How do microbiota associated with an invasive seaweed vary across scales? Accepted Article

Guido Bonthond, Till Bayer, Stacy A Krueger-Hadfield, Francisco Rafael Barboza, Masahiro Nakaoka, Myriam Valero, Gaoge Wang, Sven Künzel, Florian Weinberger

MR. GUIDO BONTHOND (Orcid ID : 0000-0002-9823-6761)

DR. STACY A KRUEGER-HADFIELD (Orcid ID : 0000-0002-7324-7448)

DR. MYRIAM VALERO (Orcid ID : 0000-0002-9000-1423)

Article type : Original Article

How do microbiota associated with an invasive seaweed vary across scales?

Guido Bonthond[1]*, Till Bayer[1], Stacy A. Krueger-Hadfield[2], Francisco Rafael Barboza[1], Masahiro Nakaoka[3], Myriam Valero[4], Gaoge Wang[5,6], Sven Künzel[7], Florian Weinberger[1]

[1] GEOMAR Helmholtz Centre for Ocean Research Kiel, Düsternbrooker Weg 20, 24105, Kiel, Germany.

[2] Department of Biology, University of Alabama at Birmingham, 1300 University Blvd, CH464, Birmingham, AL, 35294, USA

[3] Akkeshi Marine Station, Field Science Center for Northern Biosphere, Hokkaido University, Aikappu 1, Akkeshi, Hokkaido 088-1113, Japan

[4] UMI EBEA 3614, CNRS, Sorbonne Université, UCCh, UACH, Station Biologique de Roscoff, CS 90074, 29688 Roscoff, France

[5] College of Marine Life Sciences, Ocean University of China, 5 Yushan Road, Qingdao 266003, China

[6] Institute of Evolution and Marine Biodiversity, Ocean University of China, Qingdao 266003, China

[7] Max Planck Institute for Evolutionary Biology, Plön, Germany

* Corresponding author: gbonthond@geomar.de

Running title: Microbiota patterns in an invasive holobiont

Abstract

Communities are shaped by scale dependent processes. To study the diversity and variation of microbial communities across scales, the invasive and widespread seaweed *Agarophyton vermiculophyllum* presents a unique opportunity. We characterized pro- and eukaryotic communities associated with this holobiont across its known distribution range, which stretches over the northern hemisphere. Our data reveal that community composition and diversity in the holobiont vary at local but also larger geographic scales. While processes acting at the local scale (i.e., within population) are the main structuring drivers of associated microbial communities, changes in community composition also depend on processes acting at larger geographic scales. Interestingly, the largest analyzed scale (i.e., native and nonnative ranges) explained variation in the prevalence of predicted functional groups, which could suggest a functional shift in microbiota occurred over the course of the invasion process. While high variability in microbiota at the local scale supports *A. vermiculophyllum* to be a generalist host, we also identified a number of core taxa. These geographically independent holobiont members imply that co-introduction of specific microbiota may have additionally promoted the invasion process.

Keywords: microbiota, microbiome, haplodiplontic, holobiont, geographic scales, global diversity patterns, invasive species

Introduction

Virtually all multicellular organisms, including seaweeds, are holobionts (Margulis 1990, but redefined in Bordenstein & Theis 2015) and interact with complex and variable microbiota. Unlike terrestrial plants, seaweeds lack a sophisticated root system and utilize their resources from or through the surrounding medium (i.e., seawater). The community on the algal surface, mainly composed of microbes, but including multicellular eukaryotes as well, occupies a delicate niche. It colonizes the interface between organism and environment and forms a barrier which both threats (e.g., pathogens) and resources (e.g., light) need to cross to reach the host (Wahl, Goecke, Labes, Dobretsov & Weinberger 2012; Egan et al., 2013). Red macroalgae are known in particular to manipulate the composition of associated epibiota by the production of metabolites (e.g., Weinberger & Friedlander 2000; Harder, Campbell, Egan & Steinberg 2012; Saha & Weinberger 2019). The associated community structure is also dependent on local microbial availability, which in turn is shaped by interactions within the microbial community and environmental variables. In other words, microbial communities associated with seaweeds and other holobionts are shaped by processes that act on different ecological scales, such as the chemistry of the host and environmental conditions (e.g., salinity, temperature and spatial scales with which such physical parameters vary). While species abundance and distribution patterns are controlled by scale-dependent processes (McGill 2010), microbial diversity patterns across scales have only

recently received interest (e.g., Martiny et al., 2011; Tedersoo et al., 2014; Locey & Lennon 2016; Lindström & Langenheder 2012).

The invasive seaweed *Agarophyton vermiculophyllum* (Ohmi) Gurgel et al. (commonly known under the synonym *Gracilaria vermiculophylla*) represents a unique opportunity to study variation in microbiota across different ecological scales. Like most Rhodophyta, this alga has a haplodiplontic lifecycle in which haploid male gametophytes fertilize haploid female gametophytes, giving rise to diploid tetrasporophytes that in turn produce male and female gametophytes through meiosis. Both gametophytes and the tetrasporophytes develop into independent adults that are morphologically only distinguishable by reproductive structures (Krueger-Hadfield et al., 2016). *Agarophyton vermiculophyllum* originates from the Northwest Pacific, where its distribution range stretches from near the tropics (21.4° N; Hu et al., 2018) to high latitudes up to ~52° N in the cold temperate zone (Skriptsova, Titlyanova & Titlyanov 2001). Over the course of the invasion process the distribution range expanded to the Eastern Pacific (Bellorin, Oliveira & Oliveira 2004), Eastern Atlantic (Rueness 2005) and Western Atlantic (Thomsen, Gurgel, Fredericq & McGlathery 2006), where it presently also covers considerable latitudinal ranges (Krueger-Hadfield et al., 2017 and references therein). This alga thrives in a diversity of environments that differ in temperature, salinity, nutrient availability, sedimentation, grazing and epiphytic overgrowth (reviewed in Hu & Lopez-Bautista 2014). As an ecosystem engineer it has profound consequences on the local environment and can for example increase habitat availability and faunal biomass (Byers, Gribben, Yeager & Sotka 2012), change ecosystem functions (Ramus, Silliman, Thomsen & Long 2017) compete with native ecosystem engineers (Hammann, Buchholz, Karez & Weinberger 2013), alter nutrient cycling (Gonzalez, Smyth, Piehler & McGlathery 2013) and affect commercial fishing (Freshwater et al., 2006). The distribution range covers local to global geographic scales, across which different processes influence the associated microbial community. For instance, different micro-environments within one single algal individual may harbor different communities (e.g., Morrissey, Çavaş, Willems & De Clerck 2019) which at the same time depend on environmental parameters acting at local (e.g., Campbell, Marzinelli, Gelber & Steinberg 2015) and regional (e.g., Lindström & Langenheder 2012) scales. Microbial availability and connectivity may still be important at the coastal or continental scale (e.g., Martiny et al., 2011; Sunagawa et al., 2015) and effects related to the invasion process, such as genetic diversity of the host or potential adaptations, are likely more relevant at continental and global scales (e.g., Arnaud-Haond et al., 2017). Since the nonnative range is spatially far larger and the invasion process has been facilitated by multiple introductions (Krueger-Hadfield et al., 2017) it is probable that one or multiple selection processes have occurred, both in the native and nonnative ranges. Although such adaptions have not been identified genetically, several lines of evidence corroborate that populations from the nonnative range tend to be more tolerant to several stressors, including grazing (Hammann et al., 2016a), temperature and salinity extremes (Hammann, Wang, Boo, Aguilar-Rosas & Weinberger 2016b; Sotka et al., 2018) and epiphytic overgrowth (Wang et al., 2017a; Wang et al., 2017b).

Experimental work has indicated that *A. vermiculophyllum* populations from native and nonnative ranges are chemically better defended against epiphytic bacterial settlers from their own range (Saha, Wiese, Weinberger & Wahl 2016). This does not only suggest that *A. vermiculophyllum* is (chemically) well-equipped to manipulate epibiota, but also implies that the defense is fine-tuned to the environment at some spatial scale. The synthesis of molecules that interact with the microbial community is plastic, as the production or increase in production may also depend on external conditions. These traits might importantly aid a seaweed host in maintaining an epiphytic community that is overall beneficial. Whether such plasticity, increased plasticity through adaptation, or more specific adaptations have contributed to its success as an invader remains unclear. However, the abovementioned studies strongly suggest that interactions with the associated microbial community have been important in the invasion process. Whether changes in the interaction between host and microbes (such as shown in Wang et al., 2017a; Wang et al., 2017b; Saha et al., 2016), reflect a single adaptation of the host that occurred during the invasion, or rather adaptation or acclimation acting on more local scales is an important question to which the answer may importantly contribute to our understanding of the invasion success of *A. vermiculophyllum* and other invasive species.

While specific host-microbe interactions between *A. vermiculophyllum* and epibiota have been studied (Saha & Weinberger 2019; Wang et al., 2017a; Wang et al., 2017b; Saha et al., 2016), the composition of the epiphytic and endophytic microbial communities and variation therein across the native and nonnative ranges and smaller scales has not been characterized. It is unknown what commonly the most abundant microbial endo- and epiphytic taxa are and whether *A. vermiculophyllum* hosts core microbiota (i.e., a subset of taxa that is persistently associated, independent of the geographic origin; Shade & Handelsman 2011). Characterizing the putative core microbiota and how associated microbial communities vary across scales is crucial to identifying the ecological processes and symbionts that are relevant to the invasion process.

To this end, we collected individuals from fourteen populations and characterized microbiota using high-throughput amplicon sequencing of the prokaryotic 16S-V4 and the eukaryotic 18S-V7 ribosomal DNA regions aiming to answer the following two primary questions: **(*i*)** Do microbial communities differ between the algal surface (epibiota) and the internal tissue (endobiota) micro-environments? **(*ii*)** How do epiphytic communities vary across geographic scales? Furthermore, we expanded subsets of the sampling with a second year and the collection of haploid individuals, to address two additional questions: ***(iii)*** How do microbial communities vary between years within populations? ***(iv)*** Do microbial communities differ among host ploidies and sexes (*sensu* Hughes & Otto, 1999)?

Methods

Sampling design

Since *A. vermiculophyllum* spores require hard substratum to settle and germinate and adults lack the ability to produce a new holdfast (Krueger-Hadfield et al., 2016), we specifically collected individuals that were fixed to hard substratum indicating spore recruitment to avoid sampling fragments recently separated from the same individual or clonally reproduced drifting individuals. To avoid confounding possible effects of life cycle stage (i.e., ploidy and/or sex) with other effects related to the geographic scales, we identified life cycle stages and compared micro-environments and geographic scales using only diploid individuals, processing six diploids per collection site. To sample groups that can be considered populations, only sites where individuals that were presumably sexually reproductive were visited (i.e., sites known to include diploids and haploid females and males fixed to hard substrata; Krueger-Hadfield et al., 2016). We selected the collection sites such that the overall design was organized into five nested scales, which in ascending hierarchy were: (*i*) individuals, (*ii*) populations, (*iii*) ecoregions (as defined in Spalding et al., 2007), (*iv*) continental coasts, and (*v*) ranges (native and nonnative; Figure 1). To distinguish local from regional effects, all populations were sampled in pairs, separated by approximately 100 km. Three ecoregions from the Asian coast (Yellow Sea, Oyashio Current and Northeastern Honshu), two from the European coast (North Sea and Celtic Seas) and one from the Atlantic and Pacific coasts of the North American continent (Virginian and Northern California, respectively) were visited. The highest hierarchical scale (i.e., range) included one native and three nonnative continental coasts (Figure 1).

