Journal of Eukaryotic Microbiology
June 2023, Volume 70, Issue 3, Pages e12966 (14p.)
https://doi.org/10.1111/jeu.12966
https://archimer.ifremer.fr/doc/00820/93216/

Archimer
https://archimer.ifremer.fr

High- resolution phylogenetic analysis reveals long- term microbial dynamics and microdiversity in phytoplankton microbiome

Choi Chang Jae [1, *], Jauzein Cecile [1, 2], Erdner Deana L. [1]

[1] The University of Texas at Austin Marine Science Institute Port Aransas TX, USA
[2] Ifremer, Dynamiques des Ecosystèmes Côtiers (DYNECO), Laboratoire d'Ecologie Pélagique (PELAGOS) Plouzané ,France

* Corresponding author : Chang Jae Choi, email address : changjae.choi@gmail.com

Abstract :

Phytoplankton-bacteria interactions represent the evolution of complex cross-kingdom networks requiring niche specialization of diverse microbes. Unraveling this co-evolutionary process has proven challenging because microbial partnerships are complex, and their assembly can be dynamic as well as scale- and taxon-dependent. Here we monitored long-term experimental evolution of phytoplankton-bacteria interactions by reintroducing the intact microbiome into an axenized dinoflagellate Alexandrium tamarense to better understand microbiome assembly dynamics and how microbiome composition could shift and stabilize over 15 months. We examined host functioning by growth rate, photosynthetic capability, cell size, and other physiological signatures and compared it to associated microbial communities determined by 16S rRNA gene sequences. Our results showed that microbiome reconstitution did not restore the intact microbiome, instead a distinct microbial community shift to Roseobacter clade was observed in the re-established cultures. In-depth comparisons of microbial interactions revealed no apparent coupling between host physiology and specific bacterial taxa, indicating that highly represented, abundant taxa might not be essential for host functioning. The emergence of highly divergent Roseobacter clade sequences suggests fine-scale microbial dynamics driven by microdiversity could be potentially linked to host functioning. Collectively, our results indicate that functionally comparable microbiomes can be assembled from markedly different, highly diverse bacterial taxa in changing environments.

Keywords : Alexandrium, Roseobacter, phytoplankton-bacteria interactions, 16S rRNA gene V4 amplicon sequencing

Please note that this is an author-produced PDF of an article accepted for publication following peer review. The definitive publisher-authenticated version is available on the publisher Web site.

Accepted Article

15507408, ja, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/jeu.12966 by IFREMER Centre Bretagne BLP, Wiley Online Library on [13/02/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

INTRODUCTION

MOST eukaryotic organisms, if not all, have associated microbial communities termed microbiomes which exert marked influences on host functioning and health (Hacquard et al., 2015; Bourne et al., 2016). The host, in turn, selects the microbiome, and their interconnected evolutionary histories have undoubtedly shaped ecological interactions and ecosystem functioning (Fraune and Bosch 2007; Zeng et al., 2017; Bourne et al., 2016). Microbiome composition can be highly dynamic and variable, comprising hundreds or even thousands of species across a diverse array of taxa from three domains of life, as well as innumerable viruses (Taylor et al., 2007; Paez-Espino et al., 2016). Such complexity poses a number of conceptual, methodological, and analytical challenges, with each microbe having the potential to interact with all others and with allochthonous community members that may have transient effects on host function (Walter and Ley 2011). The challenges are further amplified by inherent characteristics of microbial taxa, such as rapid evolution, horizontal gene transfer, and feedback loops between the microbe and its host (Koskella et al., 2017). The concept of a 'core' microbiome, which refers particularly to the microbial community that is systematically associated with a given host, and its coupling with host functioning has thus garnered increasing attention for making sense of microbiome dynamics (Turnbaugh et al., 2009).

Given that the microbiome plays a critical role in host functioning, selection should favor those associations that increase host fitness (Gould et al., 2018). However, evolutionary dynamics are reciprocal and cannot be considered from the host side alone; microbiomes also quickly respond to such selection. There is a further layer of complexity in this co-evolutionary process since it operates on profoundly different timescales, with hosts often having far longer generation times than species within their microbiome (Koskella et al., 2017). On the other hand, host-associated microbiome composition can be conserved over evolutionary time scales. Growing evidence indicates that the composition of a host-associated microbiome harbors a phylogenetic signal, a pattern referred to as "phylosymbiosis," that correlates host and microbial community composition (Brucker and Bordenstein 2013). Significant degrees of phylosymbiosis are prevalent across diverse systems (Fitzpatrick et al., 2018; Pollock et al., 2018). However, this pattern of more similar community composition in more closely related hosts can result from completely different underlying processes, such as ecological or habitat filtering, which do not imply that microbes coevolve with hosts (Mazel et al., 2018). In the context of ecological complexity within the microbiome, all of these compound the challenges for microbiome studies.

