Permanent residents or temporary lodgers: characterizing intracellular bacterial communities in the siphonous green alga Bryopsis.

Joke Hollants, Frederik Leliaert, Heroen Verbruggen, Anne Willems, Olivier De Clerck.

Proc R Soc B (2013) 280: 20122659

DOI 10.1098/rspb.2012.2659

Permanent residents or temporary lodgers: characterizing intracellular bacterial communities in the siphonous green alga *Bryopsis*

Joke Hollants[a,b,1], Frederik Leliaert[b,1], Heroen Verbruggen[c], Anne Willems[a2], and Olivier De Clerck[b2]

[a] Laboratory of Microbiology, Department of Biochemistry and Microbiology, Ghent University, K.L. Ledeganckstraat 35, 9000 Ghent, Belgium
[b] Phycology Research Group, Department of Biology, Ghent University, Krijgslaan 281 (S8), 9000 Ghent, Belgium
[c] School of Botany, University of Melbourne, Victoria 3010, Melbourne, Australia
[1] JH and FL contributed equally to this work

[2] Corresponding authors
Anne Willems
E-mail address: anne.willems@UGent.be

Olivier De Clerck
E-mail address: olivier.declerck@UGent.be

Abstract

The ecological success of giant celled, siphonous green algae in coastal habitats has repeatedly been linked to endophytic bacteria living within the cytoplasm of the hosts. Yet, very little is known about the relative importance of evolutionary and ecological factors controlling the intracellular bacterial flora of these seaweeds. Using the marine alga *Bryopsis* (Bryopsidales, Chlorophyta) as a model, we explore the diversity of the intracellular bacterial communities and investigate whether their composition is controlled by ecological and biogeographical factors rather than the evolutionary history of the host. Using a combination of 16S rDNA clone libraries and DGGE analyses, we show that *Bryopsis* harbors a diverse mixture of bacteria. Variation partitioning analyses show a strong impact of local environmental factors on bacterial community composition for generalist species, while specialists reflect a predominant imprint of evolutionary history. The results highlight the importance of interpreting the presence of individual bacterial phylotypes in the light of ecological and evolutionary principles such as phylogenetic niche conservatism to understand complex endobiotic communities and the parameters shaping them.

Keywords: algae, bacteria, biogeography, endosymbiosis, seaweed, variation partitioning

Introduction

Variation in traits across species or populations is influenced by their ecology and evolutionary history [1]. Organisms are shaped by the environment in which they live, with species residing in similar environments having common adaptations [2]. They are also the product of their evolutionary history, and closely related species have the tendency to be more similar than distantly related species [3]. This tendency for related species to resemble each other more in a trait than expected by chance is referred to as phylogenetic signal or phylogenetic conservatism [4]. Applying these principles to host-bacterial relationships, one might presume that obligate, vertically inherited bacteria (specialists) are phylogenetically structured, while facultative endobiotic bacteria (generalists) are expected to be more randomly dispersed among host species [5] (Fig. 1). In this study, we assess for the first time the combined effect of host dependency, ecology and biogeography on the structure of a complex endobiotic community in an algal model.

Marine macroalgae (seaweeds) are commonly associated with bacteria that either live on the surface or in the cytoplasm and/or vacuolar systems of the cells [6-8]. These bacteria are able to influence the morphogenesis and life cycle of their algal host [9-11] and are linked with various metabolic functions such as the production of growth factors, fixed nitrogen and antimicrobial compounds [12-14]. Siphonous green seaweeds, consisting of a single giant tubular cell, form a benevolent biotic environment for endobiotic bacterial communities [15-17]. The siphonous cells, which range from centimeters to meters in length, typically exhibit vigorous cytoplasmic streaming to transport organelles, photosynthates and nutrients [18]. Chisholm *et al.* [19]

demonstrated that siphonous algae take up nutrients from the sediment by a root-like system containing intracellular bacteria and translocate them throughout the thallus. These cellular innovations alongside unique mechanisms of wounding response [20, 21] and the close interactions with bacteria may provide a physiological explanation for the successful spread of invasive siphonous green algae such as *Caulerpa* and *Codium* in marine coastal habitats [19, 22, 23].

Very little is known about the factors controlling the presence of bacteria inside siphonous seaweeds. Two host-related mechanisms may affect the intracellular bacterial composition. Firstly, siphonous seaweeds readily regenerate from protoplasts, facilitating environmental uptake of bacteria into the cell [24, 25]. Secondly, endogenous bacteria can persist by vertical inheritance through gametes [26]. Beside the question of whether the endobionts are acquired vertically or horizontally from the environment, ecological parameters and geographic aspects may also need to be considered to explain the bacterial composition, as some bacteria (or hosts) are likely to be geographically restricted or occur only in particular niches. Although a previous study suggested that seaweed-associated bacterial communities are biogeographically structured [23], it is not known whether ecological or historical factors cause this structure.

The goal of this study is to investigate the relative roles of host, environment and geography in determining the intracellular bacterial flora of siphonous seaweeds, focusing on the genus *Bryopsis* (Bryopsidales, Chlorophyta) as a case study. This genus is known to harbor several types of endogenous bacteria and protocols are in place to study them [17, 27]. *Bryopsis* is known to possess mechanisms for environmental uptake as well as vertical inheritance of

bacteria [25, 26]. This combination of features, combined with the large collection of available cultures, makes the genus an ideal case study to address our goal. The experimental approach consisted of molecular characterization of host samples and their intracellular bacterial flora. The molecular identification of bacterial phylotypes, along with the host phylogeny and environmental data, were explored and analyzed with statistical techniques designed to disentangle the effects of host phylogeny, geography and the external environment on the intracellular bacterial composition.

Methods

Algal material

The 20 *Bryopsis* samples analyzed in this study are listed in Table S1 and their sampling sites are depicted in Figure S1. All samples were transferred to and maintained as unialgal cultures under the conditions described by Hollants *et al.* [17].

Molecular approach

Bryopsis samples were subjected to a surface sterilization step to eliminate epiphytic bacterial contamination [27] prior to total DNA extraction [28]. The host *rbc*L and bacterial 16S rRNA genes were PCR amplified as described by Hollants *et al.* [17]. The endophytic bacterial diversity was assessed by creating 16S rRNA gene clone libraries and performing nested PCR denaturing gradient gel electrophoresis (DGGE) analyses as described previously [17, 25]. Sequences were submitted to EMBL under accession numbers HE648924-HE648948.

Sequence data analyses

*Bryopsis rbc*L and bacterial 16S rRNA gene sequences were assembled, checked for chimeras, compared with nucleotide databases and aligned as previously described [17]. Phylogenetic trees were inferred with maximum likelihood (ML) implemented in PhyML v3.0 [29] and Bayesian inference (BI) using MrBayes [30], via the University of Oslo Bioportal website [31]. Both analyses were performed under a HKY+G model as determined by the Akaike Information Criterion in JModeltest v0.1.1 [32]. Bacterial phylotypes or operational taxonomic units (OTUs) were identified based on 97% sequence similarity.

Statistical analysis

The influence of environmental, geographic, and host phylogenetic factors on the endophytic bacterial diversity in *Bryopsis* was analyzed using multivariate statistical and comparative phylogenetic approaches. The response table was represented by a presence/absence matrix of the seven bacterial phylotypes in the 20 host samples (Fig. 2). The three explanatory matrices (environment, geography and phylogeny) were prepared as follows. The environmental component was represented by seven macro-ecological variables (see Fig. 2) extracted from Bio-ORACLE [33]. The geographic component was represented by a set of orthogonal spatial variables extracted from geographic coordinates by Moran's Eigenvector Maps (MEM) analysis [34] using 'codep' in R [35]. The geographic matrix was represented by the first two eigenvectors, which were the only ones having positive eigenvalues (6.54 and 1.52). The phylogenetic component was expressed as principal coordinates via a principal coordinate

analysis (PCoA) [36] computed from a distance matrix [37]. A corrected distance matrix of the *Bryopsis rbc*L alignment was calculated in MEGA [38]; the PCoA analysis was performed in PCO [39]. The phylogenetic matrix was represented by the first four principal coordinates, representing 98% of the total variation.

To study the influence of environment, geography and host phylogeny on the endophytic bacterial diversity, we first performed data ordinations and calculated phylogenetic signals of the bacterial community composition. Ordination of *Bryopsis* samples based on endophytic bacterial community composition was performed using a principal component analysis (PCA) in CANOCO for Windows 4.5 [40]. Environmental variables were plotted on the PCA graph as supplementary information. Phylogenetic signal was assessed for (i) the environmental variables, (ii) geography, (iii) the total endophytic bacterial community (i.e. represented by principal components 1 and 2 calculated as described above) and (iv) the presence/absence of each of the endophytic bacterial OTUs separately. P-values were calculated using randomizations of the K-statistic [41] in the R package Picante [42] (for i - iii) and the D statistic [43] in the R package 'caper' (http://cran.r-project.org/web/packages/ caper/) (for iv). We quantified the common and unique influences of host phylogeny, geography and environment on the endophytic flora variation using variation partitioning analyses [2, 44] using the varpart function in the R package 'vegan'. The total bacterial diversity, as well as presence/absence data of the seven individual phylotypes was considered as response tables. We performed variation partitioning analyses using three (phylogeny, environment and geography) and two (phylogeny, environment) explanatory tables, respectively.

Results

Bryopsis host phylogeny

Based on the phylogenetic analysis of host *rbc*L sequences (Fig. 2) we assigned the seaweed samples to nine *Bryopsis* species, numbered sp. 1 through 9. The host phylogeny shows three main clades. Clades A and B include *Bryopsis* samples isolated from cold to temperate regions, whereas clade C is warm-temperate to tropical. The phylogenetic signal in annual mean sea surface temperature, as well as annual mean photosynthetically available radiation and dissolved oxygen levels, which are inversely proportional to each other, is statistically significant ($P < 0.01$, Table S2), suggesting that the structure of the *Bryopsis* phylogeny reflects temperature-related environmental variables. Conversely, geographic location (represented by Moran's Eigenvector Maps) did not show a significant phylogenetic structure (Table S2).