Collection and sample processing

Agarophyton vermiculophyllum specimens were collected from 14 September to 21 September 2016 and 24 August to 23 September 2017. At each site at least ten individuals and three water samples of 50 mL were taken. Algae were sampled with gloves, directly stored in new and separate freezer storage bags in a cooling box until arrival at the laboratory where they were immediately stored at 4 ^{0}C until processing. All processing occurred within a maximum of 12 hours after collection.

One gram of each alga was taken from the terminal part of the thallus (including branches with apices) and transferred to a 50 mL tube containing 15 sterile glass beads of 4 mm in 7.5 mL of artificial seawater. The artificial seawater was prepared from sterile distilled water and 24 gL^{-1} NaCl. To separate the epiphytic community, we followed the same method as Saha et al. (2016) with slight modifications. Tubes were placed on a vortexer for 3 min after which the water was filtered through a 0.2 µL PCTE filter with a vacuum pump. Subsequently, a new volume of 7.5 mL sterile distilled water was added, vortexed for 3 min and filtered through the same filter. Water samples were filtered directly. Filters were stored in absolute ethanol until DNA extraction. From the remaining algal tissue, a small fragment (2-4 cm) was preserved in DESS solution (25% DMSO, 2.5M EDTA and NaCl saturated) for DNA extraction of endobiota. One additional fragment was preserved in silica gel for ploidy identification. Diploids were identified morphologically with a dissecting

microscope before sample processing or by *post-hoc* microsatellite genotyping, following the amplification and genotyping methods described in Krueger-Hadfield et al. (2016). In brief, DNA was extracted from the thalli preserved in silica and ploidy was determined using ten polymorphic microsatellite loci (Kollars et al., 2015). We considered an individual diploid if at least one locus was heterozygous (Krueger-Hadfield et al., 2016). Based on either morphological or microsatellite identification of life cycle stage, six diploids were selected for amplicon sequencing. In order to test differences in associated microbiota among life cycle stages, we collected at least thirty individuals from the four North American populations and identified diploids, haploid females and haploid males based on the presence of tetrasporangial sori (diploid tetrasporophytes), cystocarps (female gametophytes) or spermatangial sori (male gametophytes) using a dissecting microscope (see Krueger-Hadfield et al., 2018). From these sites, six diploids and five of each haploid stage (i.e., sexes) were processed for characterizing microbial communities. Finally, from the four populations collected in 2016, ploidies and sexes were not identified (morphologically nor molecularly) and the microsatellite method was unsuccessful for one of the Japanese populations from 2017 (Futatsuiwa). Therefore, instead of six diploids all ten individuals were included in the sequencing for the populations from 2016 and the Futatsuiwa population from 2017 (Table S1).

DNA extraction and sequencing to characterize microbial communities

In the laboratory, ethanol was evaporated from preserved filters with a vacuum centrifuge at 30 °C for one to two hours. The tissue fragments were taken from the DESS solution and rinsed two times with DNA free water. We note that it is possible that certain firmly attached epibiota could have remained associated with the surface throughout this treatment and that differences between endo- and epiphytic communities may be underestimated. DNA was extracted with the ZYMO Fecal/soil microbe kit (D6102; ZYMO Research, Irvine, CA, USA) following the manufacturer's protocol. The 16S-V4 region was amplified with the primers 515F (S-*-Univ-0515-a-S-19) and 806R (S-D-Arch-0786-a-A-20; see Klindworth et al., 2013) and 18S-V7 region with F-1183mod and R-1443mod (Ray et al., 2016). To prepare amplicon libraries, we applied the two-step PCR strategy from Gohl et al. (2016), using the KAPA HIFI HotStart polymerase (Roche, Basel, Switzerland) and the same indexing primers. The procedure for the 18S-V7 fragment was identical, except for the target primer sequences. Amplicon concentrations from the second PCR were estimated from gel pictures and amplicons were pooled accordingly into 16S and 18S libraries which were purified by a gel extraction step, combined in a 5:1 ratio and sequenced at the Max-Plank-Institute for Evolutionary Biology (Plön, Germany) on the Illumina 2x300 MiSeq platform. Libraries included three negative controls from the DNA extractions, three negative PCR controls and two mock communities as positive control (HM-782D, Bei Resources, Manassas, VA, USA). The fastq files were demultiplexed (0 mismatches) after which sequence assemblage and quality filtering was performed using Mothur software (v1.40.5; Schloss et al., 2009) and the SILVA alignment (Quast et al., 2013 release 132) to denoise and classify the full set of reads. Unique sequences were clustered into OTUs with the

opticlust algorithm based on the traditional 3% dissimilarity criterion. Mitochondrial, chloroplast, eukaryotic and unclassified sequences were removed from the 16S dataset. We also removed sequences from the 18S dataset that were unclassified, classified as Bacteria, Archaea, higher plants and the genus of the host (*Gracilaria* in the SILVA 132 release). Since we were primarily interested in epiphytes (that is, sessile taxa growing on the surface of the alga) and sessile animals are not known as prevalent epifauna of *A. vermiculophyllum* (Nyberg, Thomson & Wallentinus, 2009), sequences classified to Animalia were also discarded from the 18S dataset. While in the present work referred to as the eukaryotic microbiota, this constituent of the community also contains many multicellular epiphytes. Finally, OTUs that were singletons in the full dataset and samples with less than a 1,000 16S sequence counts and samples with less than a 100 18S sequence counts were also excluded from downstream analyses. Raw de-multiplexed amplicon reads and metadata were deposited in the SRA database (accession: PRJNA564581).

Functional profiling

Functional profiles were obtained with tax4fun software (Aßhauer, Wemheuer, Daniel & Meinicke 2015) which uses the SILVA reference alignment and Kyoto encyclopedia of Genes and Genomes (KEGG; Kanehisa et al., 2013) to predict metagenomics profiles. Since the package was not yet compatible with the Silva 132 release at the time of analysis, OTUs were reclassified in Mothur using the Silva 123 release. Aiming to characterize general functional groups within the communities we selected genes essential to autotrophy (ribulose-1,5-biphosphate carboxylase oxygenase; K01601), aerobic heterotrophy (cytochrome c oxidase subunit III; K02276) and nitrogen fixation (nitrogenase iron protein; K02586). We also included anaerobic heterotrophs as a functional group by combining the mean gene count values for sulfate reduction (adenylylsulfate reductase; K00394), denitrification (methane/ammonia monooxygenase; K10944), acetogenesis (formate-tetrahydrofolate ligase; K01938) and fumarate respiration (fumarate reductase; K00244), which we presumed constitute the majority of anaerobic respiration metabolisms.

Data analysis

For a total of 273 samples the 16S-V4 and 18S-V7 regions were sequenced and the models explained below were applied to both datasets. Importantly, to adjust for the effect of differences in sequencing depth, inherent to amplicon sequencing based community data, we included the natural logarithm of the sequencing depth (LSD) as a continuous variable in all multi- and univariate models. To reduce complexity and computation time, datasets were trimmed for all multivariate analyses to the 95% most abundant OTUs. Sequence counts were expressed in proportion to the total sample count (relative counts) and sorted in descending order after which cumulative percentages were computed. To compare community abundance-based compositions, multivariate

generalized linear models (mGLMs) were fitted with the manyglm() function from the R package mvabund (Wang, Naumann, Wright & Warton 2012; Warton, Wright & Wang 2012).

The mGLMs assumed a negative binomial distribution with a natural logarithm as link function. In addition to providing a deviance and p-value at the multivariate level (the 'community response'), the mGLM computes the same statistics at the univariate level (each OTU) which we used to count OTUs that were differentially abundant across factors of interest ($p_{unadjusted} < 0.05$). Furthermore, by obtaining both uni- and multivariate outputs simultaneously, we could use the mGLMs to compare communities and characterize core microbiota coherently using a single model. To confirm whether mGLMs satisfied the distribution and mean-variance assumptions, *QQ-plots* and *residuals vs. fitted-plots* were visually inspected.

To address the first question – whether microbial communities are different among micro-environments – we tested the alternative hypothesis that microbial abundance-based community composition varies among micro-environments. Since mixed models are currently not available in the mvabund package for mGLMs, we used a two-step modelling approach. To first adjust the data for geographic effects we included the factor 'population identity' in the first model along with the LSD. On the residuals of this first model we ran a second model with micro-environment (including water, surface and tissue) as response variable and a Gaussian distribution (the residuals from this first model were in the link function and therefore normally distributed). The p-value of the overall effect of micro-environment was obtained using the anova.manylm() function, using 500 bootstrap iterations. We obtained the between-group comparisons and corresponding p-values of the levels within micro-environment by running summaries with different reference levels using the summary.manylm() function and also 500 bootstrap iterations (see Table S2a for the full output). Multivariate dissimilarities were visualized with non-metric dimensional scaling (nMDS, with the R package *vegan*; Oksanen et al., 2013), using the raw data and rescaled residuals from the mGLMs.

OTU richness, evenness and predicted functionality were compared with mixed linear models using the lme4 R package (Bates, Mächler, Bolker & Walker 2015). Since richness and evenness are the variables of which diversity is essentially a function, we used both parameters as indicators of structural differences in communities. We calculated asymptotic richness (S_{chao}, based on the chao1 estimator, Chao & Bunge 2002) and rarified richness (S_n). For S_n all samples were rarified to the read count of the sample with the lowest sequencing depth (1,008 for prokaryotes and 110 for eukaryotes). From a strictly theoretical point of view, both S_{chao} and S_n are expected to be independent of the sequencing depth (but the accuracy of the asymptotic estimate should increase with more reads) and are not necessarily correlated. We used the probability of interspecific encounter (PIE; Hurlbert 1971) as a measure of evenness (McGlinn et al., 2018). To compare functional group abundances between micro-environments and variation in group abundances across geographic scales the same model structures were used. LSD and micro-environment were included as fixed variables and population and individual identity as random variables. To meet the assumption of normality PIE

was logit transformed and the gene counts representing the predicted functional group abundances were log transformed. Assumptions on the distribution were verified and p-values for differences among water, epi- and endobiota were obtained by *post-hoc* TukeyHSD tests (see Table S2b for the full statistical output). Marginal and conditional R^2 values (variation explained by fixed effects and fixed plus random effects, respectively) were computed with the r.squaredGLMM function from the lme4 R package (Nakagawa & Schielzeth 2013).