Phytoplankton represent a microhabitat for aquatic microbes. The association of phytoplankton with bacteria is ubiquitous in natural phytoplankton communities and mediated by diverse molecules and sensing mechanisms that function in signaling and nutrient acquisition (Seymour et al., 2017; Amin et al., 2015; Zhou et al., 2016). A prerequisite for such exchange networks is the spatial assemblage of interacting microbes inhabiting the "phycosphere," which is defined as a microscopic region rich in organic molecules that surrounds phytoplankton cells (Bell and Mitchell 1972). This chemically enriched zone represents the key interface where tight interactions between phytoplankton and other microbes are controlled by exuded chemicals (Seymour et al., 2017). The inference of 'archetypal phytoplankton-associated taxa' derived from consistent detection of particular microbial taxa from phytoplankton cultures and field samples supports the presence of intimate and selective associations between phytoplankton and bacteria (Buchan et al., 2014). In line with these observations, there is an emerging view that phytoplankton-bacteria interactions need to be understood within the framework of symbiosis (Cooper and Smith 2015). For example, *Roseobacters* have been shown to be potential

mutualistic partners of phytoplankton and key players in the transformation of phytoplankton-derived dissolved organic matter (DOM) and dimethylsulfoniopropionate (DMSP) (Luo and Moran 2014). Despite these findings, the influence of these bacteria on the physiology and ecology of phytoplankton and the evolutionary forces that shape the relationship are still not well understood. Furthermore, the effects of bacteria vary with the physiological state of the phytoplankton (Seyedsayamdost et al., 2011), indicating that it may depend on interspecific competition or community assembly processes in the bacterial community (Smriga et al., 2016). A higher phylogenetic diversity at broad taxonomic scales in aquatic realms compared to land, with much of the aquatic diversity yet to be uncovered (Sunagawa et al., 2015), suggests a great potential for complex cross-kingdom interactions yet to be explored in phytoplankton communities.

Given the complexity of natural samples and the difficulty in isolating live (phytoplankton) particles, we investigated these phytoplankton-bacteria partnerships in the context of a microbiome framework under defined laboratory conditions. We reintroduced the intact microbiome into an axenized phytoplankton strain and monitored microbial community changes alongside phytoplankton physiology to understand how the microbial communities residing with phytoplankton shaped and were shaped by host functioning. It will provide a reference for identifying microbial influences on host physiology at the systems biology level as well as place these findings in an ecologically relevant framework.

MATERIALS AND METHODS

Culture acquisition and maintenance

The *Alexandrium tamarense* strain, CCMP1493 (xenic) and its sub-clone CCMP1598 (axenic), were acquired from the National Center for Marine Algae and Microbiota (NCMA, East Boothbay, ME, USA). Both strains were maintained as 1 L batch cultures in f/2 medium (minus Si) (Guillard 1975) at 16°C under ~100 µmol photons m^{-2} s^{-1} with a 12:12h light:dark diurnal cycle. Culture transfers were carefully done under sterile conditions with an initial cell abundance of ~200 cells ml^{-1} to freshly prepared f/2 medium at monthly intervals. Axenicity was periodically assessed by epifluorescence microscopy with DAPI (4′, 6-diamidino-2-phenylindole) staining.

Experimental design

Overall experimental design is summarized in Fig. 1 and more details are described in Jauzein et al., 2015. Briefly, a culture of xenic CCMP1493 was lysed by vortexing with 0.5 mm silica-zirconium beads (BioSpec Products, Inc., Bartlesville, OK, USA). The resulting lysate was filtered through a 20 µm Nitex mesh to remove intact phytoplankton cells (cellular dimensions of *Alexandrium tamarense* are generally ranged in length from 28 to 35 µm and width from 35 to 44 µm) (John et al., 2014). The filtrate was centrifuged at 3,000 × *g* for 5 min to remove large cell fragments then filtered through a 5.0 µm Nitex mesh to remove small cell fragments as well. The resulting bacterial filtrate was added into the culture of axenic CCMP1598 to create "re-xenic" cultures.

Measurement of host physiological parameters

Multiple physiological parameters representing host physiology were assessed and the details are described in the previous study (Jauzein et al., 2015). Briefly, growth rate was calculated based on the changes in cell abundance and cell volume was estimated from the measurements of two

Accepted Article

15507408, ja, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/jeu.12966 by IFREMER Centre Bretagne BLP, Wiley Online Library on [13/02/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

perpendicular cell diameters using an ellipsoid model. Photosynthetic capacity was assessed using fluorescence-based maximum quantum yield of photosystem II (F_v/F_m). Chlorophyll *a* content was measured using methanol extraction method, and carbon uptake rate was calculated from ^{13}C-labeled bicarbonate uptake.