Endophytic bacterial diversity

The results from the clone libraries and DGGE analyses showed the presence of seven unique endophytic bacterial phylotypes or operational taxonomic units (OTUs) within *Bryopsis* (Table S3). Five could be identified as Flavobacteriaceae (OTU-1), *Mycoplasma* (OTU-2), Bacteroidetes (OTU-3), Phyllobacteriaceae (OTU-4) and *Labrenzia* (OTU-7) species, which were previously shown to occur in *Bryopsis* [17] (Table S3, Fig. S2). In addition, two new endophytic phylotypes were identified, OTU-5 and OTU-6 (Table S3, Fig. S2). OTU-5 showed high sequence similarities with Rhizobiaceae strains isolated from root nodules of leguminous plants, and represents two distinct clusters that include *Rhizobium leguminosarum* and *Ensifer meliloti* type strains,

144 respectively. OTU-6 is allied to uncultured Rickettsiales bacteria associated with the coral
145 *Montastraea faveolata* and the marine ciliate *Diophrys appendiculata*. All OTU-6 sequences
146 formed a distinct and well-supported clade closely related to the genus *Rickettsia* and most
147 likely represent at least a new species based on their low sequence similarities (≤93%) with
148 *Rickettsia* type strains.
149
150 *Endophytic bacterial composition*
151 Figure 2 schematizes the endophytic bacterial diversity (blue boxes) in *Bryopsis*. Composition of
152 the endophytic community varied between host species, and samples from the same host
153 species harbored diverse combinations of one to four different endophytic phylotypes. Different
154 host species with the same geographic origin commonly displayed differences in their
155 intracellular bacterial community composition (e.g. samples MZ1 and MZ4). This apparent lack
156 of correlation between total bacterial diversity and *Bryopsis* host species and geography is
157 confirmed by the PCA plot which illustrates that the ordination of the different *Bryopsis* species
158 is not fully explained by their similarity in endophytic bacterial community composition (Fig. 3).
159 This PCA plot, however, clearly indicates a correlation between the presence of individual
160 endophytic phylotypes and certain environmental variables. Flavobacteriaceae, Bacteroidetes
161 and *Mycoplasma* endophytes were only present in *Bryopsis* species isolated from tropical or
162 warm-temperate seas, *Labrenzia* species were more often found in algal samples isolated from
163 temperate regions, and *Rickettsia* endophytes were only present in *Bryopsis* species inhabiting
164 seas with a low mean sea surface temperature (11.7-12.8°C) and high chlorophyll, nitrate and

phosphate levels (Figs. 2 and 3). These correlations suggest that the distribution of individual bacterial OTUs may be more predictable than the total bacterial community composition. Individual bacterial endophyte groups also appear to be more strongly correlated with the host phylogeny than the overall bacterial composition. Flavobacteriaceae and Bacteroidetes species displayed a significant phylogenetic signal ($P \leq 0.01$, see Table S2) while Rhizobiaceae, Phyllobacteriaceae, *Mycoplasma, Rickettsia* and *Labrenzia* species did not. Because the host phylogeny is correlated with ecological features as a consequence of niche conservatism (Fig. 1), it is not obvious whether the latter pattern is due to ecological preferences of the endophytic bacteria or their host.

Host versus environmental influences

In order to disentangle the influences of different factors shaping the endophytic bacterial diversity, we performed variation partitioning analyses. In the first set of analyses we partitioned the variation of the bacterial diversity data with respect to the ecological, geographic and host-phylogenetic factors into different portions: a part strictly influenced by environmental variables, a part strictly influenced by the *Bryopsis* host phylogeny, a part strictly explained by geography, four parts explained by the shared influence of these three factors, and an unexplained part of the variation. When considering the total endophytic bacterial diversity, more or less equal parts of the variation (ca. 30%) were explained by environmental and phylogenetic factors, while the strict influence of geography was low; most of the variance, however, remained unexplained (Fig. 4A). Analyses of the seven bacterial phylotypes separately

showed that the influence of environment, phylogeny and geography was very different between the seven phylotypes. The influence of geography was, in most cases, low and highly correlated with environment and/or host phylogeny (Fig. 4A, Table S4). For this reason, we excluded geography in a second set of analyses (Fig. 4B). The independent effects of host phylogeny and environment had little influence on the presence of Phyllobacteriaceae, Rhizobiaceae and *Labrenzia* phylotypes. The shared influence of host phylogeny and environment was larger than their individual effects for these bacterial types. The occurrence of *Mycoplasma* and *Rickettsia* species, on the other hand, was in part strictly determined by environmental factors, whereas the distribution of Bacteroidetes could to a large extent be explained by host phylogenetic factors only. Most of the variance in presence of these six endophytic phylotypes, however, remained unexplained, suggesting that factors other than host phylogeny and environment (at least the seven variables sampled) determine their occurrence within particular *Bryopsis* samples (Fig. 4). This is in contrast with the situation for Flavobacteriaceae endophytes, whose presence could be entirely explained by host phylogenetic factors, which partly overlapped with environmental factors.

Discussion

Community structure and variation in traits across species are the outcome of environmental, geographical and historical factors which are clearly interwoven with each other. Bacterial communities associated with eukaryotic hosts are influenced by similar factors which need to be identified separately. Besides serving as baseline knowledge of the bacterial diversity

occurring inside the siphonous cells of *Bryopsis*, our results provide insights into the various elements that contribute to the composition of the endogenous bacterial flora of siphonous green seaweeds.

Diversity of endogenous bacteria

Besides the five bacterial phylotypes that were previously characterized in *Bryopsis* (*Labrenzia, Mycoplasma,* Phyllobacteriaceae, Bacteroidetes and Flavobacteriaceae) [17], we identified two additional phylotypes related to Rhizobiaceae and *Rickettsia* species. These bacteria have been especially well studied from terrestrial habitats [45, 46], but have also been reported from marine habitats. Rhizobiales are common epiphytes of *Ulva* seaweeds [47-50] and have also been isolated from the surface of kelps where they display antimicrobial activity [13]. Additionally, a *Rhodopseudomonas* species with the potential to fix nitrogen was isolated from the rhizoidal cytoplasm of the siphonous green seaweed *Caulerpa taxifolia* [19]. We presume that also *Bryopsis* hosts Rhizobiaceae species with nitrogen fixing capacities as we were able to amplify *Ensifer*-like nitrogenase reductase genes (EMBL accession numbers HE649370-HE649371) from *Bryopsis* samples 4718 and MZ4 [51]. Obligate intracellular *Rickettsia* species, on the other hand, have not previously been described from macroalgae but have been characterized through 16S rRNA gene analysis within freshwater green microalgae [52], marine ciliates [53] and coral tissue [54].

Factors affecting bacterial composition

Even though each bacterial phylotype was encountered in at least three *Bryopsis* samples, the total endophytic bacterial diversity per host sample showed no clear pattern. All algal samples harbored diverse combinations of one to four endophytic phylotypes regardless of their phylogenetic affiliation, geographic origin or macro-ecological niche. On the other hand, when the presence of individual endophytic phylotypes rather than the total bacterial composition was analyzed, host phylogenetic, geographic and environmental influences could be determined more clearly. These three factors, however, are inevitably interrelated as a result of phylogenetic niche conservatism, i.e. the tendency of closely related species to be ecologically similar [55], and historical factors such as dispersal limitation, resulting in geographic proximity of closely related species (Fig. 1). The *Bryopsis* host phylogeny was found to be mainly correlated with temperature-dependent variables, and to a lesser extent geography (Table S2). To disentangle the effects of host phylogeny, geography and environment, we performed a variation partitioning analysis [2, 56]. This technique has been proven useful in quantifying independent influences of host phylogeny and other traits like habitat and morphology in host-parasite [57-59] and arbuscular mycorrhizal symbiosis studies [60]. Our analyses shed light on the symbiotic nature and on potential modes of transmission of the individual endophytic phylotypes.

The presence of endophytic Phyllobacteriaceae, Rhizobiaceae and *Labrenzia* phylotypes was not separately determined by host phylogenetic, geographic and ecological factors, suggesting these endophytes are true generalists adapted to both free-living and host associated lifestyles

along with a wide variety of environmental conditions. This is consistent with our previous observations that *Labrenzia* and Phyllobacteriaceae endophytes can survive outside their *Bryopsis* host and are reacquired from the local environment after repeated wounding events in culture [25]. Also the close phylogenetic relatedness of all three endophytic phylotypes with sequences from free-living bacteria (Fig. S2) indicates a recently initiated, facultative association with the *Bryopsis* host. These generalist phylotypes may be selectively acquired by *Bryopsis* hosts to fulfill specific metabolic requirements, such as nitrogen-fixation (Rhizobiaceae, [45]), anoxygenic photosynthesis (Phyllobacteriaceae, [17]) or CO-oxidation (*Labrenzia*, [61]).

The occurrence of *Mycoplasma* and *Rickettsia* endophytes was to some extent strictly influenced by environmental factors. *Mycoplasma* endophytes were only present in *Bryopsis* samples from tropical regions, whereas *Rickettsia* bacteria were only found in algal samples isolated from temperate seas. This environmental influence suggests the acquisition of habitat-specific endophytes by *Bryopsis* hosts. In addition, the phylogenies of these more specialized endophytic phylotypes show a close relatedness with symbiotic *Rickettsia* and *Mycoplasma* species isolated from the cytoplasm of the marine ciliate *Diophrys appendiculata* [53] and the intestinal bacterial flora of the *Bryopsis*-feeding abalone *Haliotis diversicolor* [62], respectively, suggesting the uptake of these endophytes could be vector dependent. This hypothesis is likely as both endophytes belong to orders that are well-known as obligate intracellular parasites of plants and animals [63, 64]. Also within sponge hosts, horizontal symbiont transmission has been proposed to occur through vectors including sponge-feeding animals [65].

The presence of Bacteroidetes endophytes within *Bryopsis* was to a large degree influenced by host phylogenetic factors, indicating that they may be vertically transmitted. This may take place through sexual reproduction or asexual proliferation by fragmentation or extruded protoplasts that regenerate into new *Bryopsis* plants [66]. Additional evidence for vertical transmission is provided by culture studies, which showed that the Bacteroidetes endophytes were present within *Bryopsis* cells but not in the surrounding seawater, suggesting that they may be obligate endosymbionts [25].

The presence of Flavobacteriaceae was found to be influenced by host phylogenetic factors in part combined with environmental influences, suggesting that these bacteria are specialized and obligate endosymbionts, which are vertically transmitted via asexual and/or sexual reproductive stages [67]. This is in line with results from culture experiments, which showed that these bacterial species are strictly dependent on the *Bryopsis* host for their growth and survival [25].

Overall, the variation partitioning analyses showed a high fraction of unexplained variation. This was true when considering the total endophytic bacterial diversity and the individual bacterial phylotypes, with the notable exception of the Flavobacteriaceae endophytes. This unexplained variation indicates that other factors may be important in explaining endophytic bacterial composition in *Bryopsis*. The variables included in this study are situated at the macroecological level and are considered suitable for explaining broad scale distribution patterns [33]. Ecological variables associated with microhabitat preferences , biotic interactions (e.g. bacterial transmission as a result of grazer-induced wounding) and physiological state of

288 the host might also be important for interpreting fine scale bacterial community structure. In
289 addition, endophytic community composition may be a result of stochastic processes. The few
290 samples collected from one region and the same host species in this study indicate that a
291 considerable variation in bacterial community composition may exist. Further studies, including
292 many samples from a single host species at a single locality will be required to shed light on
293 variation between co-occurring *Bryopsis* plants. However, to do so, technical difficulties
294 inherent to this study (e.g. time consuming culturing and surface sterilization of host plants, and
295 a large (>95%) fraction of host chloroplast 16S rDNA in clone libraries) will need to be overcome,
296 for example by applying species specific primers or high-throughput sequencing techniques. The
297 unexplained variation in endophytic bacterial species composition is also relevant in the light of
298 recent studies providing evidence that functional genes and transcriptomes, rather than species
299 identified through rRNA taxonomy may be important in understanding bacterial community
300 structure [68, 69]. For example, bacterial communities associated with the green seaweed *Ulva*
301 were found to be largely determined by function, rather than taxonomic identity [68]. It might
302 thus be possible that functional genes rather than symbiont species or phylotypes are
303 influenced by evolutionary (host-phylogeny) and ecological factors [69].
304
305 In conclusion, characterization of *Bryopsis* algae sampled worldwide revealed the presence of
306 complex endobiotic bacterial communities. Evaluation of host phylogenetic, geographic and
307 ecological factors revealed the presence of a mix of generalist and specialist bacteria. These
308 observations, however, were only evident when subdividing the total endophytic diversity into

its individual bacterial phylotypes, suggesting that both the whole community and individual community members need to be considered in host-symbiont studies.