To characterize how variation in epibiotic communities is partitioned across geographic scales, we used the subset of data representing diploid algal surface samples, comprising all fourteen populations. We fitted mGLMs including the LSD and the nested structure of spatial scales (range/coast/region/population) as explanatory variables. OTU richness, PIE and predicted functional group abundances were analyzed using mixed linear models (Table S2c). The LSD was included as fixed effect and the nested structure of spatial scales as random effect, resulting in the following model structure: ~ LSD + (1|range/coast/region/population), based on the approach used by Messier, McGill & Lechowicz (2010). We extracted variance estimates and calculated the corresponding 95% confidence intervals by bootstrapping the models with a 1,000 iterations. Multivariate GLMs with the same model structure were also fitted on subsets of the data containing only the water samples to verify the assumption that environmental microbiota vary across geographic scales (see Table S2a).

The four populations collected in both 2016 and 2017 were first analyzed with an mGLM. The lowest spatial scale (population identity) was included to represent spatial effects, under which the year was nested. Multivariate distances were visualized in dendrograms using a hierarchical clustering approach for which we used the residuals from an mGLM with LSD as single response variable to adjust for the effect of sequencing depth. The clustering was conducted with the R package pvclust, resampling with multiscale bootstrapping and a 1,000 iterations (Suzuki & Shimodaira 2006). The nested structure was specified as random in univariate models to analyze how much variation was explained by the year after explaining the geographic variation (see Table S2d).

To answer the question whether communities vary among host ploidies and sexes we tested the alternative hypothesis that different lifecycle stages host different microbial communities, using the four North American populations, for which all lifecycle stages had been collected. The mGLMs included the variables LSD, micro-environment (from which the water level was removed), life cycle stage and population identity. Based on the lowest AICsum we did not consider the interaction between micro-environment and lifecycle stage in the mGLMs for both pro- and eukaryotic datasets. In the univariate models ‘population’ and ‘individual identity’ were specified as random variables while ‘LSD’, ‘micro-environment’ and ‘lifecycle stage’ were fixed. As the interaction between micro-environment and lifecycle stage yielded a lower AICc for richness estimates, we retained the interaction in all univariate models (see Table S2e for statistical output and Table S2f for model comparisons).

Since the concept of 'cores' is somewhat flexible (Shade & Handelsman 2011), we characterized core microbiota based on two approaches, applying an occurrence criterion (occurrence core) and a model-based composition criterion (composition core). Both approaches were performed on the subset of our data containing the six populations for which tissue and surface samples were available. We considered OTUs detected in strictly 100% of the water, algal surface or tissue samples as members of the water, epi- and endophytic occurrence cores. Those present in 100% of both tissue and surface samples were considered as part of the algal core. The composition core was determined based on the univariate results of the mGLMs applied on the same dataset to compare microbial communities between micro-environments. As these mGLMs were run on data that had been adjusted for spatial effects (by using the residuals from mGLMs including population identity as explanatory variable) we considered OTUs that were significantly more abundant ($p_{unadjusted} < 0.05$) to be micro-environment specific and geographically independent. OTUs fulfilling this criterion were identified using the coefficients and p-values from the mGLM summaries and classified accordingly to either water, surface, tissue or alga (no difference between surface and tissue but in both more abundant than in the water).

We explored correlations between OTUs within the holobiont following the joint modeling approach (Warton et al., 2015). Under the assumption that the residuals have been adjusted for the effects of micro-environments (water, surface and tissue) and population identity (and all physical-chemical variables confounded in them), these correlations would represent the variation driven by biological variables and could thus hint at patterns of co-occurrence or mutual exclusion among microbes. The untransformed residuals of pro- and eukaryotic mGLM models (normally distributed in the scale of the link function) were concatenated and all Pearson correlation coefficients and corresponding p-values were computed between the 25 most abundant pro- and eukaryotic OTUs. Correlations with p-values > 0.05 were discarded.

Results

Sequencing summary

A total of 4,537,292 prokaryotic reads and 24,793 OTUs remained from 40 water, 42 endo- and 147 epiphytic samples after removal of low quality sequences, singletons and 16S samples with less than a 1,000 reads. Samples ranged from 1008 to 112,921 in read counts with a median of 13,534. Of the eukaryotic dataset 576,509 sequences and 1,919 OTUs from 37 water, 40 endo- and 138 epiphytic samples remained after quality filtering and removal of 18S samples with less than a 100 reads. Samples ranged from 110 to 15,883 in read counts with a median of 1,868. The three most abundant prokaryotic OTUs were classified to the genera *Pleurocapsa* (Cyanobacteria) and *Granulosicoccus* (*γ*-Proteobacteria) and to the family Rhodobacteraceae (*α*-Proteobacteria). The three most abundant eukaryotic OTUs were classified to the family Bacillariophyceae

(Diatomea), the order Ectocarpales (Phaeophyceae) and the genus *Navicula* (Diatomea). The most abundant phyla were Proteobacteria, Bacteriodetes and Cyanobacteria. While in epiphytic communities, Bacteriodetes was the second most abundant and Cyanobacteria were third, in endophytic communities this was reversed. The most abundant phylum in both endo- and epiphytic eukaryotic communities was the Ochrophyta (including Diatomea) and was followed by Florideophycidae. While in endophytic communities the third most abundant phylum was Chlorophyta, Peronosporomycetes (Oomycetes) was the third phylum in epiphytic communities (Figure S1). The removal of low read count samples reduced the number of individuals (initially 6) in a few populations (see Table S3).

(i) *Are microbial communities different among micro-environments?*

For both pro- and eukaryotic communities, the null-hypothesis that there are no differences in abundance-based community composition among micro-environments was rejected (p_{pro} and $p_{euk} < 0.01$). *Post-hoc* comparisons, using summaries with different reference levels resolved significant p-values for all pairwise combinations (p_{pro} and $p_{euk} < 0.01$). These differences are also apparent in the nMDS plots (Figure 2a and b). Overall, 28.9% and 29.6% of 3,054 and 226 prokaryotic and eukaryotic OTUs were differentially abundant, respectively (Figure 3a). Also for the diversity metrics the null-hypothesis was rejected (p_{pro} and $p_{euk} < 0.01$). Asymptotic and rarefied OTU richness of epibiota were higher than of endobiota in both pro- and eukaryotic communities. *Post-hoc* pairwise comparisons showed that in prokaryotic communities, richness of the water and epibiotic communities was similar, while richness of eukaryotic communities was higher in the water than both endo- and epiphytic communities (Figure 2c, e, only asymptotic richness shown, see Table S2b for the statistical output). The sequencing depth had a significant effect on S_{chao} in both pro- and eukaryotic communities. Despite S_{chao} being an estimate of the asymptotic richness (species count of the overall theoretical community), this is not unexpected for sequencing based community data (e.g., Smith & Peay 2014). The sequencing depth had no effect on the rarified richness ($p_{pro} = 0.30$, $p_{euk} = 0.73$) or the evenness parameter PIE (which is mainly depending on the shape of the rarefaction curve; McGlinn et al. 2018; $p_{pro} = 0.34$, $p_{euk} = 0.24$). Differences in predicted functional group abundances were found for autotroph, anaerobic heterotroph and diazotroph groups ($p < 0.01$). Whereas in auto- and diazotroph groups, counts increased from the water to epiphytic communities and from epi- to endophytic communities, predicted anaerobic heterotrophy was highest in the water and decreased in epi- and endophytic communities. The null-hypothesis was, however, not rejected for aerobic heterotrophy ($p = 0.48$; Figure 2g-j).

(*ii*) *How do epiphyte communities vary across geographic scales?*

Differences in epiphyte community composition were found across all geographic scales ($p_{range} = 0.03$, $p_{coast} = 0.03$, $p_{region} < 0.01$, $p_{population} < 0.01$ for the prokaryotic communities and $p_{range} < 0.01$, $p_{coast} < 0.01$, $p_{region} < 0.01$, $p_{population} < 0.01$ for eukaryotic communities). Differentially abundant OTUs ranged from 35.9% to 69.5% of 3251

prokaryotic OTUs and 36.4% to 64.9% of 231 eukaryotic OTUs, with for both datasets more differentially abundant OTUs at lower spatial scales (Figure 3b). A similar pattern can be observed from nMDS plots (Figure S2) where between-level differences are more pronounced at lower scales. Also across scales the sequencing depth was only significant for S_{chao} ($p_{LSD} < 0.01$ in both pro- and eukaryotic models) and explained considerable variation (47.0% in prokaryotic and 27.9% in eukaryotic communities). In line with the multivariate results, the smallest geographic scale (population) was the most important, explaining variation ranging from 27.5% prokaryotic S_{chao} to 61.7% in prokaryotic S_n (Figure 4). The highest scale (range) was estimated to explain a minor fraction of variation in prokaryotic S_{chao} and S_n (4.59% and 4.16% percent, respectively). For eukaryotic S_n, coast was estimated to explain 18.1% and range to explain 2.94% in eukaryotic PIE. Most variation in predicted functional group abundance was explained at the population level with percentages ranging from 25.2% (phototrophs) to 60.7% (anaerobic heterotrophs; without variation explained by LSD). Predicted autotroph and diazotroph abundances also varied at the highest hierarchical level (range) and explained 42.1% of the variation in autotrophs and 33.0% in diazotrophs (Figure 4). Within the water, pro- and eukaryotic microbiota varied compositionally at all scales ($p < 0.01$) and this was most pronounced at the population level where a number of OTUs (106 pro- and 11 eukaryotic) differed significantly (Table S2c, Figure S2i).

(*iii*) *How do microbial communities vary between years?*

The mGLMs resolved significant p-values for both population (representing the geographic effects) and year (in all cases $p < 0.01$). However, in prokaryotic communities geographic effects were relatively more pronounced with 64.9% of 2,849 OTUs differentially abundant among populations compared to 14.0% between years. In the eukaryotic communities 65.4% of 179 OTUs were differentially abundant among populations and within populations 26.8% between years (Figure 3c). In the hierarchical clustering of the prokaryotic OTU table all individuals clustered within their respective populations with high bootstrap values (≥ 99, Figure S3). Within the population clusters, with some exceptions, samples from different years clustered adjacently. For the eukaryotic data, the overall pattern was similar but with less bootstrap support for the nodes combining both years of the same populations and a few more individuals clustering within different populations. The effect of the higher geographic scales detected with the subset of the data including all 14 populations was not visible in these dendrograms. Variation in prokaryotic OTU richness and evenness was explained majorly by geographic effects (S_{chao} = 42.9%, S_n = 63.6% and PIE = 65.6%), while variation explained by year within population was little to none (S_{chao} = 0%, S_n = 4.2% and PIE = 5.7%). In eukaryotic communities, however, more variation was explained by the inter-annual changes (Figure S4, Table S2e).

(*iv*) *Are microbial communities different among host ploidies and sexes?*

In the mGLMs for prokaryotes and eukaryotes life cycle stage explained the least deviance of all included variables and was only significant for eukaryotic communities ($p_{eukaryote}$ = 0.02). However, while in this case the alternative hypothesis was not rejected, *post-hoc* comparisons did not resolve any significant pairwise differences (see Table S2a). The hypothesis that S_{chao}, S_n and PIE differ among life cycle stages was in all cases rejected. The interaction between life cycle stage and micro-environment was not significant either in any of the models (Figure S5).