Sample collection

Phytoplankton-attached (attached) bacterial assemblages were collected by filtering 300 mL of culture using a 20 µm Nitex mesh. The free-living microbial communities were collected from the filtrate by filtering through a 0.2 µm pore size polyethersulfone membrane filters (Supor 200, Pall Gelman, East Hills, NY, USA). Filters with attached or free-living communities were placed in sterile cryovials, flash-frozen in liquid nitrogen and transferred to -80°C until further use. All samples were collected when cells were in mid-exponential phase of growth. For the xenic "parent" clone CCMP1493, sampling was conducted at the beginning of the experiment and six years later in order to assess long-term community dynamics in the absence of major manipulation. For the rexenic clones, sample collection was performed after 9 and 15 months of culture maintenance after reintroducing bacteria.

DNA extraction, 16S rRNA gene V4 amplicon PCR and sequencing

DNA was extracted using the DNeasy Blood and Tissue Kit (Qiagen, Valencia, CA, USA) according to the manufacturer's protocol, with the addition of a bead beating step after the addition of buffer ATL. Phytoplankton-attached extracts were lysed with a mixture of 0.1 and 0.5 mm silica-zirconium beads (BioSpec Products, Inc., Bartlesville, OK, USA) while free-living extracts had only 0.1 mm beads. DNA extracts were quantified with the Qubit dsDNA HS (High Sensitivity) Assay Kit (Invitrogen, Carlsbad, CA, USA) and diluted to 20 ng μl^{-1} with TE pH 8.0 as template for PCR. The V4 hypervariable region of 16S rRNA gene was amplified with universal primers 515F (5′-GTGCCAGCMGCCGCGGTAA-3′) and 806R (5′-GGACTACHVGGGTWTCTAAT-3′) at RTL Genomics (Research and Testing Laboratory, Lubbock, TX, USA). Paired-end (PE) library sequencing (PE 2×250 bp) was performed using an Illumina MiSeq platform (MiSeq Reagent Kit v3, Illumina, Inc., San Diego, CA, USA).

Sequences were demultiplexed and assigned to corresponding samples using CASAVA (Illumina, Inc., San Diego, CA, USA) and primer sequences were trimmed using Cutadapt v.1.13 (Martin 2011). Quality of reads was assessed using FastQC v.0.11.8 (Andrews 2010) and low-quality sequence ends were trimmed at a Phred quality (Q) threshold of 25 using a 10 bp sliding window in Sickle 1.33 (Joshi and Fass 2011). Paired-end reads were merged using USEARCH v.11.0.667 (Edgar 2010) when reads had a ≥ 50 bp overlap with maximum 5% mismatch and those with maximum error rate >0.001 or shorter than 200 bp were discarded. After the removal of single sequence reads (using USEARCH), exact sequence variants (or amplicon sequence variants, ASVs) were identified using UNOISE3 algorithm with the default values of parameters (minimum abundance of 8) and OTU table was generated using the otutab command. Taxonomy was assigned by comparing each ASV against the SILVA 16S rRNA gene database (release 128, 29 September 2016, https://www.arb-silva.de/documentation/release-128) using a classifier SINTAX algorithm implemented in USEARCH (Edgar 2016). The ASV abundance distribution for each sample was assessed by generating octave plots on log2 scale as described by Edgar and Flyvbjerg 2015 using otutab_octave command, and low-abundance ASVs (ASVs with abundances less than 128) were discarded based on the histogram cutoffs (Supplementary Figure S1). Rarefaction curves were generated for each sample using the rarefy function from the R

Accepted Article

15507408, ja, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/jeu.12966 by IFREMER Centre Bretagne BLP, Wiley Online Library on [13/02/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

package vegan (Supplementary Figure S2). In addition, alpha diversity was checked against each sample read numbers to confirm that sequence depth was not correlated with alpha diversity (Supplementary Figure S3). The number of reads per sample ranged from 10,118 to 64,676 (36,651 ± 12,615), resulting in between 9,455 to 60,308 (34,176 ± 11,825) reads per sample after QC, and 9,330 to 58,544 (33,215 ± 11,347) reads were mapped back to ASV (Supplementary Data S1). Sequences were deposited in the SRA under SRR12228759-SRR12228781 (BioProject PRJNA646482).