Acknowledgements

This research was funded by Research Foundation - Flanders project G.0045.08. We sincerely thank Gayle Hansen, Lennert Tyberghein, John West and Joe Zuccarello for providing *Bryopsis* cultures or collecting specimens. We acknowledge Frederik Hendrickx, Liam Revell and Guillaume Guénard for useful comments on the statistical design. FL and HV are postdoctoral fellows of the Research Foundation - Flanders.

References

1. Freckleton R.P., Jetz W. 2009 Space versus phylogeny: disentangling phylogenetic and spatial signals in comparative data. *Proceedings of the Royal Society B: Biological Sciences* **276**(1654), 21-30. (doi:10.1098/rspb.2008.0905).
2. Borcard D., Legendre P., Drapeau P. 1992 Partialling out the spatial component of ecological variation. *Ecology* **73**(3), 1045-1055. (doi:10.2307/1940179).
3. Harvey P.H., Pagel M.D. 1991 *The comparative method in evolutionary biology*. Oxford, UK, Oxford University Press.
4. Blomberg S.P., Garland T. 2002 Tempo and mode in evolution: phylogenetic inertia, adaptation and comparative methods. *J Evol Biol* **15**(6), 899-910. (doi:10.1046/j.1420-9101.2002.00472.x).
5. Ronquist F. 1997 Phylogenetic approaches in coevolution and biogeography. *Zool Scr* **26**(4), 313-322. (doi:10.1111/j.1463-6409.1997.tb00421.x).
6. Goecke F., Labes A., Wiese J., Imhoff J.F. 2010 Chemical interactions between marine macroalgae and bacteria. *Mar Ecol-Prog Ser* **409**, 267-299. (doi:10.3354/meps08607).
7. Hollants J., Leliaert F., De Clerck O., Willems A. 2012 What we can learn from sushi: a review on seaweed-bacterial associations. *FEMS Microbiol Ecol*. (doi:10.1111/j.1574-6941.2012.01446.x).
8. Friedrich M.W. 2012 Bacterial communities on macroalgae. In *Seaweed Biology* (eds. Wiencke C., Bischof K.), pp. 189-201, Springer Berlin Heidelberg.
9. Marshall K., Joint I., Callow M.E., Callow J.A. 2006 Effect of marine bacterial isolates on the growth and morphology of axenic plantlets of the green alga *Ulva linza*. *Microb Ecol* **52**(2), 302-310. (doi:10.1007/s00248-006-9060-x).
10. Patel P., Callow M.E., Joint I., Callow J.A. 2003 Specificity in the settlement – modifying response of bacterial biofilms towards zoospores of the marine alga *Enteromorpha*. *Environ Microbiol* **5**(5), 338-349. (doi:10.1046/j.1462-2920.2003.00407.x).
11. Weinberger F., Beltran J., Correa J.A., Lion U., Pohnert G., Kumar N., Steinberg P., Kloareg B., Potin P. 2007 Spore release in *Acrochaetium* sp. (Rhodophyta) is bacterially controlled. *J Phycol* **43**(2), 235-241. (doi:10.1111/j.1529-8817.2007.00329.x).

12. Dimitrieva G.Y., Crawford R.L., Yüksel G.Ü. 2006 The nature of plant growth-promoting effects of a pseudoalteromonad associated with the marine algae *Laminaria japonica* and linked to catalase excretion. *J Appl Microbiol* **100**(5), 1159-1169. (doi:10.1111/j.1365-2672.2006.02831.x).
13. Wiese J., Thiel V., Nagel K., Staufenberger T., Imhoff J.F. 2009 Diversity of antibiotic-active bacteria associated with the brown alga *Laminaria saccharina* from the Baltic Sea. *Mar Biotechnol* **11**(2), 287-300. (doi:10.1007/s10126-008-9143-4).
14. Rosenberg G., Paerl H.W. 1981 Nitrogen fixation by blue-green algae associated with the siphonous green seaweed *Codium decorticatum*, effects on ammonium uptake. *Mar Biol* **61**(2), 151-158. (doi:10.1007/bf00386654).
15. Colombo P.M. 1978 Occurrence of endophytic bacteria in siphonous algae. *Phycologia* **17**(2), 148-151. (doi:doi:10.2216/i0031-8884-17-2-148.1).
16. Delbridge L., Coulburn J., Fagerberg W., Tisa L.S. 2004 Community profiles of bacterial endosymbionts in four species of *Caulerpa*. *Symbiosis* **37**(1-3), 335-344.
17. Hollants J., Leroux O., Leliaert F., Decleyre H., De Clerck O., Willems A. 2011 Who is in there? Exploration of endophytic bacteria within the siphonous green seaweed *Bryopsis* (Bryopsidales, Chlorophyta). *PLoS ONE* **6**(10), e26458. (doi:10.1371/journal.pone.0026458).
18. Cocquyt E., Verbruggen H., Leliaert F., De Clerck O. 2010 Evolution and cytological diversification of the green seaweeds (Ulvophyceae). *Mol Biol Evol* **27**(9), 2052-2061. (doi:10.1093/molbev/msq091).
19. Chisholm J.R.M., Dauga C., Ageron E., Grimont P.A.D., Jaubert J.M. 1996 'Roots' in mixotrophic algae. *Nature* **381**(6581), 382-382.
20. Menzel D. 1988 How do giant plant-cells cope with injury - the wound response in siphonous green algae. *Protoplasma* **144**(2-3), 73-91.
21. Welling M., Pohnert G., Kupper F.C., Ross C. 2009 Rapid biopolymerisation during wound plug formation in green algae. *J Adhes* **85**(11), 825-838. (doi:10.1080/00218460903291452).
22. Williams S.L., Smith J.E. 2007 A global review of the distribution, taxonomy, and impacts of introduced seaweeds. *Annual Review of Ecology, Evolution, and Systematics* **38**(1), 327-359. (doi:doi:10.1146/annurev.ecolsys.38.091206.095543).
23. Meusnier I., Olsen J.L., Stam W.T., Destombe C., Valero M. 2001 Phylogenetic analyses of *Caulerpa taxifolia* (Chlorophyta) and of its associated bacterial microflora provide clues to the origin of the Mediterranean introduction. *Mol Ecol* **10**(4), 931-946.
24. Kim G.H., Klotchkova T.A., West J.A. 2002 From protoplasm to swarmer: regeneration of protoplasts from disintegrated cells of the multicellular marine green alga *Microdictyon umbilicatum* (Chlorophyta). *J Phycol* **38**(1), 174-183. (doi:10.1046/j.1529-8817.2002.01053.x).
25. Hollants J., Decleyre H., Leliaert F., De Clerck O., Willems A. 2011 Life without a cell membrane: challenging the specificity of bacterial endophytes within *Bryopsis* (Bryopsidales, Chlorophyta). *BMC Microbiol* **11**, e255.
26. Burr F.A., West J.A. 1970 Light and electron microscope observations on the vegetative and reproductive structures of *Bryopsis hypnoides*. *Phycologia* **9**(1), 17-37. (doi:doi:10.2216/i0031-8884-9-1-17.1).
27. Hollants J., Leliaert F., De Clerck O., Willems A. 2010 How endo- is endo-? Surface sterilization of delicate samples: a *Bryopsis* (Bryopsidales, Chlorophyta) case study. *Symbiosis* **51**(1), 131-138. (doi:10.1007/s13199-010-0068-0).
28. Doyle J.L., Doyle J.J. 1987 A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochem Bull* **19**(1), 11-15.

29. Guindon S., Gascuel O. 2003 A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Syst Biol* **52**(5), 696-704. (doi:10.1080/10635150390235520).
30. Ronquist F., Huelsenbeck J.P. 2003 MrBayes 3: bayesian phylogenetic inference under mixed models. *Bioinformatics* **19**(12), 1572-1574. (doi:10.1093/bioinformatics/btg180).
31. Kumar S., Skjaeveland A., Orr R., Enger P., Ruden T., Mevik B.-H., Burki F., Botnen A., Shalchian-Tabrizi K. 2009 AIR: a batch-oriented web program package for construction of supermatrices ready for phylogenomic analyses. *BMC Bioinformatics* **10**(1), 357.
32. Posada D. 2008 jModelTest: phylogenetic model averaging. *Mol Biol Evol* **25**(7), 1253-1256. (doi:10.1093/molbev/msn083).
33. Tyberghein L., Verbruggen H., Pauly K., Troupin C., Mineur F., De Clerck O. 2012 Bio-ORACLE: a global environmental dataset for marine species distribution modelling. *Glob Ecol Biogeogr* **21**, 272-281. (doi:10.1111/j.1466-8238.2011.00656.x).
34. Dray S., Legendre P., Peres-Neto P.R. 2006 Spatial modelling: a comprehensive framework for principal coordinate analysis of neighbour matrices (PCNM). *Ecol Model* **196**, 483-493.
35. Guénard G., Legendre P., Boisclair D., Bilodeau M. 2010 Multiscale codependence analysis: an integrated approach to analyze relationships across scales. *Ecology* **91**, 2952-2964.
36. Gower J.C. 1966 Some distance properties of latent root and vector methods used in multivariate analysis. *Biometrika* **53**, 325-338.
37. Diniz-Filho J.A.F., De Sant'Ana C.E.R., Bini L.M. 1998 An eigenvector method for estimating phylogenetic inertia. *Evolution* **52**(1247-1262).
38. Tamura K., Dudley J., Nei M., Kumar S. 2007 MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. *Mol Biol Evol* **24**(8), 1596-1599. (doi:10.1093/molbev/msm092).
39. Anderson M.J. 2003 PCO: a FORTRAN computer program for principal coordinate analysis. (New Zealand, Department of Statistics, University of Auckland.
40. Ter Braak C.J.F. 1986 Canonical Correspondence Analysis: a new eigenvector technique for multivariate direct gradient analysis. *Ecology* **67**(5), 1167-1179.
41. Blomberg S.P., Garland T., Ives A.R. 2003 Testing for phylogenetic signal in comparative data: behavioral traits are more labile. *Evolution* **57**(4), 717-745. (doi:10.1111/j.0014-3820.2003.tb00285.x).
42. Kembel S.W., Cowan P.D., Helmus M.R., Cornwell W.K., Morlon H., Ackerly D.D., Blomberg S.P., Webb C.O. 2010 Picante: R tools for integrating phylogenies and ecology. *Bioinformatics* **26**(11), 1463-1464. (doi:10.1093/bioinformatics/btq166).
43. Fritz S.A., Purvis A. 2010 Selectivity in mammalian extinction risk and threat types: a new measure of phylogenetic signal strength in binary traits. *Conserv Biol* **24**(4), 1042-1051. (doi:10.1111/j.1523-1739.2010.01455.x).
44. Peres-Neto P.R., Legendre P., Dray S., Borcard D. 2006 Variation partitioning of species data matrices: estimation and comparison of fractions. *Ecology* **87**, 2614-2625.
45. Kuykendall L.D. 2005 Order VI. Rhizobiales ord. nov. In *Bergey's Manual of Systematic Bacteriology (The Proteobacteria), part C (The Alpha-, Beta-, Delta-, and Epsilonproteobacteria)* (eds. Brenner D.J., Krieg N.R., Staley J.T., Garrity G.M.), p. 324. New York, Springer.
46. Dale C., Moran N.A. 2006 Molecular interactions between bacterial symbionts and their hosts. *Cell* **126**(3), 453-465. (doi:10.1016/j.cell.2006.07.014).
47. Tujula N.A., Crocetti G.R., Burke C., Thomas T., Holmstrom C., Kjelleberg S. 2010 Variability and abundance of the epiphytic bacterial community associated with a green marine Ulvacean alga. *ISME J* **4**(2), 301-311. (doi:10.1038/ismej.2009.107).