Core microbiota and correlations

For each of the micro-environments (i.e., water, surface and tissue) a number of prokaryotic OTUs was present in 100% of the samples from the six populations from which both endo- and epiphytic samples were sequenced (Table S4). Based on the occurrence approach we counted two endophytic, five epiphytic, seven algal (present in 100% endo- and epiphytic samples) and two water core OTUs, which were all prokaryotes. The composition core counted 141 endophytic, 26 epiphytic, 123 algal and 166 water prokaryotic OTUs and one endophytic, 6 epiphytic, 6 algal and 32 eukaryotic OTUs (Figure 3a, Table S4). After removing the effects of geography and micro-environment, the correlations among the 25 most abundant pro- and eukaryotic OTUs with p-values < 0.05 were mostly positive for within domain comparisons (within prokaryotes 61 positive and 7 negative correlations and within eukaryotes 19 positive and 10 negative correlations). Between domain correlations were majorly negative (28 positive and 34 negative correlations; Figure S6).

Discussion

Endo- and epiphytic microbial communities

The tissue and surface of *A. vermiculophyllum* harbor specific communities that differ in terms of composition, diversity and predicted functionality from each other and from the surrounding water. This is perhaps not surprising as distinct endo- and epibiota were also recognized in green algae (Morrissey et al., 2019; Aires, Moalic, Serrão & Arnaud-Haond 2015). Overall, epiphytic communities are characterized by higher richness (two-fold or more) and evenness compared to endophytic communities. The decreased richness and evenness in the tissue compared to the surface may reflect that the surface is more exposed to the environment and more prone to community coalescence (Rillig et al., 2015). Further, chemical control by the host might be higher within the tissue and the biofilm on the algal surface could act as a biotic filter that prevents potential symbionts from reaching the endophytic microbial niche. Although richness and evenness are merely diversity parameters, both have been associated with functional community properties, including productivity (Bell, Newman, Silverman, Turner & Lilley 2005) and community stability (Wittebolle et al., 2009; Coyte, Schluter & Foster 2015). Differences in the predicted functional groups between tissue and surface suggest that endophytic communities are more autotrophic and diazotrophic, while epiphytic communities are more

heterotophic (anaerobically). It is noteworthy that one of the most abundant OTUs, which was also a member of the algal core in both the occurrence and compositional approaches, was classified to the *Pleurocapsa* (Cyanobacteria; Figure S1a). Speculatively, such cyanobacteria might be the main group of potentially autotrophic endophytes and may even be able to fix inorganic nitrogen. Many terrestrial plants (e.g., legumes) are known to form mutualistic relationships with diazotrophic bacteria and therefore have an advantage in nitrogen limiting environments. For seaweeds, such associations have only been found in few occasions (Rosenberg & Paerl 1981; Gerard, Dunham & Rosenberg 1990; Raut, Morando & Capone 2018). Some terrestrial plants hosting diazotrophs are highly successful invaders (e.g., *Myrica faya* in Hawai; Vitousek, Walker, Whiteaker, Mueller-Dombois & Matson 1987). Hypothetically, an intimate relationship with Cyanobacteria capable of fixing dinitrogen – coastal marine waters are commonly nitrogen limited (Vitousek & Howarth 1991) – could provide *A. vermiculophyllum* a substantial advantage to survive in nonnative habitats. However, specialized structures such as nodules like in *M. faya* have not been identified in *A. vermiculophyllum* and future experimental work is required to determine the functional activity of these prevalent Cyanobacteria and other members of endo- and epiphytic cores.

Core microbiota

Within the high variability of microbial communities, core microbiota (or core microbiomes) represent a signature of stability. Different concepts are used to define microbial community cores based on occurrence, composition, persistence or connectivity (see Shade & Handelsman 2011). For occurrence cores, a range of percentage thresholds has been applied (e.g., 12% to 100%, Astudillo-García et al., 2017), and OTUs may be collapsed into higher taxonomic levels. We applied a highly conservative 100% occurrence threshold without collapsing OTUs. Since the caveat of an occurrence core is that the probability of detecting core taxa is directly correlated with abundance we also defined a composition core with a model based approach, using the univariate results of the mGLMs comparing composition between micro-environments. With this approach we detected substantially more core taxa (303 compared to 14 associated to tissue, surface or alga). While this may be less strict compared to the occurrence criterion, the model controls for effects from sampling location and the increased abundances of these OTUs endophytically, epiphytically or in the alga compared to the water, thus representing a geographically stable signature within the high amount of variation.

Despite the considerable geographic distribution, with both occurrence and composition based criteria a number of core taxa was detected. This contrasts somewhat with recent studies on seaweed associated microbiota in the green algal genera *Caulerpa* and *Ulva* (Burke, Steinberg, Rusch, Kjelleberg & Thomas 2011; Roth-Schulze et al., 2018; but see Arnaud-Haond et al., 2017). In these algae functional – but not taxonomic – cores were detected and it was suggested that the assemblage of microbial communities in seaweeds might be shaped by stochasticity and abiotic environmental filtering, selecting for functions rather than specific taxa

(Burke et al., 2011; Roth-Schulze et al., 2018). In contrast, a recent study comparing microbiota associated with the kelp *Ecklonia radiata* showed that the microbial community composition was more dependent on the host-condition rather than the local and regional environments (Marzinelli et al., 2015). For red algae, taxonomic and functional cores have, to our knowledge, not been characterized, nor have associated microbiota been compared across multiple geographic scales. Our results may reflect that the environment (i.e., a composite of the host and the local environment) filters taxa more strongly in *A. vermiculophyllum* than in green algae that have been studied previously. However, function and taxonomy are correlated and the filtering of specific functions might indirectly also filter for certain taxa and *vice versa*.

The detected core microbes are either globally available taxa or represent associations that predate the invasion process and originate from the native range. The co-introduction of microbes with specific functions in the holobiont is thought to be an important facilitator of successful invasion (Arnaud-Haond et al., 2017; Rodríguez-Echeverría, Le Roux, Crisóstomo & Ndlovu 2011). Only few of the detected core microbiota were found in the water column and *A. vermiculophyllum*'s core may thus largely constitute co-introduced microbes. The ability of an invasive seaweed to maintain specific taxa with beneficial functions might provide an advantage over competitors and perhaps even protect against disadvantageous microbiota from the environment that are unable to settle in the already populated niche. As shown in Saha & Weinberger (2019), *A. vermiculophyllum* is associated with cultivatable microorganisms that provide it with protection from microbial pathogens. At the same time the host produces metabolites that selectively promote the settlement of these protective bacteria. Such mechanisms may be general and directed at taxonomic or functional groups or be specific and targeted at microbes that are part of the core microbiota. However, to test such hypotheses it would be necessary to disentangle the filtering by different environmental components from one another (i.e., that from the local abiotic and biotic environments and that of the host itself).

Variation in epibiota across scales

Given that previous studies indicated that interactions between host and epibiota might have been involved in the invasion process of *A. vermiculophyllum* (Saha et al., 2016; Wang et al., 2017a; Wang et al., 2017b), we were specifically interested in epiphytic communities from a geographic perspective. While microbial communities have been studied across local and regional scales in other seaweeds (e.g., Lindström & Langenheder 2012; Campbell et al., 2015; Roth-Schulze et al., 2018), in this study we also compared variation in microbiota across more global geographic scales, reaching to that of a hemisphere. To a certain extent our results are in line with previous studies (Campbell et al., 2015; Roth-Schulze et al., 2018), which suggest that seaweed microbiota strongly depend on the local scale as we found for both pro- and eukaryotic datasets that most variation in OTU richness and evenness is explained among populations. In addition, abundances of all considered predicted functional groups varied mostly at the local scale. However, some functional groups

(autotrophs and diazotrophs) varied also substantially between ranges and the overall community composition differed among levels of all geographic scales. This indicates that epibiotic communities, in terms of composition, diversity and function, are mostly defined at the local scale, but also shows that processes acting at larger scales contribute to shaping the microbial community. What these processes are and from where they originate (the environment, the host or an interaction between environment and host) is important to identify as they may represent general mechanisms of relevance to other microbiota or holobionts. While compositionally all scales are important, in terms of predicted function we found that only the most local (population) and the most global scale (the native and nonnative ranges) mattered. On this scale, the invasion process itself is most relevant and may thus hint at a functional shift in the epibiotic community that occurred during the process. Given the strong genetic signature of the host itself between the ranges (Krueger-Hadfield et al., 2017), it is tempting to speculate that genetic differentiation may be an important process acting at this scale, underlying functional differences within epibiota. However, to test this hypothesis, the confounding genetic and geographic patterns need to be disentangled. To achieve this, an experimental approach is required where different populations are translocated or compared in the context of a common garden.

In addition to comparing geographic scales, we collected four populations in two consecutive summers and were able to characterize how these populations varied between years in terms of pro- and eukaryotic epibiota (Figure S3 and S4). Prokaryotic communities were relatively similar between years and these results suggest a temporally stable signature within the *A. vermiculophyllum* holobiont. Our result is in line with Lachnit, Meske, Wahl, Harder & Schmitz (2011), where denaturing gradient gel electrophoresis was used to characterize *A. vermiculophyllum* associated epibiota. Although based on only twelve samples from a single location Lachnit et al. (2011) observed that *A. vermiculophyllum* communities differed between summer and winter, but returned to similar compositions during the summer and winter of the subsequent year. While this pattern was especially clear within prokaryotic communities in our data, richness and evenness were less stable between years within eukaryotic communities. Therefore, seasonal stability in the structure of the associated community might be considerably lower for eukaryotic epibiota and the presence of these taxa may be less specific and more sensitive to stochasticity and local environmental conditions. The results from Lachnit et al. (2011) and this study emphasize the need to also consider temporal scales. Important environmental variables such as temperature that can strongly impact microbial communities (Sunagawa et al. 2015) vary both at geographic (e.g., with latitude) and temporal scales (e.g., with season). However, the effect of temperature on microbial communities is scale-dependent and may be unambiguous (Nottingham et al., 2018) or not (Hendershot et al., 2017). In the case of holobionts it is also important to identify processes originating from the host that affect diversity and composition of associated microbiota. To identify, disentangle and classify the importance of these abiotic and biotic processes, more scales must be considered and microbial communities should be studied along natural gradients in for example temperature, salinity and latitudinal space.

Agarophyton vermiculophyllum life cycle stages are characterized by morphological and ecological differences (Krueger-Hadfield et al., 2016). For the haplodiplontic life cycle to be stabilized over evolutionary timescales, ecological differentiation is essential (Hughes & Otto 1999). However, we found no differences in associated microbiota among reproductive diploids, haploid females and haploid males. This could indicate that traits related to chemical control of endo- and epiphytes are conserved across life cycle stages and the dominance of diploids in the nonnative range (Krueger-Hadfield et al., 2016) may thus be related to physiological traits (e.g., thallus strength, Lees et al., 2018) rather than traits associated to manipulation of microbiota. However, we sampled during a time of the year when populations tend to be healthy and with substantial biomass (Weinberger, Buchholz, Karez & Wahl 2008; Muangmai, Vo & Kawaguchi 2014) and we cannot predict whether microbiota remain similar among lifecycle stages under stressful conditions.