Roseobacter reference alignment and tree construction

For high-resolution *Roseobacter* diversity analysis, 78 near full-length 16S rRNA gene sequences (70 described species and eight environmental sequences from undescribed taxa) representing the *Roseobacter* clade were retrieved from the most recent SILVA database (release 132, 13 December 2017, https://www.arb-silva.de/documentation/release-132) and through BLASTN searches against the GenBank nr database with a particular focus on sequences from cultured representatives. These sequences (and one outgroup sequence) were aligned using MAFFT v7.271 (Katoh and Standley 2013) with default parameters, and gaps were masked using trimAl v1.4 (Capella-Gutiérrez et al., 2009). Phylogenetic inferences were made by Maximum Likelihood methods implemented in RAxML v8.0.0 (Stamatakis 2014) under gamma corrected GTR model of evolution with 1,000 bootstrap replicates based on 1161 homologous positions as well as with PhyML v3.0.1 (Guindon et al., 2010) with the same substitution model and 100 bootstrap replicates. Additional phylogenetic reconstructions were performed in MrBayes v3.2.6 (Ronquist et al., 2012) with the parameters of lset nst=6 rates=invgamma ncat=6, and ngenval=10,000,000 samplefreqval=1,000 and tempval=0.200, and the final tree was produced with assistance from FigTree v1.4.3 (Rambaut 2007) (http://tree.bio.ed.ac.uk/software/figtree) and topology from RAxML (Fig. 5a).

Fine-scale taxonomic assignment of *Roseobacter* amplicons

Amplicon sequence variants (ASVs) representing *Rhodobacteraceae* from this study and other studies (Hasegawa et al., 2007; Jasti et al., 2005; Hold et al., 2001) were further analyzed using phylogenetic methods implemented in PhyloAssigner v.6.166 (Vergin et al., 2013), which performs a profile alignment of amplicons to a multiple sequence alignment using HMMER and assigns phylogenetic positions in an unmasked reference tree based on maximum likelihood methods using pplacer (Matsen et al., 2010). *Rhodobacteraceae* amplicons were classified with the *Roseobacter* reference alignment and tree developed in this study. In order to ascertain the accuracy of placement of amplicon sequences on the unmasked alignment by PhyloAssigner, we performed quality control with a set of known sequences. Specifically, 10 sets of V4 sequences were generated from 20 near-full length 16S rRNA gene sequences and used as a test set to verify correct classification by PhyloAssigner with the new reference alignment and tree. The placement results were 100% correct.

Hierarchical clustering

Hierarchical cluster analyses were carried out with log-transform normalized relative abundance of amplicons. The approximately unbiased *p*-values (%) as well as bootstrap probabilities were computed via multiscale bootstrap resampling with 10,000 replications using the R package Pvclust (Suzuki and Shimodaira 2006), modified to allow Bray-Curtis dissimilarity for distance calculations.

Accepted Article

15507408, ja, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/jeu.12966 by IFREMER Centre Bretagne BLP, Wiley Online Library on [13/02/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Accepted Article

15507408, ja, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/jeu.12966 by IFREMER Centre Bretagne BLP, Wiley Online Library on [13/02/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Microbial diversity and indicator species analyses

The non-parametric species richness estimator Chao1 and the Shannon diversity index (H′) were calculated for the bacterial communities using the R package phyloseq (McMurdie and Holmes 2013). Principal coordinate analysis (PCoA) based on Bray-Curtis dissimilarity was performed using ASV relative abundance. Principal component analysis (PCA) of phytoplankton physiological characteristics was conducted using the R package vegan (Oksanen et al., 2007). The most characteristic taxa based on their habitat preference (attached or free-living) and impact on host functioning (high function, transition, or low function) were identified through indicator values, implemented as 'indval' function in labdsv package in R (Roberts 2016).

Data availability and deposition

Microbiome sequence data are deposited in the Sequence Read Archive (SRA) under SRR12228759-SRR12228781. A complete PhyloAssigner package used for *Roseobacter* clade amplicon analysis including *Roseobacter* clade reference alignment and tree are deposited in Figshare (https://doi.org/10.6084/m9.figshare.12652988.v1).

