org/web/packages/vegan/index.html (26 June 2015, date last accessed).

Oliveros JC. *VENNY. An Interactive Tool for Comparing Lists with Venn Diagrams*, 2007. http://bioinfogp.cnb.csic.es/tools/venny/index.html (26 June 2015, date last accessed).

Ollivier G. *Etude de la flore marine de la Côte d'Azur*. Monaco, Paris: Ann Inst Océanogr, 1929;**8**(3).

Orole O, Adejumo T. Bacterial and fungal endophytes associated with grains and roots of maize. *J Ecol Nat Environ* 2011;**3**:298–303.

Panayotidis P. On the enigmatic origin of the Mediterranean invasive Caulerpa racemosa (Caulerpales, Chlorophyta). *Mediterr Mar Sci* 2006;**7**:119–21.

Pérez-Ruzafa A, Marcos C, Bernal CM, *et al. Cymodocea nodosa* vs. *Caulerpa prolifera*: causes and consequences of a long term history of interaction in macrophyte meadows in the mar Menor coastal lagoon (Spain, Southwestern Mediterranean). *Estuar Coast Shelf S* 2012;**110**:101–15.

Price MN, Dehal PS, Arkin AP. FastTree 2—Approximately maximum-likelihood trees for large alignments. *PLoS One* 2010;**5**:e9490, DOI:10.1371/journal.pone.0009490.

Polzin T, Daneschmand SV. On Steiner trees and minimum spanning trees in hypergraphs. *Oper Res Lett* 2003;**31**:12–20.

Raghavan UN, Albert R, Kumara S. Near linear time algorithm to detect community structures in large-scale networks. *Phys Rev E* 2007;**76**:1–12.

Rayss T. Sur les Caulerpes de la côte palestinienne. *Palest J Bot Jerus Ser* 1941;**2**:103–24.

Reichenbach H. Order I. Cytophagales Leadbetter 1974, 99AL. In: Staley JT, Bryant MP, Pfennig N et al. (ed). *Bergey's Manual of Systematic Bacteriology*. Baltimore, MD: Williams & Wilkins, 1989;**3**:2011–3.

Reshef L, Koren O, Loya Y, *et al.* The coral probiotic hypothesis. *Environ Microbiol* 2006;**8**:2068–73.

Rosenberg E, Koren O, Reshef L, *et al.* The role of microorganisms in coral health, disease and evolution. *Nat Rev Microbiol* 2007;**5**:355–62.

Rosvall M, Bergstrom CT. Community structure in directed networks. *P Natl Acad Sci USA* 2008;**105**:1118–23.

Sachs JL, Skophammer RG, Regus JU. Evolutionary transitions in bacterial symbiosis. *P Natl Acad Sci USA* 2011;**108**(Suppl):10800–7.

Schmitt S, Tsai P, Bell J, *et al.* Assessing the complex sponge microbiota: core, variable and species-specific bacterial communities in marine sponges. *ISME J* 2012;**6**:564–76.

Stewart FJ, Young CR, Cavanaugh CM. Lateral symbiont acquisition in a maternally transmitted chemosynthetic clam endosymbiosis. *Mol Biol Evol* 2008;**25**:673–87.

Tonon T, Eveillard D, Prigent S, *et al*. Toward systems biology in brown algae to explore acclimation and adaptation to the shore environment. *OMICS* 2011;**15**:883–92.

Tujula NA, Crocetti GR, Burke C, *et al*. Variability and abundance of the epiphytic bacterial community associated with a green marine Ulvacean alga. *ISME J* 2010;**4**:301–11.

Turnbaugh PJ, Ridaura VK, Faith JJ, *et al*. Metagenomic analysis in humanized gnotobiotic mice. 2010;**1**:1–19.

Verlaque M, Boudouresque CF, Meinesz A, *et al*. The Caulerpa racemosa complex (Caulerpales, Ulvophyceae) in the Mediterranean Sea. *Bot Mar* 2000;**43**:49–68.

Wakelin SA, Colloff MJ, Kookana RS. Effect of wastewater treatment plant effluent on microbial function and community structure in the sediment of a freshwater stream with variable seasonal flow. *Appl Environ Microb* 2008;**74**:2659–68.

Wang Q, Garrity GM, Tiedje JM, *et al*. Naive Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. *Appl Environ Microb* 2007;**73**:5261–7.

Weinberger F, Beltran J, Correa JA, *et al*. Spore release in Acrochaetium sp. (Rhodophyta) is bacterially controlled. *J Phycol* 2007;**43**:235–41.

Wiese J, Thiel V, Gärtner A, *et al*. Kiloniella laminariae gen. nov., sp. nov., an alphaproteobacterium from the marine macroalga Laminaria saccharina. *Int J Syst Evol Micr* 2009;**59**:350–6.

Zilber-Rosenberg I, Rosenberg E. Role of microorganisms in the evolution of animals and plants: the hologenome theory of evolution. *FEMS Microbiol Rev* 2008;**32**:723–35.