Holobiont and invasion biology

Communities are structured by processes that are scale dependent (McGill 2010 and references therein). In the case of communities associated with a holobiont these processes may originate from the host, the local environment in which the host resides or the combined environment of host and local conditions. Evidently, a microbial community that is benign, or even beneficial, contributes to the health and success of the host. An ability to exert a certain degree of control over the associated community could therefore be a valuable strategy. Adaptations to associate specific beneficial symbionts could have occurred over time and if such (core) microbes accompany the host to nonnative ranges, they may provide an advantage and promote invasiveness (i.e., accompanying mutualist hypothesis; Rodríguez-Echeverría et al., 2010). Given the tremendous environmental variability of microbes it might be advantageous to adapt simultaneously a more flexible strategy in associating beneficial microbial taxa and invest energy in mechanisms targeting more general groups (generalist host hypothesis; Rodríguez-Echeverría et al., 2011). While maintaining essential associations by hosting certain microbes as core members, a high degree of flexibility towards microbial associations may thus be of additional benefit for an invasive species. Arnaud-Haond et al. (2017) identified core microbiota from the invasive green alga *Caulerpa taxifolia* which they showed were associated with (or had accompanied) the host across native and nonnative ranges. Our data show high variability at the local scale and suggest that *A. vermiculophyllum* is a generalist host. However, the substantial number of core microbiota we found implies that this alga was also accompanied over the course of the invasion. *Agarophyton vermiculophyllum* may therefore qualify for both a generalist and accompanied host. Although originally based on legumes and rhizobial symbionts, Rodríguez-Echeverría et al. (2011) classified such accompanied generalist hosts as species with a high probability to invade, and with a typically fast invasion process. Given its current wide latitudinal distribution ranges along the nonnative coasts and in some cases rapid regional expansion (e.g., Nyberg et al,. 2009; Byers et al., 2012), the invasion process of *A. vermiculophyllum* may well

have been accelerated by accompanying symbionts. However, to characterize the functions of these core microbiota and to assess whether they have promoted the invasion, the *A. vermiculophyllum* holobiont needs to be studied in an experimental context.

Conclusions

This is the first study addressing seaweed microbiota with a sampling design at the scale of the known distribution range. The results presented here show that *A. vermiculophyllum* associated microbial communities differ within the individual between surface and tissue in composition, diversity and predicted functionality. Our data also reveal that *A. vermiculophyllum* hosts microbial core taxa, some of which are specific to the tissue or the surface, while others are associated to both niches. These taxa are conserved across the distribution range and may represent functionally important symbionts. Variation in epibiota is compositionally partitioned across all local and global scales, but mostly at the local scale (i.e., among populations), at which OTU richness and evenness varied most substantially. At the scale of the invasion (i.e., native and nonnative ranges), however, differences in community composition and predicted functional groups could suggest that a shift in epibiota occurred over the course of the invasion process. Finally, with microbiota that are locally highly variable and a geographically conserved core *A. vermiculophyllum* matches criteria of both the generalist host and the accompanying mutualist hypotheses, which may in part explain its remarkable success as an invasive species.

Acknowledgements

We thank Caitlin Cox and Mike Crowley at the Heflin Center for Genomic Sciences (University of Alabama at Birmingham) for use of the capillary sequencer for fragment analysis. We are also grateful to Nadja Stärck and Anna Neu for support during the fieldwork. This study was funded by grants from the Deutsche Forschungsgemeinschaft awarded to FW (DFG grant number WE2700/5-1) and TB (DFG grant number BA5508/2-1). Start-up funds from the University of Alabama at Birmingham to SAKH supported North American field work and fragment analysis. FRB acknowledges the financial support of the German Academic Exchange Service (DAAD) through the project Doctoral Programs in Germany 2015/16 (57129429).

References

Aires T., Moalic Y, Serrão EA, Arnaud-Haond S (2015) Hologenome theory supported by cooccurrence networks of species-specific bacterial communities in siphonous algae (*Caulerpa*). *FEMS microbiology ecology,* **91**, fiv067. https://doi.10.1093/femsec/fiv067.

Arnaud-Haond S., Aires T, Candeias R, Teixeira S, Duarte CM, Valero M, ... Serrão E (2017) Entangled fates of holobiont genomes during invasion: nested bacterial and host diversities in *Caulerpa taxifolia*. *Molecular ecology,* **26**, 2379-2391.

Aßhauer K. P., Wemheuer B, Daniel R, Meinicke P (2015) Tax4Fun: predicting functional profiles from metagenomic 16S rRNA data. *Bioinformatics,* **31**, 2882-2884. https://doi.10.1093/bioinformatics/btv287.

Astudillo-García C., Bell JJ, Webster NS, Glasl B, Jompa J, Montoya JM, ... Taylor MW (2017) Evaluating the core microbiota in complex communities: a systematic investigation. *Environmental microbiology,* **19**, 1450-1462. https://doi.10.1111/1462-2920.13647.

Bates D., Mächler M, Bolker B, Walker S (2015) Fitting linear mixed-effects models using lme4. *arXiv preprint arXiv:1406.5823,* **67**, 1-48. https://doi.10.18637/jss.v067.i01.

Bell T., Newman JA, Silverman BW, Turner SL, Lilley AK (2005) The contribution of species richness and composition to bacterial services. *Nature,* **436**, 1157. https://doi.10.1038/nature03891.

Bellorin A. M., Oliveira MC, Oliveira EC (2004) *Gracilaria vermiculophylla*: a western Pacific species of Gracilariaceae (Rhodophyta) first recorded from the eastern Pacific. *Phycological Research,* **52**, 69-79. https://doi.10.1111/j.1440-183.2004.00330.x.

Bordenstein S. R., Theis KR (2015) Host biology in light of the microbiome: ten principles of holobionts and hologenomes. *PLoS biology,* **13**, e1002226. https://doi.10.1371/journal.pbio.1002226.

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 of the United States of America,* **108**, 14288-14293. https://doi.10.1073/pnas.1101591108.

Byers J. E., Gribben PE, Yeager C, Sotka EE (2012) Impacts of an abundant introduced ecosystem engineer within mudflats of the southeastern US coast. *Biological Invasions,* **14**, 2587-2600. https://doi.10.1007/s10530-012-0254-5.

Campbell A. H., Marzinelli EM, Gelber J, Steinberg PD (2015) Spatial variability of microbial assemblages associated with a dominant habitat-forming seaweed. *Frontiers in microbiology,* **6**, 230. https://doi.10.3389/fmicb.2015.00230.

Chao A., Bunge J (2002) Estimating the number of species in a stochastic abundance model. *Biometrics,* **58**, 531-539. https://doi.10.1111/j.0006-341X.2002.00531.x.

Coyte K. Z., Schluter J, Foster KR (2015) The ecology of the microbiome: Networks, competition, and stability. *Science,* **350**, 663-666. https://doi.10.1126/science.aad2602.

Egan S., Harder T., Burke C., Steinberg P., Kjelleberg S., Thomas T. (2013) The seaweed holobiont: understanding seaweed–bacteria interactions. FEMS microbiology reviews, 37, 462-476. https://doi:10.1111/1574-6976.12011.

Freshwater D. W., Montgomery F, Greene JK, Hamner RM, Williams M, Whitfield PE (2006) Distribution and identification of an invasive *Gracilaria* species that is hampering commercial fishing operations in southeastern North Carolina, USA. *Biological Invasions,* **8**, 631-637. https://doi.10.1007/s10530-005-1809-5.

Gerard V., Dunham S, Rosenberg G (1990) Nitrogen-fixation by cyanobacteria associated with *Codium fragile* (Chlorophyta): environmental effects and transfer of fixed nitrogen. *Marine Biology,* **105**, 1-8.

Gohl D. M., Vangay P, Garbe J, MacLean A, Hauge A, Becker A, ... Hunter R (2016) Systematic improvement of amplicon marker gene methods for increased accuracy in microbiome studies. *Nature biotechnology,* **34**, 942. https://doi.10.1038/nbt.3601.

Gonzalez D. J., Smyth AR, Piehler MF, McGlathery KJ (2013) Mats of the nonnative macroalga, *Gracilaria vermiculophylla*, alter net denitrification rates and nutrient fluxes on intertidal mudflats. *Limnology and Oceanography,* **58**, 2101-2108. https://doi.10.4319/lo.2013.58.6.2101.

Gurgel C. F. D., Norris JN, Schmidt WE, Le HN, Fredericq S (2018) Systematics of the Gracilariales (Rhodophyta) including new subfamilies, tribes, subgenera, and two new genera, *Agarophyton* gen. nov. and *Crassa* gen. nov. *Phytotaxa,* **374**, 1-23. https://doi.10.11646/phytotaxa.374.1.1.

Hammann M., Buchholz B, Karez R, Weinberger F (2013) Direct and indirect effects of *Gracilaria vermiculophylla* on native *Fucus vesiculosus*. *Aquatic Invasions,* **8**, 121-132. https://doi.10.3391/ai.2013.8.2.01.

Hammann M., Rempt M, Pohnert G, Wang G, Boo SM, Weinberger F (2016a) Increased potential for wound activated production of Prostaglandin E2 and related toxic compounds in non-native populations of *Gracilaria vermiculophylla*. *Harmful algae,* **51**, 81-88. https://doi.10.1016/j.hal.2015.11.009.

Hammann M., Wang G, Boo SM, Aguilar-Rosas LE, Weinberger F (2016b) Selection of heat-shock resistance traits during the invasion of the seaweed *Gracilaria vermiculophylla*. *Marine Biology,* **163**, 104. https://doi.10.1007/s00227-016-2881-3.

Harder T., Campbell AH, Egan S, Steinberg PD (2012) Chemical mediation of ternary interactions between marine holobionts and their environment as exemplified by the red alga *Delisea pulchra*. *Journal of chemical ecology,* **38**, 442-450. https://doi.10.1007/s10886-012-0119-5.

Hendershot J. N., Read QD, Henning JA, Sanders NJ, Classen AT (2017) Consistently inconsistent drivers of microbial diversity and abundance at macroecological scales. *Ecology*, **98**, 1757-1763. https:/doi.org/10.1002/ecy.1829.

Hu Z., Liu R, Zhang J, Duan D, Wang G, Li W (2018) A unique genetic lineage at the southern coast of China in the agar-producing *Gracilaria vermiculophylla* (Gracilariales, Florideophyceae). *Algae,* **33**, 269-278. https://doi.10.4490/algae.2018.33.8.30.

Hu Z., Lopez-Bautista J (2014) Adaptation mechanisms and ecological consequences of seaweed invasions: a review case of agarophyte *Gracilaria vermiculophylla*. *Biological Invasions,* **16**, 967-976. https://doi.10.1007/s10530-013-0558-0.

Hughes J. S., Otto SP (1999) Ecology and the Evolution of Biphasic Life Cycles. The American Naturalist, 154, 306-320. https://doi:10.1086/303241.