RESULTS

A unique system for the study of dinoflagellate-microbe interactions

The system that formed the basis of our results came from the microbiome comparison of genetically identical, monoclonal strains of the dinoflagellate, *Alexandrium tamarense* CCMP1493 and CCMP1598. CCMP1493 was isolated from the China Sea and deposited at the NCMA in July 1992. CCMP1493 was rendered axenic at the NCMA, and the resulting axenic strain has been maintained since April 1993 as strain CCMP1598. CCMP1598 (hereafter "axenic") is a sub-culture of the original CCMP1493 (hereafter "xenic"), thus they are genetically identical and differ only in the presence or absence of their associated microbes. The morphology and physiology of these two strains are different; the axenic strain is smaller than the xenic strain, has lower chlorophyll and carbon content, and decreased photosynthetic efficiency (Table 1, $p < 0.0001$, two-tailed Mann-Whitney *U* tests). To determine whether the observed differences were caused by the associated microbial communities, microbiome reconstitution experiments were conducted by isolating the microbial community from the xenic "parent" strain, inoculating it into the axenic strain to create three "re-xenic" strains (Fig. 1). Our previous study examining the recovery of host physiological measures revealed that the three rexenic strains followed different trajectories. One failed to survive (Fig. 1, rexenic 3), and the remaining two rexenic strains showed different host functioning and health between the two replicate microbiome reconstitutions (Table 1, rexenic 1 and rexenic 2) (Jauzein et al., 2015).

Linking phytoplankton physiology to microbial dynamics

To gain insights into microbial interactions and their influences on host physiological phenotype, taxonomic characterization of both free-living and phytoplankton-attached (attached) microbial communities was performed using 16S rRNA gene V4 amplicon sequencing (Supplementary Data S1). The xenic "parent" strain was analyzed six years apart (time 1 and time 2) and the two replicate rexenic strains were examined at 9 and 15 months (time 1 and time 2) after microbiome re-addition. We first performed hierarchical clustering on the host functional signatures, such as growth rate, photosynthetic capability, carbon flux, and cell size (Table 1) (Jauzein et al., 2015) to characterize the changes in host physiology, for further comparison to associated microbial

communities. This resulted in two statistically supported groupings, a physiologically competent host group (Fig. 2a, smiling face) comprised of the xenic strain at time 1 and time 2, and the rexenic strain 2 at time 2, and a less competent host group (Fig. 2a, neutral face) composed of the rexenic strain 1 at time 2 and the rexenic strain 2 at time 1, here labeled 'high function' and 'transition', respectively. These terms reflect the host physiological state, as indistinguishable from the xenic parent (high function), or intermediate between the axenic and xenic parental states (transition). The axenic parent strain and the rexenic strain 1 at time 1 were separated from these groupings and did not form a cluster but both showed physiological indices of stress (Table 1) (Jauzein et al., 2015) and were thus labeled 'low function' (Fig. 2a, frowning face). A statistical ordination of physiological parameters clustered similarly to the hierarchical clustering (Supplementary Fig. S4). Next we examined the associated microbial communities to understand whether the changes in host physiology were derived from changes in microbial communities.

In hierarchical clustering of ASV frequencies, a major division was observed in the microbiomes of the xenic and rexenic strains, and there was no apparent coupling of host physiology to the associated microbial communities (Fig. 2b). Microbiomes from the xenic strain at time 1 and time 2 and the rexenic strain 2 at time 2 were markedly different from each other even though they were all associated with highly functioning hosts. Similarly, microbiomes derived from 'transition' states of host physiology (rexenic strain 1 time 2 and rexenic strain 2 time 1) were not clustered together, instead they grouped with other rexenic strains from different host physiological conditions. These ASV profiles indicated that distinct microbiomes were observed primarily between the xenic and rexenic strains, and that communities from rexenic strains further clustered based on temporal changes from time 1 to time 2. While there is a growing literature on the importance of attached over free-living communities for phytoplankton functioning and dynamics (Rooney-Varga et al., 2005; Grossart et al., 2005), analyses based on attached- (Supplementary Fig. S5a) and free-living-specific (Supplementary Fig. S5b) ASVs gave the same clustering results, indicating that the observed decoupling between host physiology and the associated microbiome was not solely driven by attached or free-living lifestyle.

To further investigate the effect of microbiome reconstitution on microbial diversity, we next explored microbiome alpha and beta diversity. Principal coordinates analysis (PCoA) and the Shannon diversity index (H′) combined with multivariate analysis of variance revealed differences between xenic and rexenic strains including attached-, free-living-, and time-dependent effects on microbial composition (Fig. 3a). Overall, microbiomes from the xenic and rexenic strains were distinctly different, and a clear separation between attached and free-living communities was observed as well. The xenic strain showed microbiome changes between time 1 and time 2 (a six-year period) in a similar fashion in both attached and free-living communities. A parallel time-dependent microbial shift was observed in free-living communities in both rexenic 1 and rexenic 2 strains (a six-month period), while attached communities showed some difference at time 1 then completely converged at time 2 (Fig. 3a). Shannon diversity was higher in the xenic strain than in the rexenic strains (Fig. 3b, $p < 0.005$, Kruskal-Wallis with Dunn's post-hoc test). This was mostly driven by free-living communities, which showed much higher diversity in the xenic strain than the rexenic strains (Fig. 3c, $p < 0.0001$, Kruskal-Wallis with Dunn's post-hoc test). The xenic strain also showed increased microbiome diversity over time (Fig. 3d), in both attached and free-living communities (Supplementary Fig. S6). In contrast, diversity decreased over time in rexenic strain 1 and showed no distinct change in rexenic strain 2 (Fig. 3d). In fact, both rexenic strains showed decreased diversity in free-living