48. Lachnit T., Meske D., Wahl M., Harder T., Schmitz R. 2011 Epibacterial community patterns on marine macroalgae are host-specific but temporally variable. *Environ Microbiol* **13**(3), 655-665. (doi:10.1111/j.1462-2920.2010.02371.x).
49. Burke C., Thomas T., Lewis M., Steinberg P., Kjelleberg S. 2011 Composition, uniqueness and variability of the epiphytic bacterial community of the green alga *Ulva australis*. *ISME J* **5**(4), 590-600. (doi:10.1038/ismej.2010.164).
50. Longford S.R., Tujula N.A., Crocetti G.R., Holmes A.J., Holmström C., Kjelleberg S., Steinberg P.D., Taylor M.W. 2007 Comparisons of diversity of bacterial communities associated with three sessile marine eukaryotes. *Aquat Microb Ecol* **48**(3), 217-229. (doi:10.3354/ame048217).
51. Hollants J. 2012 Endophytic bacteria within the green siphonous seaweed *Bryopsis*: Exploration of a partnership, Dissertation, Ghent University.
52. Kawafune K., Hongoh Y., Hamaji T., Nozaki H. 2012 Molecular identification of rickettsial endosymbionts in the non-phagotrophic volvocalean green algae. *PLoS ONE* **7**(2), e31749. (doi:10.1371/journal.pone.0031749).
53. Vannini C., Petroni G., Verni F., Rosati G. 2005 A bacterium belonging to the Rickettsiaceae family inhabits the cytoplasm of the marine ciliate *Diophrys appendiculata* (Ciliophora, Hypotrichia). *Microb Ecol* **49**(3), 434-442. (doi:10.1007/s00248-004-0055-1).
54. Sunagawa S., DeSantis T.Z., Piceno Y.M., Brodie E.L., DeSalvo M.K., Voolstra C.R., Weil E., Andersen G.L., Medina M. 2009 Bacterial diversity and White Plague Disease-associated community changes in the Caribbean coral *Montastraea faveolata*. *ISME J* **3**(5), 512-521.
55. Losos J.B. 2008 Phylogenetic niche conservatism, phylogenetic signal and the relationship between phylogenetic relatedness and ecological similarity among species. *Ecol Lett* **11**(10), 995-1003. (doi:10.1111/j.1461-0248.2008.01229.x).
56. Desdevises Y., Legendre P., Azouzi L., Morand S. 2003 Quantifying phylogenetically structured environmental variation. *Evolution* **57**(11), 2647-2652. (doi:10.1111/j.0014-3820.2003.tb01508.x).
57. Timi J.T., Lanfranchi A.L. 2009 The importance of the compound community on the parasite infracommunity structure in a small benthic fish. *Parasitol Res* **104**(2), 295-302. (doi:10.1007/s00436-008-1191-1).
58. Krasnov B.R., Poulin R., Mouillot D. 2011 Scale-dependence of phylogenetic signal in ecological traits of ectoparasites. *Ecography* **34**(1), 114-122. (doi:10.1111/j.1600-0587.2010.06502.x).
59. Desdevises Y., Morand S., Legendre P. 2002 Evolution and determinants of host specificity in the genus *Lamellodiscus* (Monogenea). *Biol J Linn Soc* **77**(4), 431-443. (doi:10.1046/j.1095-8312.2002.00114.x).
60. Stajerova K., Smilauerova M., Smilauer P. 2009 Arbuscular mycorrhizal symbiosis of herbaceous invasive neophytes in the Czech Republic. *Preslia* **81**(4), 341-355.
61. Weber C.F., King G.M. 2007 Physiological, ecological, and phylogenetic characterization of *Stappia*, a marine CO-oxidizing bacterial genus. *Appl Environ Microbiol* **73**(4), 1266-1276. (doi:10.1128/aem.01724-06).
62. Huang Z.-B., Guo F., Zhao J., Li W.-D., Ke C.-H. 2010 Molecular analysis of the intestinal bacterial flora in cage-cultured adult small abalone, *Haliotis diversicolor*. *Aquac Res* **41**(11), e760-e769. (doi:10.1111/j.1365-2109.2010.02577.x).
63. Weinert L., Werren J., Aebi A., Stone G., Jiggins F. 2009 Evolution and diversity of *Rickettsia* bacteria. *BMC Biol* **7**(1), 6.
64. Fraune S., Zimmer M. 2008 Host-specificity of environmentally transmitted *Mycoplasma*-like isopod symbionts. *Environ Microbiol* **10**(10), 2497-2504. (doi:10.1111/j.1462-2920.2008.01672.x).

65. Taylor M.W., Radax R., Steger D., Wagner M. 2007 Sponge-associated microorganisms: evolution, ecology, and biotechnological potential. *Microbiol Mol Biol Rev* **71**(2), 295-347. (doi:10.1128/mmbr.00040-06).
66. Morabito M., Gargiulo G., Genovese G. 2010 A review of life history pathways in *Bryopsis*. *AAPP | Physical, Mathematical, and Natural Sciences* **88**(2), doi 10.1478/C1471A1002005.
67. Vignon M., Pariselle A., Vanhove M.P.M. 2011 Modularity in attachment organs of African *Cichlidogyrus* (Platyhelminthes: Monogenea: Ancyrocephalidae) reflects phylogeny rather than host specificity or geographic distribution. *Biol J Linn Soc* **102**(3), 694-706. (doi:10.1111/j.1095-8312.2010.01607.x).
68. Burke C., Steinberg P., Rusch D., Kjelleberg S., Thomas T. 2011 Bacterial community assembly based on functional genes rather than species. *Proceedings of the National Academy of Sciences* **108**(34), 14288-14293. (doi:10.1073/pnas.1101591108).
69. Fan L., Reynolds D., Liu M., Stark M., Kjelleberg S., Webster N.S., Thomas T. 2012 Functional equivalence and evolutionary convergence in complex communities of microbial sponge symbionts. *Proceedings of the National Academy of Sciences* **109**(27), E1878–E1887. (doi:10.1073/pnas.1203287109).

Figure legends

Figure 1. Relationships between host phylogeny, environment and geography on endophytic bacterial composition in *Bryopsis* seaweeds and relations between these three factors. 1: phylogenetic structured variation, 2: ecological structured variation and 3: geographic structured variation. The shared influence of phylogeny and environment (1+2) is known as "phylogenetically structured environmental variation".

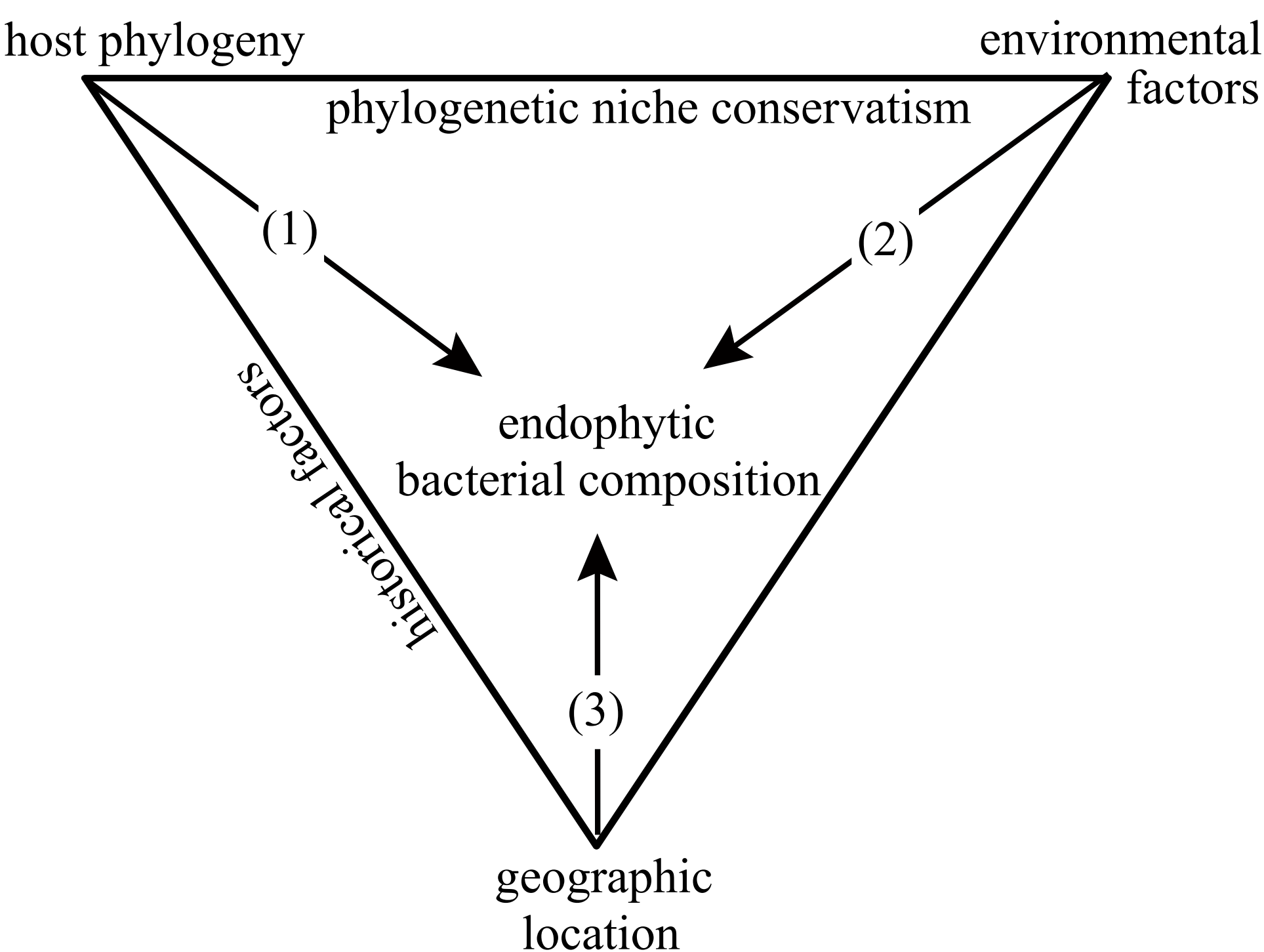
Figure 2. Endophytic diversity results, geographic data and environmental variables plotted against the *Bryopsis* host phylogram. The endophytic bacterial diversity displayed by blue boxes summarizes the diversity results from the 16S rRNA gene clone libraries and DGGE analyses. Environmental variables were extracted from the host sampling sites using Bio-ORACLE: salinity (PSS); chlo_mean: annual mean chlorophyll ($mg.m^{-3}$); nitrate ($\mu mol.l^{-1}$); phosphate ($\mu mol.l^{-1}$); dissolved oxygen ($ml.l^{-1}$); PAR_mean: annual mean photosynthetically