Hurlbert S. H. (1971) The nonconcept of species diversity: a critique and alternative parameters. *Ecology,* **52**, 577-586. https://doi.10.2307/1934145.

Kanehisa M., Goto S, Sato Y, Kawashima M, Furumichi M, Tanabe M (2013) Data, information, knowledge and principle: back to metabolism in KEGG. *Nucleic acids research,* **42**, D199-D205. https://doi.10.1093/nar/gkt1076.

Kim S. Y., Weinberger F, Boo SM (2010) Genetic data hint at a common donor region for invasive Atlantic and Pacific populations of *Gracilaria vermiculophylla* (Gracilariales, Rhodophyta). *Journal of Phycology,* **46**, 1346-1349. https://doi.10.1111/j.1529-8817.2010.00905.x.

Klindworth A., Pruesse E, Schweer T, Peplies J, Quast C, Horn M, ... Glöckner FO (2013) Evaluation of general 16S ribosomal RNA gene PCR primers for classical and next-generation sequencing-based diversity studies. *Nucleic acids research,* **41**, e1-e1. https://doi.10.1093/nar/gks808.

Kollars N. M., Krueger-Hadfield SA, Byers JE, Greig TW, Strand AE, Weinberger F, ... Sotka EE (2015) Development and characterization of microsatellite loci for the haploid–diploid red seaweed *Gracilaria vermiculophylla*. *PeerJ,* **3**, e1159. https://doi./10.7717/peerj.1159.

Krueger-Hadfield S. A., Kollars NM, Byers JE, Greig TW, Hammann M, Murray DC, ... Weinberger F (2016) Invasion of novel habitats uncouples haplo-diplontic life cycles. *Molecular ecology,* **25**, 3801-3816. https://doi.10.1111/mec.13718.

Krueger-Hadfield S. A., Kollars NM, Strand AE, Byers JE, Shainker SJ, Terada R, ... Sotka EE (2017) Genetic identification of source and likely vector of a widespread marine invader. *Ecology and evolution,* **7**, 4432-4447. https://doi.10.1002/ece3.3001.

Krueger-Hadfield S. A., Stephens TA, Ryan WH, Heiser S (2018) Everywhere you look, everywhere you go, there's an estuary invaded by the red seaweed *Gracilaria vermiculophylla* (Ohmi) Papenfuss, 1967. BioInvasions Records, **7**, 343-355. https://doi.10.3391/bir.2018.7.4.01.

Lees L. E., Krueger-Hadfield SA, Clark AJ, Duermit EA, Sotka EE, Murren CJ (2018) Nonnative *Gracilaria vermiculophylla* tetrasporophytes are more difficult to debranch and are less nutritious than gametophytes. *Journal of Phycology*, **54**, 471-482. https://doi.10.1111/jpy.12746.

Lachnit T., Meske D, Wahl M, Harder T, Schmitz R (2011) Epibacterial community patterns on marine macroalgae are host-specific but temporally variable. *Environmental microbiology,* **13**, 655-665. https://doi.10.1111/j.1462-2920.2010.02371.x.

Lindström E. S., Langenheder S (2012) Local and regional factors influencing bacterial community assembly. *Environmental Microbiology Reports,* **4**, 1-9. https://doi.10.1111/j.1758-2229.2011.00257.x

Locey K. J., Lennon JT (2016) Scaling laws predict global microbial diversity. *Proceedings of the National Academy of Sciences of the United States of America,* **113**, 5970-5975. https://doi.10.1073/pnas.1521291113.

Nottingham A. T., Fierer N, Turner BL, Whitaker J, Ostle NJ, McNamara NP, ... Silman MR (2018) Microbes follow Humboldt: temperature drives plant and soil microbial diversity patterns from the Amazon to the Andes. *Ecology*, **99**, 2455-2466. https://doi.org/10.1002/ecy.2482.

Margulis L. (1990) Words as battle cries: symbiogenesis and the new field of endocytobiology. *Bioscience,* **40**, 673-677. https://doi.10.2307/1311435.

Martiny JB, Eisen JA, Penn K, Allison SD, Horner-Devine MC (2011) Drivers of bacterial beta-diversity depend on spatial scale. *Proceedings of the National Academy of Sciences of the United States of America*, 108, 7850-7854. https://doi.10.1073/pnas.1016308108

Marzinelli E. M., Campbell AH, Zozaya Valdes E, Vergés A, Nielsen S, Wernberg T, ... Thomas T (2015) Continental-scale variation in seaweed host-associated bacterial communities is a function of host condition, not geography. *Environmental microbiology,* **17**, 4078-4088. https://doi.10.1111/1462-2920.12972.

McGill B. J. (2010) Ecology. Matters of scale. *Science,* **328**, 575-576. https://doi.10.1126/science.1188528.

McGlinn D. J., Xiao X, May F, Gotelli NJ, Engel T, Blowes SA, ... McGill BJ (2018) Measurement of Biodiversity (MoB): A method to separate the scale-dependent effects of species abundance distribution, density, and aggregation on diversity change. *Methods in Ecology and Evolution,* **10**, 258-269. https://doi.10.1111/2041-210X.13102.

Messier J., McGill BJ, Lechowicz MJ (2010) How do traits vary across ecological scales? A case for trait-based ecology. *Ecology Letters,* **13**, 838-848. https://doi.10.1111/j.1461-0248.2010.01476.x.

Morrissey K. L., Çavaş L, Willems A, De Clerck O (2019) Disentangling the influence of environment, host specificity and thallus differentiation on bacterial communities in siphonous green seaweeds. *Frontiers in microbiology,* **10**, 717. https://doi.10.3389/fmicb.2019.00717.

Muangmai N., Vo TD, Kawaguchi S (2014) Seasonal fluctuation in a marine red alga, *Gracilaria vermiculophylla* (Gracilariales, Rhodophyta), from Nokonoshima Island, Southern Japan. *J Fac Agricult Kyushu Univ,* **59**, 243-248.

Nakagawa S., Schielzeth H (2013) A general and simple method for obtaining R2 from generalized linear mixed-effects models. *Methods in ecology and evolution,* **4**, 133-142. https://doi.10.1111/j.2041-210x.2012.00261.x.

Nyberg C. D., Thomsen MS, Wallentinus I (2009) Flora and fauna associated with the introduced red alga *Gracilaria vermiculophylla*. *European Journal of Phycology*, **44**, 395-403. https://doi.0.1080/09670260802592808.

Oksanen J., Blanchet FG, Kindt R, Legendre P, Minchin PR, O'hara R, ... Wagner H (2013) Package 'vegan'. Community ecology package, version 2.9.

Quast C., Pruesse E, Yilmaz P, Gerken J, Schweer T, Yarza P, ... Gloeckner FO (2013) The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic acids research,* **41**, D590-D596. https://doi.10.1093/nar/gks1219.

Ramus A. P., Silliman BR, Thomsen MS, Long ZT (2017) An invasive foundation species enhances multifunctionality in a coastal ecosystem. *Proceedings of the National Academy of Sciences of the United States of America,* **114**, 8580-8585. https://doi.10.1073/pnas.1700353114.

Raut Y., Morando M, Capone DG (2018) Diazotrophic macroalgal associations with living and decomposing *Sargassum*. *Frontiers in microbiology*, **9**, 3127. https://doi.10.3389/fmicb.2018.03127.

Ray J. L., Althammer J, Skaar KS, Simonelli P, Larsen A, Stoecker D, ... Nejstgaard JC (2016) Metabarcoding and metabolome analyses of copepod grazing reveal feeding preference and linkage to metabolite classes in dynamic microbial plankton communities. *Molecular ecology,* **25**, 5585-5602. https://doi.10.1111/mec.13844.

Rillig M. C., Antonovics J, Caruso T, Lehmann A, Powell JR, Veresoglou SD, ... Verbruggen E (2015) Interchange of entire communities: microbial community coalescence. *Trends in ecology & evolution,* **30**, 470-476. https://doi.10.1016/j.tree.2015.06.004.

Rodríguez-Echeverría S (2010) Rhizobial hitchhikers from Down Under: invasional meltdown in a plant–bacteria mutualism? *Journal of Biogeography*, **37**, 1611-1622.

Rodríguez-Echeverría S., Le Roux JJ, Crisóstomo JA, Ndlovu J (2011) Jack-of-all-trades and master of many? How does associated rhizobial diversity influence the colonization success of Australian *Acacia* species? *Diversity and Distributions,* **17**, 946-957. https://doi.10.1111/j.1472-4642.2011.00787.x.

Rosenberg G., Paerl H (1981) Nitrogen fixation by blue-green algae associated with the siphonous green seaweed *Codium decorticatum*: effects on ammonium uptake. *Marine Biology,* **61**, 151-158. https://doi.10.1007/BF00386654.

Roth-Schulze A. J., Pintado J, Zozaya-Valdés E, Cremades J, Ruiz P, Kjelleberg S, ... Thomas T (2018) Functional biogeography and host specificity of bacterial communities associated with the Marine Green Alga *Ulva* spp. *Molecular ecology,* **27**, 1952-1965. https://doi.doi.org/10.1111/mec.14529.

Rueness J. (2005) Life history and molecular sequences of *Gracilaria vermiculophylla* (Gracilariales, Rhodophyta), a new introduction to European waters. *Phycologia,* **44**, 120-128.

Saha M., Weinberger F (2019) Microbial “gardening” by a seaweed holobiont: Surface metabolites attract protective and deter pathogenic epibacterial settlement. *Journal of Ecology,* **107**, 2255-2265. https://doi.10.1111/1365-2745.13193.

Saha M., Wiese J, Weinberger F, Wahl M (2016) Rapid adaptation to controlling new microbial epibionts in the invaded range promotes invasiveness of an exotic seaweed. *Journal of Ecology,* **104**, 969-978. https://doi.10.1111/1365-2745.12590.

Schloss P. D., Westcott SL, Ryabin T, Hall JR, Hartmann M, Hollister EB, ... Weber CF (2009) Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities. *Applied and Environmental Microbiology,* **75**, 7537-7541. https://doi.10.1128/AEM.01541-09.

Shade A., Handelsman J (2011) Beyond the Venn diagramm: the hunt for a core microbiome. *Beyond the Venn diagram: the hunt for a core microbiomeemi_2585,* **14**, 4-12. https://doi.10.1111/j.1462-2920.2011.02585.x.

Skriptsova A., Titlyanova T, Titlyanov E (2001) Red algae of the genus *Gracilaria* in the south of the Russian Far East. *Russian Journal of Marine Biology,* **27**, S38-S52. https://doi.10.1023/A:1013898905437.

Smith D. P., Peay KG (2014) Sequence depth, not PCR replication, improves ecological inference from next generation DNA sequencing. *PloS one,* **9**, e90234. https://doi.10.1371/journal.pone.0090234.

Sotka E. E., Baumgardner AW, Bippus PM, Destombe C, Duermit EA, Endo H, ... Murren CJ (2018) Combining niche shift and population genetic analyses predicts rapid phenotypic evolution during invasion. *Evolutionary Applications,* **11**, 781-793. https://doi.10.1111/eva.12592.