Accepted Article

15507408, ja, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/jeu.12966 by IFREMER Centre Bretagne BLP, Wiley Online Library on [13/02/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

communities while attached communities in rexenic strain 2 showed increased diversity and rexenic strain 1 showed no distinct change (Supplementary Fig. S6). When compared to host physiology, microbiomes from hosts in the 'high function' condition exhibited higher α-diversity than those from hosts in 'transition' and 'low function' states (Fig. 3e, $p < 0.0001$, Kruskal-Wallis with Dunn's post-hoc test).

Detailed microbial community structure from microbiome reconstitution

Overall, a total of 73 ASVs were identified across all samples. We found many taxa from the source (xenic) microbiome were lost in the reconstituted communities during fifteen months of cultivation; 24 ASVs (33% of the total 73 ASVs) from the xenic strain were not observed in the rexenic strain sequence sets. Many of these (15 ASVs) were minor (relative abundance less than 1% of all reads), and therefore could easily become undetectable or fail to establish a population within the reconstituted communities. However, three ASVs contributing >10% of the total amplicons were also not recovered, indicating that even highly represented taxa could be lost or not established to a detectable level during the reassembly process, which is not easily identifiable through shorter-term studies. Conversely, 17 ASVs that were not observed from the source microbiome emerged later in the rexenic communities. Our results clearly showed taxon loss and emergence in the reconstructed microbiomes, as well as the presence of long-term (> 1 yr) assembly dynamics which cannot be resolved in short-term studies.

Patterns in relative amplicon abundance are strongly influenced by the reciprocal interplay between a taxon's own abundance and changing abundances of other taxa. To gain insights into key microbial taxa contributing to the observed diversity patterns, we examined taxonomic dynamics within attached and free-living communities over time. Based on ASV frequencies, five distinct clusters were identified: 1. ASVs with high relative abundance in both xenic and rexenic strains ("xenic and rexenic dominant"); 2. ASVs found primarily in rexenic strains at high relative abundance ("rexenic dominant"); 3. ASVs found primarily in rexenic strains at low relative abundance ("rexenic minor"); 4. ASVs found primarily in the xenic strain at high relative abundance ("xenic dominant"); and 5. ASVs found primarily in the xenic strain at low relative abundance ("xenic minor") (Fig. 4).

The xenic and rexenic dominant cluster was comprised of the five most abundant ASVs. They were highly represented in most of the xenic and rexenic strains, for example ASV1, which accounted for >90% of the total reads in some samples. ASV1, identified as *Roseobacter* sp., was the dominant taxon in rexenic strains. It was an indicator species for rexenic free-living communities, accounting for an average of 77 ± 17% of the total reads in those samples. This *Roseobacter* sp. was abundant in rexenic attached communities as well, comprising an average of 24 ± 13% of the total reads. Its relative abundance was lower in the xenic strain, where it accounted for an average of 7 ± 5% of reads across all xenic samples. ASV2, identified as *Formosa* sp. from *Bacteroidetes* at 99.6% nt identity, was the dominant taxon in the xenic strain, particularly in attached communities. It comprised 54% of the total reads in these samples at time 1, declining markedly to 13% of total reads at time 2. The ASV2 *Formosa* sp. was also abundant in rexenic attached communities at time 1 (33% and 52% of reads in rexenic strain 1 and 2, respectively) but barely detected at time 2 (<1% of reads in both rexenic strains). A similar pattern of decreasing relative abundance over time was observed for ASV5, which was identified as an uncultured *Gammaproteobacteria* and had lower average abundance than ASV1 and ASV2. Both ASV3 (*Marinobacter* sp., 100% nt identity) and ASV4 (*Aliiroseovarius* sp., 100% nt identity) had higher relative abundance in attached communities and generally increased over

Accepted Article

15507408, ja, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/jeu.12966 by IFREMER Centre Bretagne BLP, Wiley Online Library on [13/02/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

time. The abundance of these two taxa was highest in xenic attached communities at time 2 (ASV3; 26%, ASV4; 22% of total reads in those samples). These two ASVs were also abundant in rexenic 1 (ASV3; 39%, ASV4; 9% of total reads) and rexenic 2 (ASV3; 11%, ASV4; 15% of total reads) attached communities at time 2.