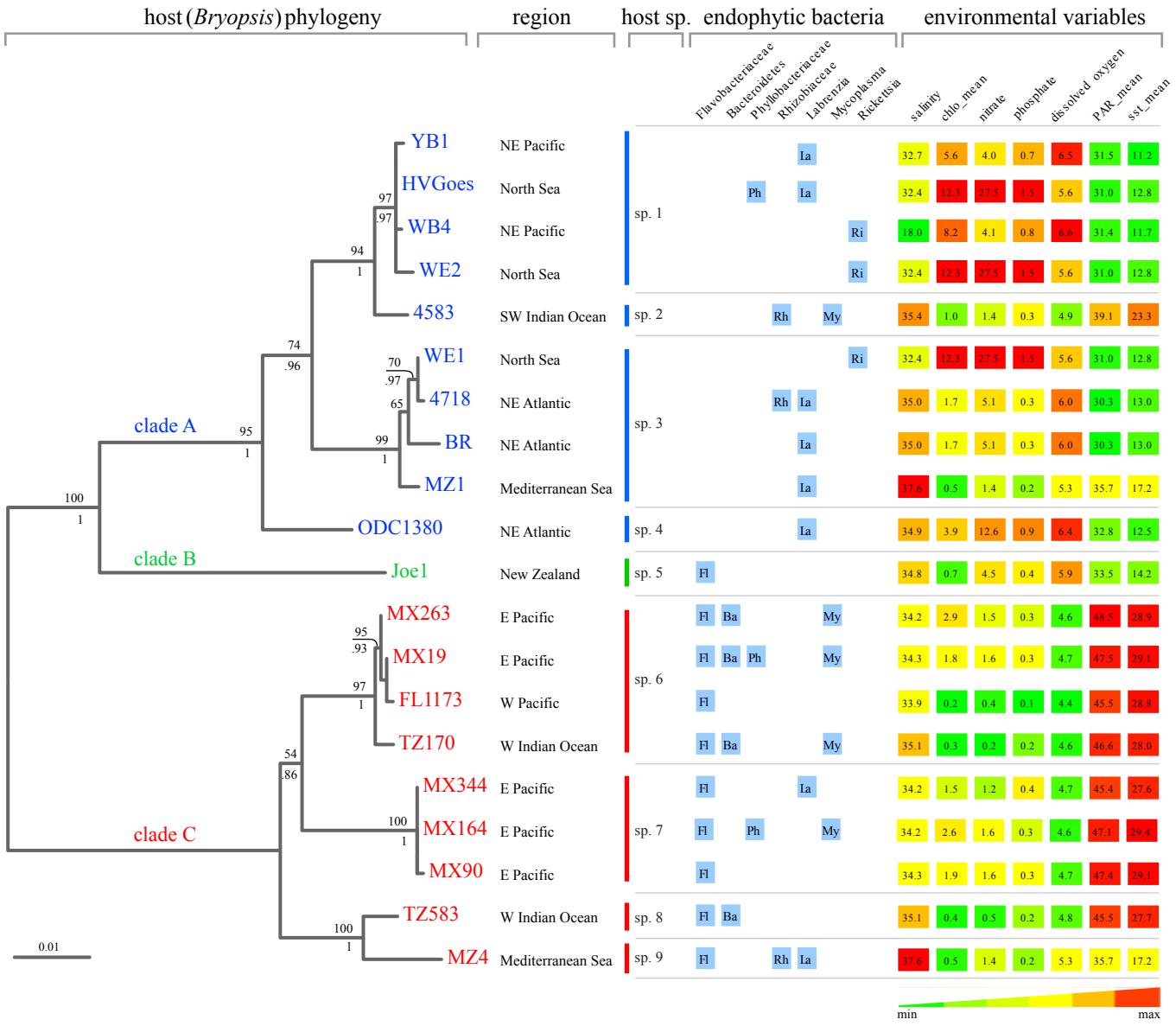
available radiation (µmol.m^{-2}.s^{-1}); sst_mean: annual mean sea surface temperature (°C). The phylogram on the left classifies the 20 algal samples for which endophytic bacterial data are available in nine different *Bryopsis* species and three distinct clades (i.e. A, B and C). These clades seem more consistent with the ecology of the host samples (environmental variables depicted on the right) than with their geographic origin (sample region). ML bootstrap values and BI posterior probabilities, respectively, are indicated above and below the branch nodes. The scale bar indicates 0.01 nucleotide changes per nucleotide position.

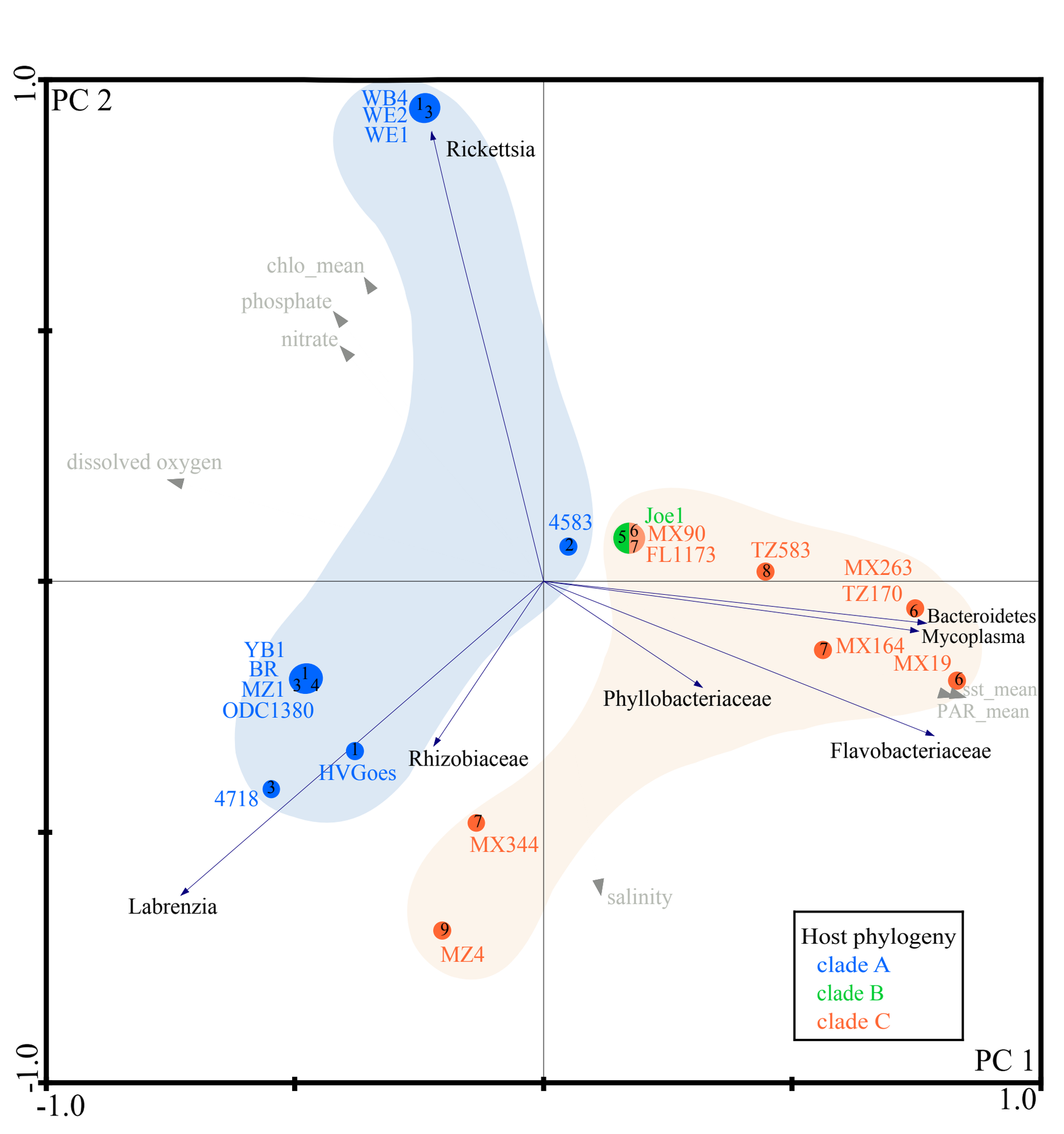
Figure 3. Principal component analysis of the 20 *Bryopsis* samples for which endophytic bacterial information is available. The PCA plot spreads the host samples in direction of maximum variance in endophytic bacterial community composition with principal component 1 explaining 41.7% and principal component 2 19.9% of the variance. *Bryopsis* species are indicated as numbers 1-9 and phylogenetic clades A, B and C are showed in blue, green and red, respectively. Environmental variables (in gray) were plotted on the PCA graph as supplementary information.