Spalding M. D., Fox HE, Allen GR, Davidson N, Ferdana ZA, Finlayson M, ... Lourie SA (2007) Marine ecoregions of the world: a bioregionalization of coastal and shelf areas. *Bioscience,* **57**, 573-583. https://doi.10.1641/B570707.

Sunagawa S., Coelho LP, Chaffron S, Kultima JR, Labadie K, Salazar G, ... Bork P (2015) Ocean plankton. Structure and function of the global ocean microbiome. *Science,* **348**, 1261359. https://doi.10.1126/science.1261359.

Suzuki R., Shimodaira H (2006) Pvclust: an R package for assessing the uncertainty in hierarchical clustering. *Bioinformatics,* **22**, 1540-1542. https://doi.10.1093/bioinformatics/btl117.

Tedersoo L., Bahram M, Polme S, Koljalg U, Yorou NS, Wijesundera R, ... Abarenkov K (2014) Fungal biogeography. Global diversity and geography of soil fungi. *Science,* **346**, 1256688. https://doi.10.1126/science.1256688.

Thomsen M. S., Gurgel CFD, Fredericq S, McGlathery KJ (2006) *Gracilaria vermiculophylla* (Rhodophyta, Gracilariales) in Hog Island Bay, Virginia: A cryptic alien and invasive macroalga and taxonomic correction. *Journal of Phycology,* **42**, 139-141. https://doi.10.1111/j.1529-8817.2006.00160.x.

Vitousek P. M., Howarth RW (1991) Nitrogen limitation on land and in the sea: how can it occur? *Biogeochemistry,* **13**, 87-115.

Vitousek P. M., Walker LR, Whiteaker LD, Mueller-Dombois D, Matson PA (1987) Biological Invasion by *Myrica faya* alters ecosystem development in Hawaii. *Science,* **238**, 802-804. https://doi.10.1126/science.238.4828.802.

Wahl M., Goecke F, Labes A, Dobretsov S, Weinberger F (2012) The second skin: ecological role of epibiotic biofilms on marine organisms. *Frontiers in microbiology,* **3**, 292. https://doi.10.3389/fmicb.2012.00292.

Wang S., Wang G, Weinberger F, Bian D, Nakaoka M, Lenz M (2017b) Anti-epiphyte defences in the red seaweed *Gracilaria vermiculophylla*: non-native algae are better defended than their native conspecifics. *Journal of Ecology,* **105**, 445-457. https://doi.10.1111/1365-2745.12694.

Wang S., Weinberger F, Xiao L, Nakaoka M, Wang G, Krueger-Hadfield SA, ... Lenz M (2017a) *In situ* common garden assays demonstrate increased defense against natural fouling in non-native populations of the red seaweed *Gracilaria vermiculophylla*. *Marine Biology,* **164**, 193. https://doi.10.1007/s00227-017-3226-6.

Wang Y., Naumann U, Wright ST, Warton DI (2012) mvabund–an R package for model-based analysis of multivariate abundance data. *Methods in Ecology and Evolution,* **3**, 471-474. https://doi.10.1111/j.2041-210X.2012.00190.x.

Warton D. I., Wright ST, Wang Y (2012) Distance-based multivariate analyses confound location and dispersion effects. *Methods in Ecology and Evolution*, **3**, 89-101. https://doi.10.1111/j.2041-210X.2011.00127.x.

Warton D. I., Blanchet FG, O'Hara RB, Ovaskainen O, Taskinen S, Walker SC, ... Hui FK (2015) So many variables: joint modeling in community ecology. *Trends in Ecology & Evolution,* **30**, 766-779. https://doi.10.1016/j.tree.2015.09.007.

Weinberger F., Buchholz B, Karez R, Wahl M (2008) The invasive red alga *Gracilaria vermiculophylla* in the Baltic Sea: adaptation to brackish water may compensate for light limitation. *Aquatic Biology,* **3**, 251-264. https://doi.10.3354/ab00083.

Weinberger F., Friedlander M (2000) Response of *Gracilaria conferta* (Rhodophyta) to oligoagars results in defense against agar-degrading epiphytes. *Journal of Phycology,* **36**, 1079-1086. https://doi.10.1046/j.1529-8817.2000.00003.x.

Wittebolle L., Marzorati M, Clement L, Balloi A, Daffonchio D, Heylen K, ... Boon N (2009) Initial community evenness favours functionality under selective stress. *Nature,* **458**, 623. https://doi.10.1038/nature07840.

[dataset] Bonthond G., Bayer T., Krueger-Hadfield S. A., Barboza F. R., Nakaoka M., … Weinberger F.; 2019; PRJNA564581; Short Read Archive; https://www.ncbi.nlm.nih.gov/bioproject/PRJNA564581/

Data Accessibility

De-multiplexed V4-16S and V7-18S rRNA gene amplicon reads and associated metadata are deposited in the SRA database under the Bioproject accession number PRJNA564581.

Author Contributions

GB, TB and FW conceptualized the study. Field collections were conducted by GB, TB, SAKH, MN, MV, GW and FW. GB, SAKH and SK conducted laboratory work. GB processed and analyzed the data. FRB helped analyze and interpret the data. GB drafted the manuscript. TB, SAKH, FRB, MN, MV, GW, SK and FW helped writing the manuscript.

Figures

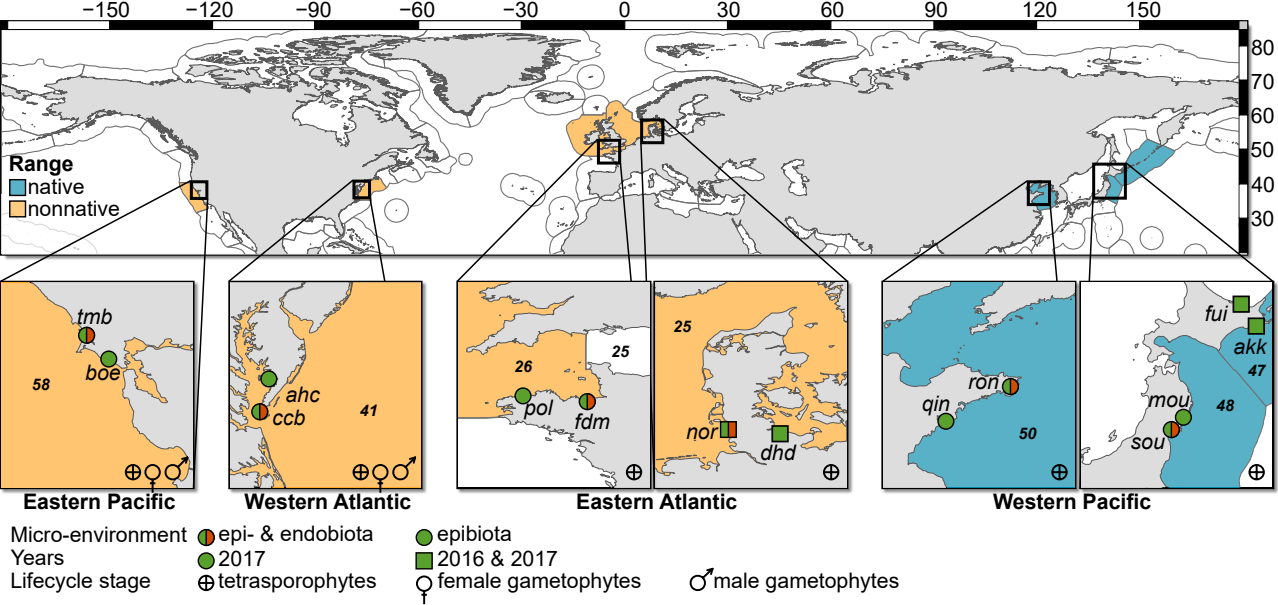
Figure 1. Geographic scales included in the sampling of this study: ranges (native and nonnative), coasts, ecoregions and populations. Sampled ecoregions are in blue (native) and orange (nonnative) and labelled with

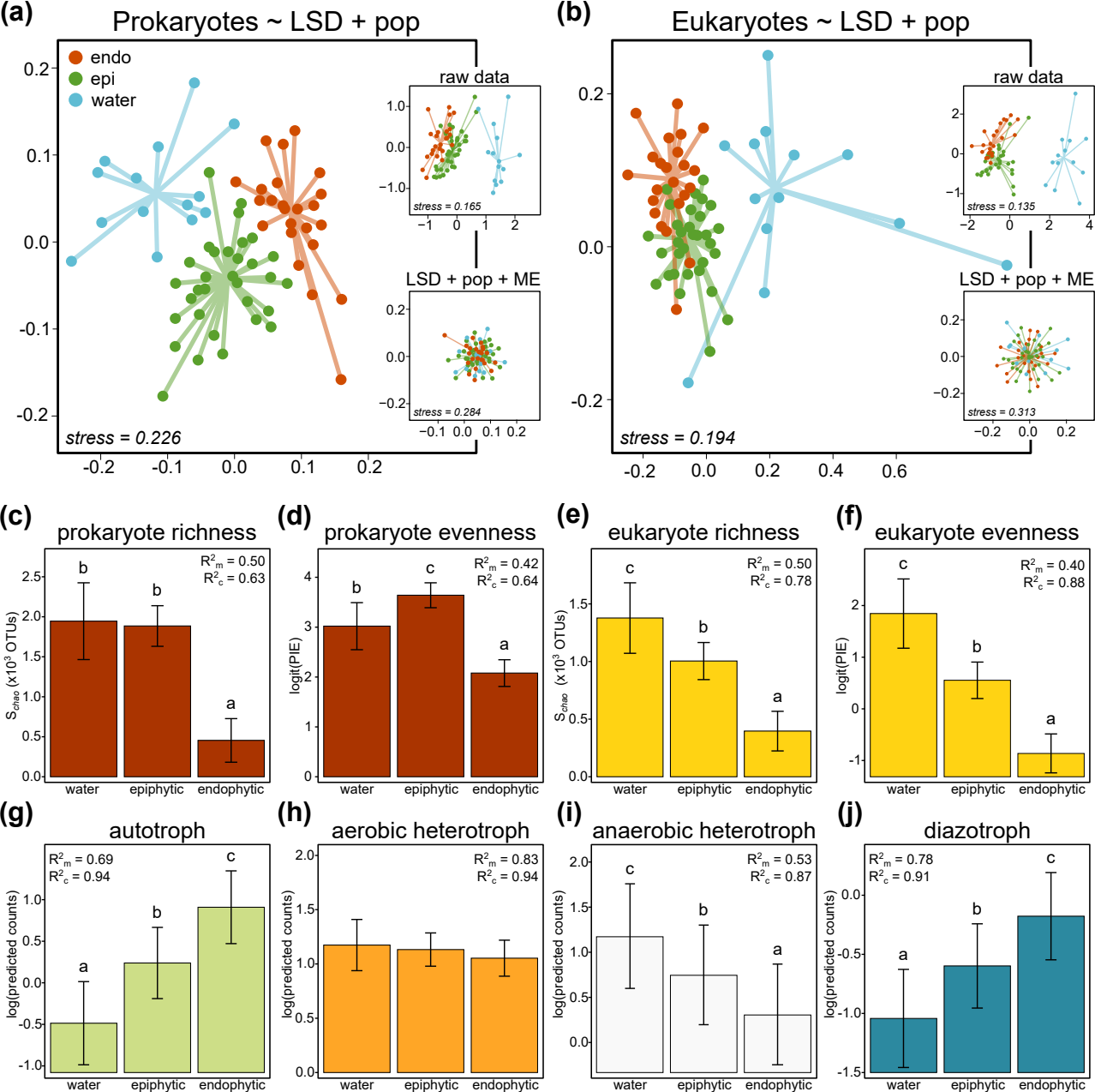
numbers corresponding to the biogeographic framework from Spalding et al. (2007). The insets are ordered by coast and show the sampling sites (labelled as in Table S1) inside the corresponding ecoregions. Sites from where only epibiotic samples were taken are labeled in green and those from which also endobiotic samples were acquired are green-red. Populations collected in both 2016 and 2017 are indicated by squares. Symbols in the lower right corner indicate which ploidies and sexes (⊕, ♀, ♂) were collected.