The rexenic dominant cluster comprised 15 ASVs, one-third of which (5 ASVs) were not detected from the xenic strain and exclusively recovered from the rexenic strains. Many ASVs showed increased abundance in attached communities at time 2. This agreed well with the result from the indicator species analysis since most indicator taxa were from attached communities at time 2. They were mostly identified as belonging to the *Roseobacter* clade (11 ASVs) of the *Alphaproteobacteria*. Other members included *Marinobacter* sp. (98% nt identity) from *Gammaproteobacteria*, *Imperialibacter* sp. (97% nt identity) from *Bacteroidetes*, and two other *Alphaproteobacteria* (*Pyruvatibacter* sp. and *Henriciella* sp.). Except for these two latter taxa, all others showed increased abundance over time.

The rexenic minor cluster included 21 ASVs; most accounted for <0.1% to the total reads across all samples. More than half of the ASVs (12 ASVs) were detected only in the rexenic strains. No particular lifestyle was evident as most ASVs were commonly detected in both attached and free-living communities. However, indicator species analysis identified 6 of the ASVs as indicator taxa for free-living communities, indicating a greater contribution from the free-living fraction. All members of this cluster were identified as *Alphaproteobacteria*, mostly *Roseobacter* clade (15 ASVs) followed by *Thalassospira* spp. (3 ASVs), and they generally showed increased abundance over time.

The xenic dominant cluster contained 14 ASVs, of which 10 ASVs (71%) were not detected from the rexenic strains. The other four ASVs were found in both xenic and rexenic strains and comprised <1% of the total reads, with declines in their relative abundance over time. No ASVs were identified as being specific to attached or free-living communities. Changes in the relative abundance varied but declines over time in free-living communities were more frequently observed, including three ASVs collectively accounting for >10% of the total reads at time 1 whose abundance was lower at time 2. Members were mostly *Alphaproteobacteria*, followed by *Gammaproteobacteria* and *Bacteroidetes*.

The xenic minor cluster comprised 18 ASVs contributing <1% of the total reads, of which 13 ASVs (72%) were not observed from the rexenic strains. This cluster exhibited a higher taxonomic diversity including phyla that were not recovered from the rexenic samples, such as *Betaproteobacteria*, *Planctomycetes*, and *Archaea*. ASVs from these clusters were mostly associated with free-living communities, which was confirmed by the indicator species analysis as well.

High sequence diversity and evolution within *Roseobacter* clade

Of the 73 total ASVs identified across all samples, 40 ASVs (55% of the total 73 ASVs) of them belonged to the *Roseobacter* clade, but many exhibited high sequence divergence from sequences in the GenBank nr database including other reference databases. To accurately determine the placement of these divergent sequences in the *Roseobacter* phylogeny, we developed a high-resolution reference phylogenetic tree based on near full-length *Roseobacter* 16S rRNA genes with particular emphasis on taxa that were captured from our amplicon dataset. The reference tree included 24 sub-clades represented at the genus level that were delineated and recovered with statistical support, containing 78 16S rRNA gene sequences from 70 described species and eight environmental sequences from undescribed taxa. A total of 40 ASVs classified

Accepted Article

15507408, ja, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/jeu.12966 by IFREMER Centre Bretagne BLP, Wiley Online Library on [13/02/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

as *Rhodobacteraceae* were placed in the reference tree using the PhyloAssigner pipeline (Vergin et al., 2013), and their phylogenetic positions were resolved at the genus level for most ASVs (Fig. 5a).

Different levels of sequence diversity could be observed from the *Roseobacter* clade members. A minority of ASVs (11 of 40 ASVs) showed 100% nt identity to cultured representatives of *Antarctobacter* sp., *Mameliella* sp., *Pseudooceanicola* sp., *Seohaeicola* sp., *Marivita* sp., *Roseovarius* sp., *Maritimibacter* sp., *Aliiroseovarius* sp., *Albimonas* sp., and *Labrenzia* sp. Twelve ASVs were assigned to *Roseobacter* spp. with a high sequence divergence, ranging from 94.1 to 99.6% nt identity, to the known cultured representatives. The presence of highly divergent sequences related to other taxa, such as *Ruegeria* sp. (94.1%, ASV76), *Phaeobacter* sp. (96.8%, ASV47), and *Aliiroseovarius* sp. (97.2%, ASV33 and 97.6%, ASV44), indicates high sequence diversity within the *Roseobacter* clade in these dinoflagellate microbiomes.