Figure 4. Variation partitioning. Adjusted R^2 values are given or illustrated. **A.** Results of the analysis with three explanatory tables: phylogeny, environment and geography. Venn diagram shows the influence of the three factors on the total bacterial diversity. Below are the variation explained by geography and the unexplained variation given for the seven bacterial phylotypes. Because the influence of geography was, in most cases, low and highly correlated with

environment and/or host phylogeny, we excluded geography in a second set of analyses shown in Fig. 4B. **B.** Results of the analysis with two explanatory tables: phylogeny and environment. Diagrams show the unique and shared influence of both factors on the variation in total endophytic bacterial diversity and the individual endophytic phylotypes. Negative fractions (which indicate that two explanatory variables have strong and opposite effects on the dependent variable) are treated as zeros in the graphs. We refer to Table S4 for a detailed overview of the variation partitioning results.


host phylogeny
environmental factors
phylogenetic niche conservatism
(1)
(2)
endophytic bacterial composition
historical factors
(3)
geographic location


host (Bryopsis) phylogeny
region
host sp.
endophytic bacteria
environmental variables
Flavobacteriaceae
Bacteroidetes
Phyllobacteriaceae
Rhizobiaceae
Labrenzia
Mycoplasma
Rickettsia
salinity
chlo_mean
nitrate
phosphate
dissolved_oxygen
PAR_mean
sst_mean
YB1
HVGoes
WB4
WE2
4583
WE1
4718
BR
MZ1
ODC1380
Joe1
MX263
MX19
FL1173
TZ170
MX344
MX164
MX90
TZ583
MZ4
clade A
clade B
clade C
NE Pacific
North Sea
NE Pacific
North Sea
SW Indian Ocean
North Sea
NE Atlantic
NE Atlantic
Mediterranean Sea
NE Atlantic
New Zealand
E Pacific
E Pacific
W Pacific
W Indian Ocean
E Pacific
E Pacific
E Pacific
W Indian Ocean
Mediterranean Sea
sp. 1
sp. 2
sp. 3
sp. 4
sp. 5
sp. 6
sp. 7
sp. 8
sp. 9
0.01
min
max


PC 2
PC 1
1.0
-1.0
-1.0
1.0
WB4
WE2
WE1
Rickettsia
chlo_mean
phosphate
nitrate
dissolved oxygen
4583
Joe1
MX90
FL1173
TZ583
MX263
TZ170
Bacteroidetes
Mycoplasma
MX164
MX19
sst_mean
PAR_mean
Phyllobacteriaceae
Flavobacteriaceae
YB1
BR
MZ1
ODC1380
HVGoes
Rhizobiaceae
4718
MX344
salinity
Labrenzia
MZ4
Host phylogeny
clade A
clade B
clade C

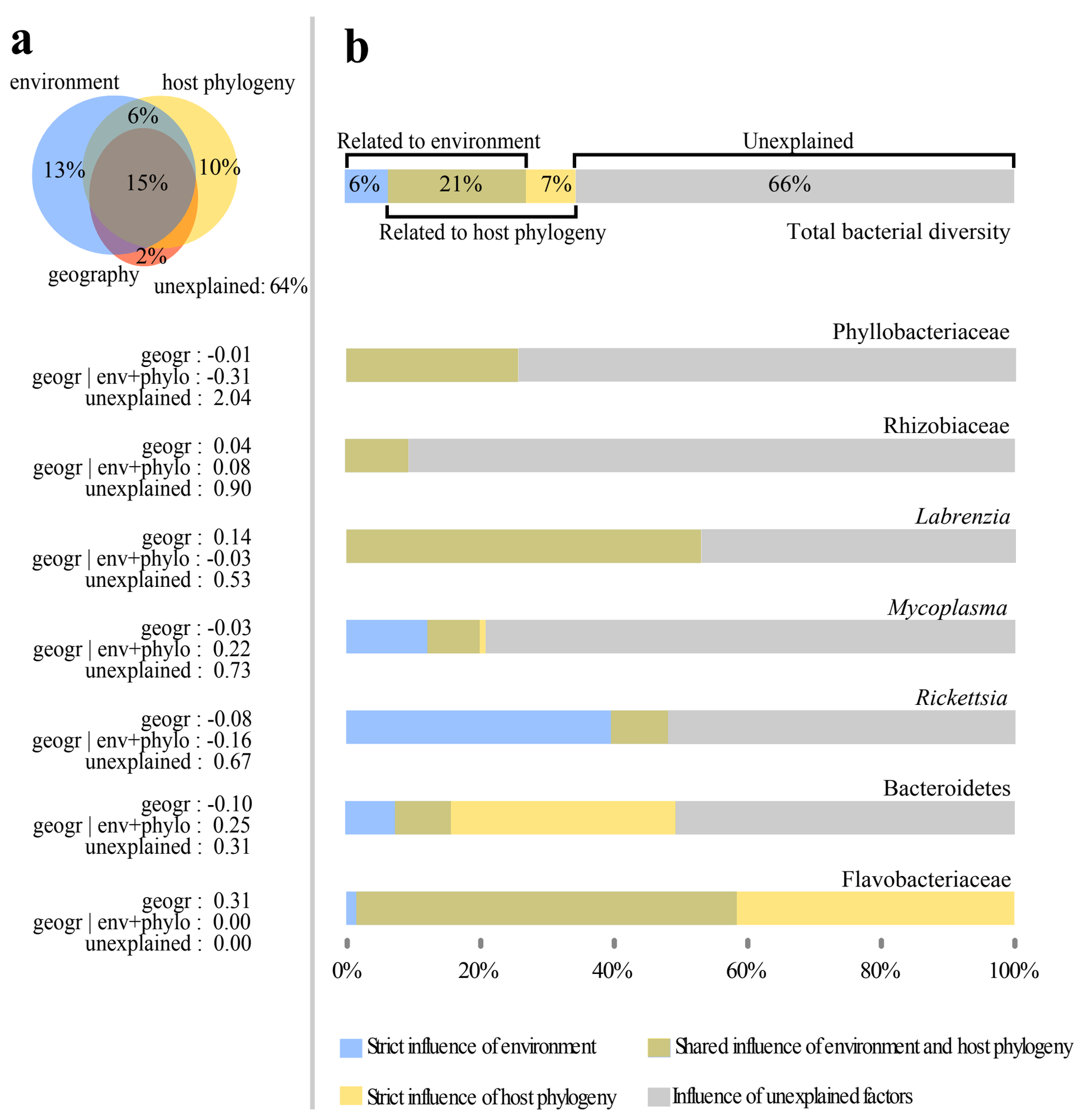
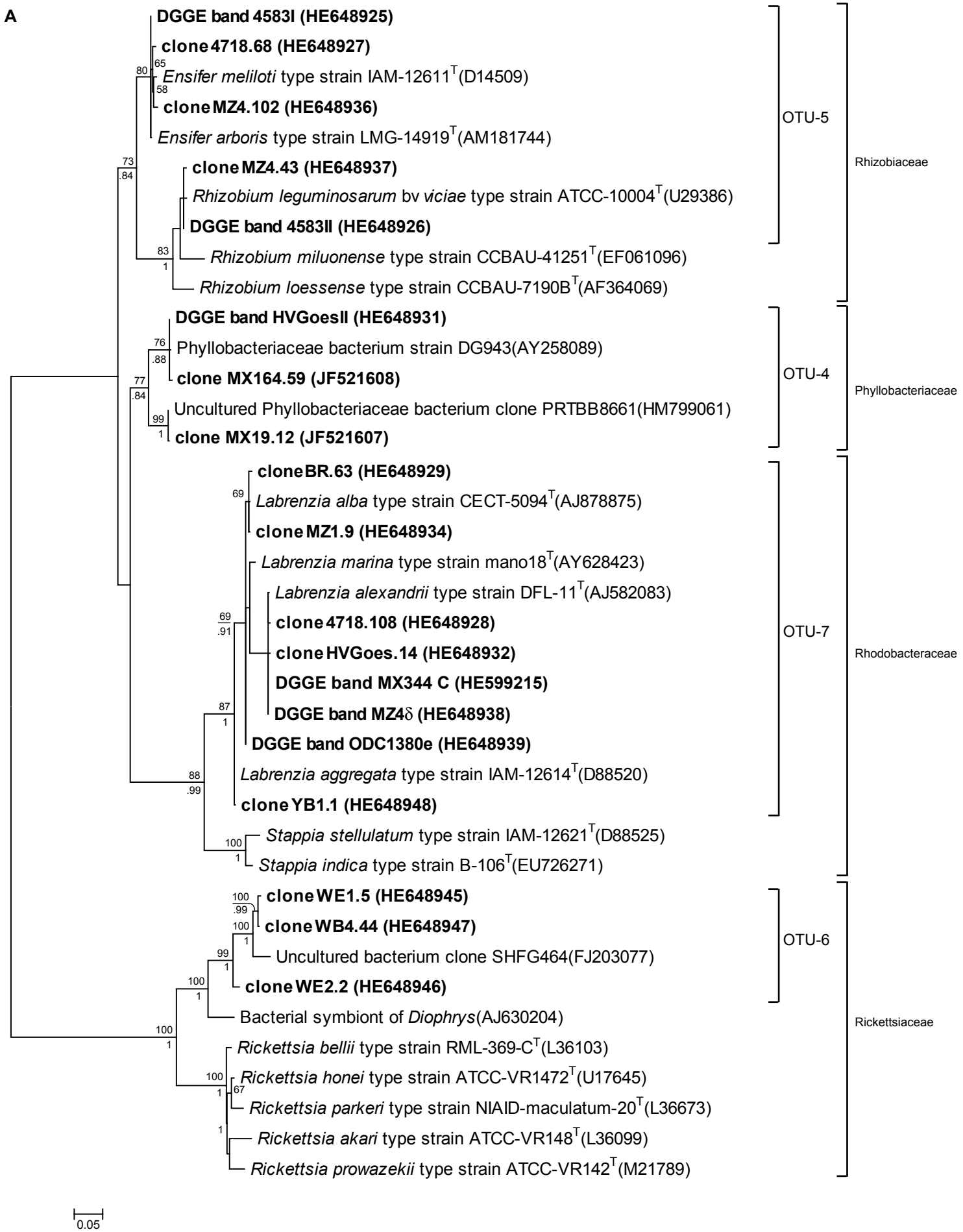

a
environment
host phylogeny
6%
13%
15%
10%
2%
geography
unexplained: 64%
geogr : -0.01
geogr | env+phylo : -0.31
unexplained : 2.04
geogr : 0.04
geogr | env+phylo : 0.08
unexplained : 0.90
geogr : 0.14
geogr | env+phylo : -0.03
unexplained : 0.53
geogr : -0.03
geogr | env+phylo : 0.22
unexplained : 0.73
geogr : -0.08
geogr | env+phylo : -0.16
unexplained : 0.67
geogr : -0.10
geogr | env+phylo : 0.25
unexplained : 0.31
geogr : 0.31
geogr | env+phylo : 0.00
unexplained : 0.00
b
Related to environment
Unexplained
6%
21%
7%
66%
Related to host phylogeny
Total bacterial diversity
Phyllobacteriaceae
Rhizobiaceae
Labrenzia
Mycoplasma
Rickettsia
Bacteroidetes
Flavobacteriaceae
0%
20%
40%
60%
80%
100%
Strict influence of environment
Shared influence of environment and host phylogeny
Strict influence of host phylogeny
Influence of unexplained factors

Figure S1. Map of *Bryopsis* sampling sites. The collection sites are marked by black circles and labelled with the *Bryopsis* sample name. In addition to the 15 *Bryopsis* samples analyzed in this study, also the five Mexican *Bryopsis* samples MX19, MX90, MX164, MX263 and MX344, which were previously studied [17], are depicted.