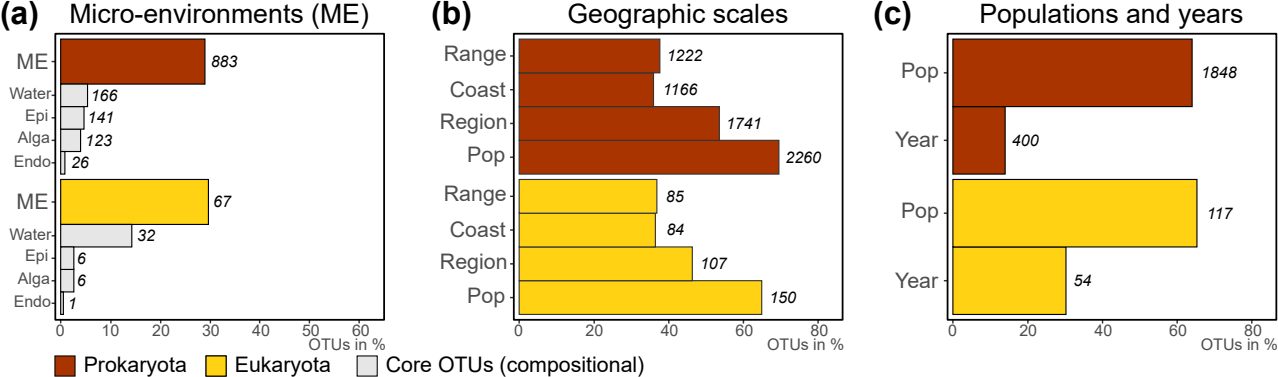
Figure 2. Differences among micro-environments. (a-b) Non-metric dimensional scaling (nMDS) plots and effects for water, epi- and endophytic microbial communities, adjusted for the effects of sequencing depth and population identity. Upper insets on the right display nMDS plots based on the raw data and lower right insets on the mGLM residuals also adjusted for effects of micro-environment (ME). (a) Prokaryotic communities. (b) Eukaryotic communities. Effects from univariate models are displayed with 95% confidence intervals for prokaryotic S_{chao} (c), prokaryotic logit PIE (d), eukaryotic asymptotic S_{chao} (e), eukaryotic logit PIE (f) and predicted abundance of functional groups (g-j). Effects are labeled with letters to indicate significant differences ($p < 0.05$). The marginal and conditional R^2 values (R^2_m and R^2_c) are displayed in the upper left or right corners.

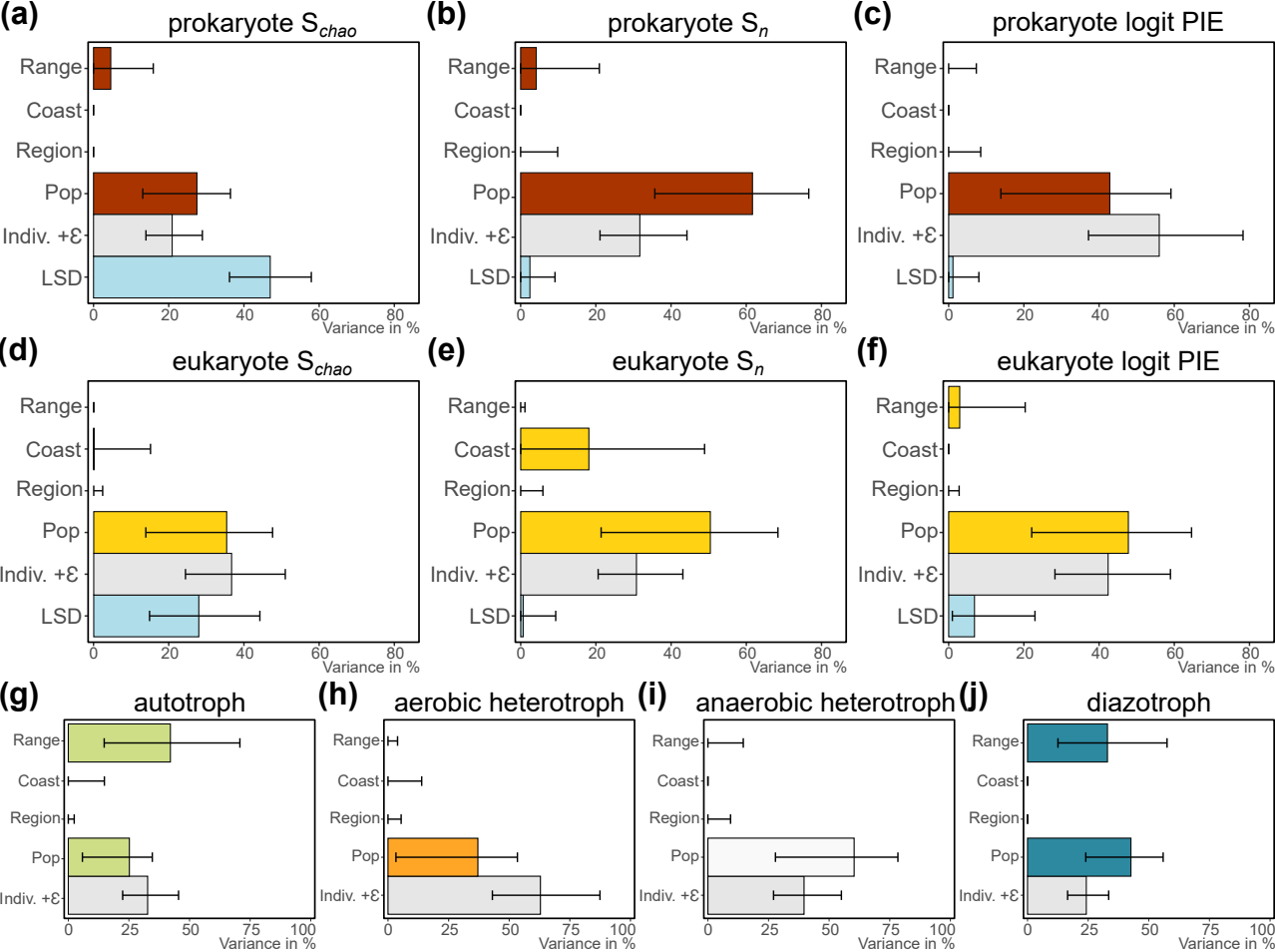
Figure 3. Differentially abundant OTUs from mGLMs (a) among micro-environments (tissue, surface and water), (b) within different geographic scales (range, coast, region and population) and (c) populations and years (population and year). Bars represent percentages of differentially abundant OTUs and numbers next to the bars to the corresponding OTU counts. Prokaryotic and eukaryotic results are shown in red and yellow, respectively. Grey bars in (a) indicate the amount of OTUs significantly associated with water, epiphytic, algal (epi and endo) and endophytic samples (composition cores, see main text for details). The results are based on a significance level of $p < 0.05$.

Figure 4. Variation in microbial communities across scales, visualized as percentages of explained variance. Variance associated to geographic scales, sequencing depth (blue) and residuals (grey) were extracted from univariate mixed models and the respective 95% confidence intervals were computed by bootstrapping with a 1,000 iterations. (a-c) Prokaryotic diversity parameters. (d-f) Eukaryotic diversity parameters. (g-j) Abundance of predicted functional groups (variation explained by the LSD is not shown).


−150
−120
−90
−60
−30
0
30
60
90
120
150
80
70
60
50
40
30
Range
native
nonnative
tmb
boe
58
Eastern Pacific
ahc
ccb
41
Western Atlantic
26
25
pol
fdm
25
nor
dhd
Eastern Atlantic
ron
qin
50
fui
akk
47
mou
48
sou
Western Pacific
Micro-environment
epi- & endobiota
epibiota
Years
2017
2016 & 2017
Lifecycle stage
tetrasporophytes
female gametophytes
male gametophytes


(a)
Prokaryotes ~ LSD + pop
endo
epi
water
raw data
stress = 0.165
LSD + pop + ME
stress = 0.284
stress = 0.226
(b)
Eukaryotes ~ LSD + pop
raw data
stress = 0.135
LSD + pop + ME
stress = 0.313
stress = 0.194
(c)
prokaryote richness
R^2_m = 0.50
R^2_c = 0.63
S_{chao} (x10^3 OTUs)
(d)
prokaryote evenness
R^2_m = 0.42
R^2_c = 0.64
logit(PIE)
(e)
eukaryote richness
R^2_m = 0.50
R^2_c = 0.78
S_{chao} (x10^3 OTUs)
(f)
eukaryote evenness
R^2_m = 0.40
R^2_c = 0.88
logit(PIE)
(g)
autotroph
R^2_m = 0.69
R^2_c = 0.94
log(predicted counts)
(h)
aerobic heterotroph
R^2_m = 0.83
R^2_c = 0.94
log(predicted counts)
(i)
anaerobic heterotroph
R^2_m = 0.53
R^2_c = 0.87
log(predicted counts)
(j)
diazotroph
R^2_m = 0.78
R^2_c = 0.91
log(predicted counts)
water
epiphytic
endophytic


(a)
Micro-environments (ME)
ME
883
Water
166
Epi
141
Alga
123
Endo
26
ME
67
Water
32
Epi
6
Alga
6
Endo
1
0
10
20
30
40
50
60
OTUs in %
(b)
Geographic scales
Range
1222
Coast
1166
Region
1741
Pop
2260
Range
85
Coast
84
Region
107
Pop
150
0
20
40
60
80
OTUs in %
(c)
Populations and years
Pop
1848
Year
400
Pop
117
Year
54
0
20
40
60
80
OTUs in %
Prokaryota
Eukaryota
Core OTUs (compositional)


(a)
prokaryote S_{chao}
(b)
prokaryote S_n
(c)
prokaryote logit PIE
(d)
eukaryote S_{chao}
(e)
eukaryote S_n
(f)
eukaryote logit PIE
(g)
autotroph
(h)
aerobic heterotroph
(i)
anaerobic heterotroph
(j)
diazotroph
Range
Coast
Region
Pop
Indiv. +ε
LSD
Variance in %