A
DGGE band 4583I (HE648925)
clone 4718.68 (HE648927)
Ensifer meliloti type strain IAM-12611T(D14509)
clone MZ4.102 (HE648936)
Ensifer arboris type strain LMG-14919T(AM181744)
clone MZ4.43 (HE648937)
Rhizobium leguminosarum bv viciae type strain ATCC-10004T(U29386)
DGGE band 4583II (HE648926)
Rhizobium miluonense type strain CCBAU-41251T(EF061096)
Rhizobium loessense type strain CCBAU-7190BT(AF364069)
DGGE band HVGoesII (HE648931)
Phyllobacteriaceae bacterium strain DG943(AY258089)
clone MX164.59 (JF521608)
Uncultured Phyllobacteriaceae bacterium clone PRTBB8661(HM799061)
clone MX19.12 (JF521607)
clone BR.63 (HE648929)
Labrenzia alba type strain CECT-5094T(AJ878875)
clone MZ1.9 (HE648934)
Labrenzia marina type strain mano18T(AY628423)
Labrenzia alexandrii type strain DFL-11T(AJ582083)
clone 4718.108 (HE648928)
clone HVGoes.14 (HE648932)
DGGE band MX344 C (HE599215)
DGGE band MZ4δ (HE648938)
DGGE band ODC1380e (HE648939)
Labrenzia aggregata type strain IAM-12614T(D88520)
clone YB1.1 (HE648948)
Stappia stellulatum type strain IAM-12621T(D88525)
Stappia indica type strain B-106T(EU726271)
clone WE1.5 (HE648945)
clone WB4.44 (HE648947)
Uncultured bacterium clone SHFG464(FJ203077)
clone WE2.2 (HE648946)
Bacterial symbiont of Diophrys(AJ630204)
Rickettsia bellii type strain RML-369-CT(L36103)
Rickettsia honei type strain ATCC-VR1472T(U17645)
Rickettsia parkeri type strain NIAID-maculatum-20T(L36673)
Rickettsia akari type strain ATCC-VR148T(L36099)
Rickettsia prowazekii type strain ATCC-VR142T(M21789)
OTU-5
OTU-4
OTU-7
OTU-6
Rhizobiaceae
Phyllobacteriaceae
Rhodobacteraceae
Rickettsiaceae
0.05

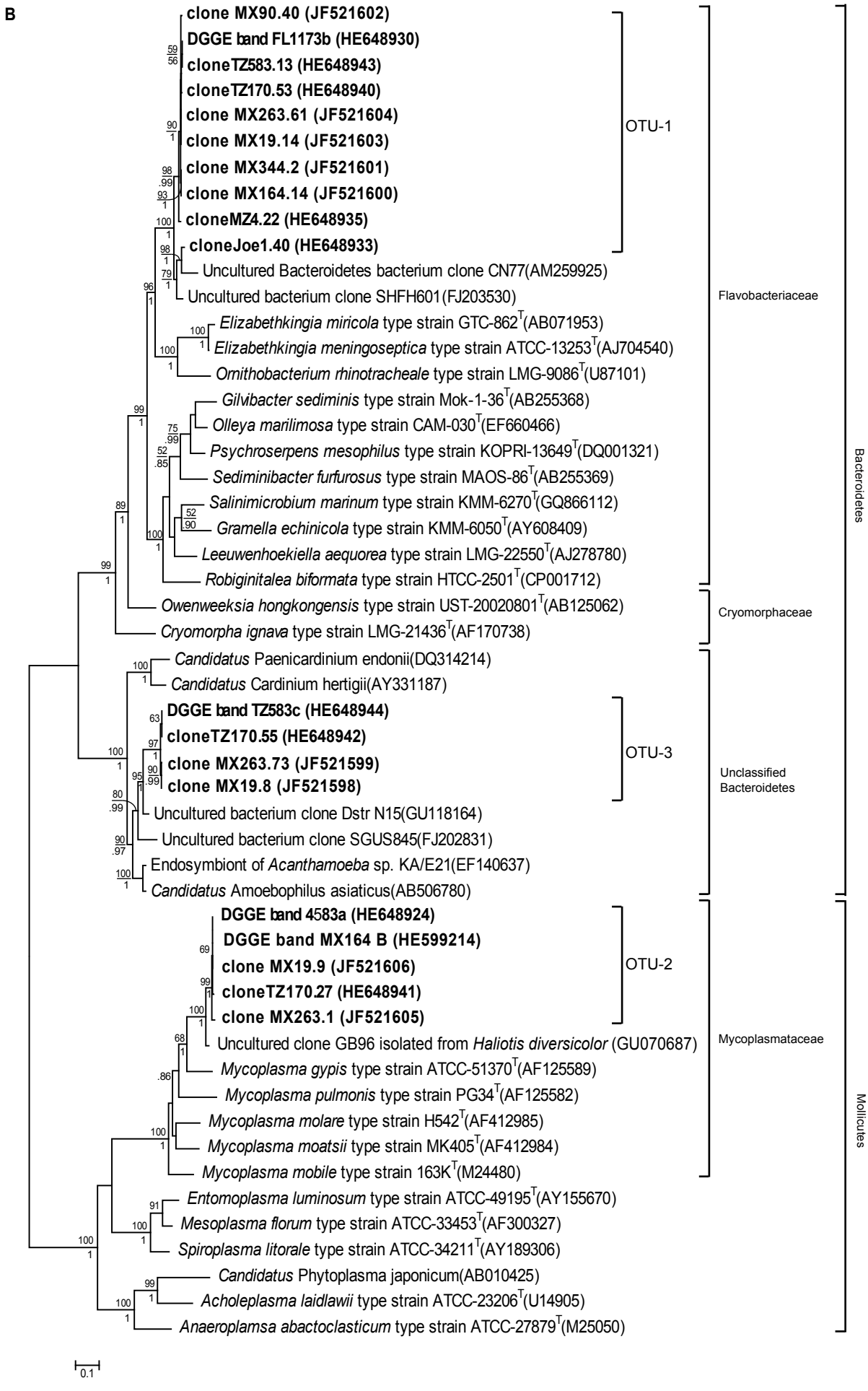


Figure S2. Wide-range ML/BI trees showing the phylogenetic positions of endophytic bacterial clones and DGGE bands. Phylogenies were inferred from 16S rRNA gene sequences determined in this and our previous study (in bold), BLAST hits (see Table S3), and Alphaproteobacterial (A) as well as Bacteroidetes and Mollicutes (B) type strains. Phylograms were generated using ML and BI under a GTR+G model. ML bootstrap values above 50 and BI posterior probabilities above 0.8, respectively, are indicated on top and beneath the branch nodes. The scale bar shows 5 (A) and 10 (B) nucleotide substitutions per 100 nucleotides.

Table S1. **Overview of the *Bryopsis* samples analyzed in this study, their collection sites and collection dates.**

***Bryopsis* sample**	**Collection site**	**Collection date**
Bryopsis 4583	Umhlanga Rocks KwaZulu Natal, South Africa	August 2005
Bryopsis 4718	Roscoff, Brittany, France	April 2008
Bryopsis BR	Roscoff, Brittany, France	July 2008
Bryopsis FL1173	Negros Oriental, Apo Island, Philippines	September 2007
Bryopsis HVGoes	Sas van Goes, The Netherlands	June 2007
Bryopsis Joe1	Moa Dt, Wellington, New Zealand	October 2008
Bryopsis MX19	Playa el Panteon, Puerto Angel, Oaxaca, Mexico	February 2009
Bryopsis MX90	Mazunte Beach, Mazunte, Oaxaca, Mexico	February 2009
Bryopsis MX164	Acapulco, Guerrero, Mexico	February 2009
Bryopsis MX263	Playa las Gatas, Zihuatanejo, Guerrero, Mexico	February 2009
Bryopsis MX344	Playa Careyero, Punta de Mita, Nayarit, Mexico	February 2009
Bryopsis MZ1 and MZ4	Begur, Catalogna, Spain	January 2008
Bryopsis ODC1380	Pointe de la Crèche, Boulogne, Nord-Pas-de-Calais, France	April 2007
Bryopsis TZ170	N tip of peninsula, Ruvula, Mtwara, Tanzania	January 2008
Bryopsis TZ583	E of lighthouse, Nungwi, Zanzibar, Tanzania	February 2008
Bryopsis WB4	Willapa Bay, SW Washington, USA	May 2008
Bryopsis WE1 and WE2	Wemeldinge, The Netherlands	May 2008
Bryopsis YB1	Yaquina Bay, Oregon, USA	May 2008

Table S2. Phylogenetic signal values calculated for the environmental variables (Fig. 2), geography (Moran's eigenvector maps, MEM 1 and 2), total bacterial composition (principal components 1 and 2) (Fig. 3) and the presence of the seven endophytic bacterial OTUs (Fig. 2). P values were calculated from randomizations using Blomberg et al.'s K [(K)] and Fritz and Purvis' D statistic [(D)]. Statistical significant p-values ≤ 0.01 are indicated in bold.

		Phylogenetic signal	P-value
Environmental variables	chlo_mean	0.07	0.18[(K)]
	dissolved oxygen	0.16	**0.00**[(K)]
	nitrate	0.04	0.49[(K)]
	PAR_mean	0.73	**0.00**[(K)]
	phosphate	0.05	0.29[(K)]
	salinity	0.04	0.55[(K)]
	sst_mean	0.70	**0.00**[(K)]
Geography	MEM 1	0.07	0.07[(K)]
	MEM 2	0.08	0.20[(K)]
Total bacterial composition	PC 1	0.07	0.09[(K)]
	PC 2	0.03	0.74[(K)]
Presence endophytic phylotypes	Bacteroidetes	-0.03	**0.01**[(D)]
	Flavobacteriaceae	-0.54	**0.00**[(D)]
	Labrenzia	1.16	0.67[(D)]
	Mycoplasma	0.63	0.12[(D)]
	Phyllobacteriaceae	1.75	0.98[(D)]
	Rhizobiaceae	1.19	0.61[(D)]
	Rickettsia	1.24	0.66[(D)]

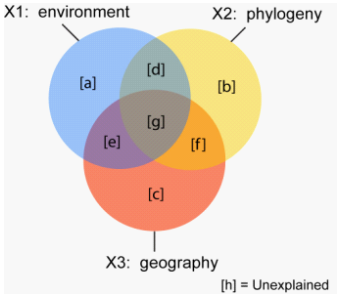
Table S3. Taxonomic affiliation of the clones and DGGE bands representing the endophytic bacterial OTUs, sorted per *Bryopsis* sample.

Host	16S rRNA gene sequence analysis of bacterial clones and DGGE bands					
***Bryopsis* sample**	**OTU* no.**	**OTU representative clone/DGGE band**	**Accession no.**	**Higher taxonomic ranks**	**Closest NCBI match**	**Accession no. (Query coverage/Maximum identity)**
4583	OTU-2	DGGE band 4583a	HE648924	Mollicutes, Mycoplasmatales, Mycoplasmataceae	Uncultured *Mycoplasma* sp. clone MX19.9	JF521606 (100/100)
	OTU-5	DGGE band 4583I	HE648925	Alphaproteobacteria; Rhizobiales; Rhizobiaceae	*Ensifer meliloti* strain RMP66	AB665549 (100/100)
	OTU-5	DGGE band 4583II	HE648926	Alphaproteobacteria; Rhizobiales; Rhizobiaceae	*Rhizobium leguminosarum* strain IPR-Pv1097	JN208903 (100/100)
4718	OTU-5	Clone 4718.68	HE648927	Alphaproteobacteria; Rhizobiales; Rhizobiaceae	*Ensifer medicae* WSM419	CP000738 (100/99)
	OTU-7	Clone 4718.108	HE648928	Alphaproteobacteria; Rhodobacterales; Rhodobacteraceae	Uncultured bacterium clone SGUS723	FJ202588 (100/99)
BR	OTU-7	Clone BR.63	HE648929	Alphaproteobacteria; Rhodobacterales; Rhodobacteraceae	*Labrenzia alba* strain CECT 5094	NR_042378 (100/99)
FL1173	OTU-1	DGGE band FL1173b	HE648930	Bacteroidetes; Flavobacteria; Flavobacteriales	Uncultured Flavobacteriaceae bacterium clone MX19.14	JF521603 (100/100)
HVGoes	OTU-4	DGGE band HVGoesII	HE648931	Alphaproteobacteria; Rhizobiales; Phyllobacteriaceae	Uncultured Phyllobacteriaceae bacterium clone MX164.59	JF521608 (100/100)
	OTU-7	Clone HVGoes.14	HE648932	Alphaproteobacteria; Rhodobacterales; Rhodobacteraceae	Uncultured bacterium clone SGUS723	FJ202588 (100/99)
Joe1	OTU-1	Clone Joe1.40	HE648933	Bacteroidetes; Flavobacteria; Flavobacteriales	Uncultured Flavobacteriaceae bacterium clone MX19.14	JF521603 (100/96)
MX19	OTU-1	Clone MX19.14	JF521603	Bacteroidetes; Flavobacteria; Flavobacteriales	Uncultured bacterium clone SHFH601	FJ203530 (99/96)
	OTU-2	Clone MX19.9	JF521606	Mollicutes, Mycoplasmatales, Mycoplasmataceae	Uncultured bacterium clone GB96	GU070687 (100/97)
	OTU-3	Clone MX19.8	JF521598	Bacteroidetes; unclassified Bacteroidetes	Uncultured bacterium clone Dstr_N15	GU118164 (99/94)
	OTU-4	Clone MX19.12	JF521607	Alphaproteobacteria; Rhizobiales; Phyllobacteriaceae	Uncultured Rhizobiales bacterium clone PRTBB8661	HM799061 (99/99)
MX90	OTU-1	Clone MX90.40	JF521602	Bacteroidetes; Flavobacteria; Flavobacteriales	Uncultured bacterium clone SHFH601	FJ203530 (99/96)
MX164	OTU-1	Clone MX164.14	JF521600	Bacteroidetes; Flavobacteria; Flavobacteriales	Uncultured bacterium clone SHFH601	FJ203530 (99/96)
	OTU-2	DGGE band MX164 B	HE599214	Mollicutes, Mycoplasmatales, Mycoplasmataceae	Uncultured bacterium clone GB96	GU070687 (100/97)
	OTU-4	Clone MX164.59	JF521608	Alphaproteobacteria; Rhizobiales; Phyllobacteriaceae	Phylobacteriaceae bacterium strain DG943	AY258089 (97/99)
MX263	OTU-1	Clone MX263.61	JF521604	Bacteroidetes; Flavobacteria; Flavobacteriales	Uncultured bacterium clone SHFH601	FJ203530 (99/96)
	OTU-2	Clone MX263.1	JF521605	Mollicutes, Mycoplasmatales, Mycoplasmataceae	Uncultured bacterium clone GB96	GU070687 (100/97)
	OTU-3	Clone MX263.73	JF521599	Bacteroidetes; unclassified Bacteroidetes	Uncultured bacterium clone Dstr_N15	GU118164 (99/94)

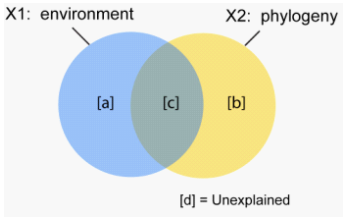
Host	16S rRNA gene sequence analysis of bacterial clones and DGGE bands					
***Bryopsis* sample**	**OTU*no.**	**OTU representative clone/DGGE band**	**Accession no.**	**Higher taxonomic ranks**	**Closest NCBI match**	**Accession no. (Query coverage/Maximum identity)**
MX344	OTU-1	Clone MX344.2	JF521601	Bacteroidetes; Flavobacteria; Flavobacteriales	Uncultured bacterium clone SHFH601	FJ203530 (99/96)
	OTU-7	DGGE band MX344 C	HE599215	Alphaproteobacteria; Rhodobacterales; Rhodobacteraceae	*Labrenzia alba* isolate CMS163	FR750958 (100/100)
MZ1	OTU-7	Clone MZ1.9	HE648934	Alphaproteobacteria; Rhodobacterales; Rhodobacteraceae	*Labrenzia alba* type strain CECT 5094^{T}	AJ878875 (100/99)
MZ4	OTU-1	Clone MZ4.22	HE648935	Bacteroidetes; Flavobacteria; Flavobacteriales	Uncultured Flavobacteriaceae bacterium clone MX19.14	JF521603 (100/99)
	OTU-5	Clone MZ4.102	HE648936	Alphaproteobacteria; Rhizobiales; Rhizobiaceae	*Ensifer meliloti* SM11	CP001830 (100/99)
	OTU-5	Clone MZ4.43	HE648937	Alphaproteobacteria; Rhizobiales; Rhizobiaceae	*Rhizobium leguminosarum* bv. *viciae* strain BIHB 1160	EU730590 (100/99)
	OTU-7	DGGE band MZ4?	HE648938	Alphaproteobacteria; Rhodobacterales; Rhodobacteraceae	Uncultured *Labrenzia* sp. DGGE band MX344 C	HE599215 (100/100)
ODC1380	OTU-7	DGGE band ODC1380e	HE648939	Alphaproteobacteria; Rhodobacterales; Rhodobacteraceae	*Labrenzia aggregata* strain KMO25	JF514325 (100/100)
TZ170	OTU-1	Clone TZ170.53	HE648940	Bacteroidetes; Flavobacteria; Flavobacteriales	Uncultured Flavobacteriaceae bacterium clone MX19.14	JF521603 (100/99)
	OTU-2	Clone TZ170.27	HE648941	Mollicutes, Mycoplasmatales, Mycoplasmataceae	Uncultured *Mycoplasma* sp. clone MX19.9	JF521606 (100/99)
	OTU-3	Clone TZ170.55	HE648942	Bacteroidetes; unclassified Bacteroidetes	Uncultured Bacteroidetes bacterium clone MX19.8	JF521598 (100/99)
TZ583	OTU-1	Clone TZ583.13	HE648943	Bacteroidetes; Flavobacteria; Flavobacteriales	Uncultured Flavobacteriaceae bacterium clone MX19.14	JF521603 (100/99)
	OTU-3	DGGE band TZ583c	HE648944	Bacteroidetes; unclassified Bacteroidetes	Uncultured Bacteroidetes bacterium clone MX19.8	JF521598 (100/99)
WE1	OTU-6	Clone WE1.5	HE648945	Alphaproteobacteria; Rickettsiales	Uncultured bacterium clone SHFG464	FJ203077 (99/98)
WE2	OTU-6	Clone WE2.2	HE648946	Alphaproteobacteria; Rickettsiales	Uncultured bacterium clone SHFG464	FJ203077 (99/97)
WB4	OTU-6	Clone WB4.44	HE648947	Alphaproteobacteria; Rickettsiales	Uncultured bacterium clone SHFG464	FJ203077 (99/98)
YB1	OTU-7	Clone YB1.1	HE648948	Alphaproteobacteria; Rhodobacterales; Rhodobacteraceae	*Labrenzia aggregata* strain 2PR58-2	EU440961 (100/99)

* OTUs were delineated at 97% sequence similarity

Table S4. Results of the variation partitioning analysis using three (phylogeny, environment and geography) and two (phylogeny, environment) explanatory tables. Adjusted R^2 values are shown, with values > 20% indicated in bold. Negative fractions indicate that two explanatory variables have strong and opposite effects on the dependent variable.



Fraction	Total bacterial diversity	Phyllobacteriaceae	Rhizobiaceae	*Labrenzia*	*Mycoplasma*	*Rickettsia*	Bacteroidetes	Flavobacteriaceae
[a+d+f+g] = X1	**0,27**	-0,19	0,16	**0,25**	0,04	**0,57**	0,14	**0,68**
[b+d+e+g] = X2	**0,28**	-0,11	-0,06	0,16	0,08	-0,01	**0,42**	**0,99**
[c+e+f+g] = X3	0,07	-0,01	0,04	0,14	-0,03	-0,08	-0,10	**0,31**
[a+b+d+e+f+g] = X1+X2	**0,34**	-0,73	0,02	**0,50**	0,05	**0,49**	**0,44**	**1,00**
[a+c+d+e+f+g] = X1+X3	**0,26**	-0,33	0,11	**0,26**	-0,07	**0,49**	**0,33**	**0,73**
[b+c+d+e+f+g] = X2+X3	**0,23**	-0,17	-0,19	0,06	0,06	-0,16	**0,47**	**0,99**
[a+b+c+d+e+f+g] = All	**0,36**	-1,05	0,10	**0,47**	**0,27**	**0,33**	**0,69**	**1,00**
[a] = X1 | X2+X3	0,13	-0,87	**0,29**	**0,41**	**0,21**	**0,49**	**0,22**	0,00
[b] = X2 | X1+X3	0,10	-0,71	-0,01	**0,22**	**0,34**	-0,16	**0,36**	**0,27**
[c] = X3 | X1+X2	0,02	-0,31	0,08	-0,03	**0,22**	-0,16	**0,25**	0,00
[d]	0,06	**0,54**	-0,22	-0,29	-0,24	0,08	**0,21**	**0,41**
[e]	-0,03	0,17	-0,13	0,03	-0,33	0,08	-0,06	0,04
[f]	-0,07	**0,25**	-0,21	-0,06	-0,24	0,01	-0,20	0,00
[g]	0,15	-0,11	**0,29**	0,20	**0,32**	-0,02	-0,09	**0,27**
[h] = Residuals	**0,64**	**2,05**	**0,90**	**0,53**	**0,73**	**0,67**	**0,31**	0,00



Fraction	Total bacterial diversity	Phyllobacteriaceae	Rhizobiaceae	*Labrenzia*	*Mycoplasma*	*Rickettsia*	Bacteroidetes	Flavobacteriaceae
[a+b] = X1	**0,30**	-0,03	0,09	**0,21**	0,20	**0,48**	0,16	**0,58**
[b+c] = X2	**0,23**	-0,11	-0,06	0,16	0,08	-0,01	**0,42**	**0,99**
[a+b+c] = X1+X2	**0,36**	-0,39	-0,07	**0,53**	**0,21**	**0,38**	**0,49**	**1,00**
[a] = X1 | X2	0,13	-0,28	0,00	**0,38**	0,12	**0,40**	0,07	0,01
[b]	0,17	**0,26**	0,09	-0,16	0,08	0,08	0,08	**0,58**
[c] = X2 | X1	0,06	-0,36	-0,16	**0,32**	0,00	-0,10	**0,34**	**0,42**
[d] = Residuals	**0,64**	**1,39**	**1,07**	**0,47**	**0,79**	**0,62**	**0,51**	0,00