The observed sequence diversity raised the question of whether it is linked to host functioning, thus we used hierarchical clustering based solely on the relative abundance of *Roseobacter* clade ASVs. Similar to what was observed from overall microbial community composition, a distinct difference in *Roseobacter* clade members was observed between the xenic and rexenic strains but revealed no coupling between host physiology and the associated *Roseobacter* communities. Instead, groupings were related by members that were more pronounced in the rexenic strains (Group A and C) or from the xenic strain (Group D) or both (Group B) (Fig. 5b). The xenic and rexenic strains harbored taxonomically diverse members within *Roseobacter* clade. Group D taxa were mostly observed from the xenic strain and exhibited less sequence divergence from the cultured representatives (average 99.5 ± 1.0% nt identity) compared to ones mostly found from the rexenic strain (p <0.001, Kruskal-Wallis with Dunn's post-hoc test, Group A, 96.5 ± 1.8; Group C, 97.7 ± 0.3% nt identity), showing evidence of increased microdiversity within the re-established rexenic strains. Group B taxa showed less sequence divergence from the cultured representatives (average 99.1 ± 1.1% nt identity) except for the two unknown ASV14 (97.2%) and ASV16 (97.6%) that could represent novel taxa. We also observed a biphasic 'swim-or-stick' dual lifestyle from this group, particularly from the rexenic strains, with a clear distinction between free-living and attached communities. Seven ASVs, represented by three *Maritimibacter* spp. (ASV11, 18, and 19), one *Aliiroseovarius* sp. (ASV4), one *Phaeobacter* sp. (ASV22), and two unknown taxa (ASV14 and 16) dominated attached communities while one ASV identified as *Roseobacter* sp. (ASV1, 99.6% nt identity to the cultured representative) dominated the free-living communities, contributing up to 94% of the total *Rhodobacteraceae* reads.

DISCUSSION

It is broadly recognized that community assembly is simultaneously influenced by stochastic and deterministic factors that vary under different environmental conditions (Stegen et al., 2012). Multiple surveys of microbiome composition over a wide range of environmental settings indicate that microbial communities may have emergent properties for their assembly and dynamics (Goldford et al., 2018; Louca et al., 2018). Thus structuring principles must exist that determine the various community configurations; however, these are presently insufficiently described. Factors shaping microbial communities and understanding the mechanisms underlying the formation and function of host-associated microbial community structures is particularly relevant, as the associated microbes contribute to the host functioning and physiology (David et

Accepted Article

15507408, ja, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/jeu.12966 by IFREMER Centre Bretagne BLP, Wiley Online Library on [13/02/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

al., 2014). Here we examine changes in the host-microbiome symbiotic relationship in an attempt to explore the bidirectional relationship in which alterations in the microbiome affect host physiology and changes in host physiology may influence the microbiome. The host-microbe system used here, with tractable bacterial communities and axenized phytoplankton as a model system, could help to define the roles of microbial dynamics in a community context. Importantly, it can provide the basis to functionally validate mechanisms and causal relationships hypothesized from environmental observations or correlations.

Phytoplankton microbiome in a case study of *Alexandrium*

Links between phytoplankton and bacterial community composition and dynamics have been shown by multiple studies (Jasti et al., 2005; Sörenson et al., 2019; Lawson et al., 2018; Grossart et al., 2005; Rooney-Varga et al., 2005). The overall microbiome composition from this study is consistent with what has been reported from cultures of *Alexandrium* species, including bacterial groups such as *Rhodobacteraceae*, *Alteromonadaceae*, and *Flavobacteriaceae* (Hasegawa et al., 2007; Hold et al., 2001; Jasti et al., 2005; Sörenson et al., 2019). *Rhodobacteraceae*, one of the major marine groups belonging to the *Alphaproteobacteria*, could be found in both free-living and attached communities and exhibits a broad metabolic potential for nutrient uptake (Buchan et al., 2005). *Alteromonadaceae* belonging to the *Gammaproteobacteria* are frequently observed in bloom-associated bacteria communities (Teeling et al., 2012). *Flavobacteriia* belonging to the *Bacteroidetes* are typically one of the most dominant bacteria in blooms, often responding to phytoplankton decay, and are commonly found in close association with phytoplankton cells in cultures (Kirchman 2002).

The microbiome composition in phytoplankton cultures can shift over time under cultivation conditions (Sapp et al., 2007) although the effect of long-term cultivation on associated bacterial communities has shown to be consistent from diverse phytoplankton isolates (Behringer et al., 2018). However, multiple questions still remain to be answered to better understand microbiome assembly and dynamics, such as the importance of indigenous microbial communities for host functioning and health, the molecular basis by which host-microbe and microbe-microbe interactions shape and maintain microbial communities, and the role of individual microbes as well as their collective ecosystem function. Our microbiome reconstitution study showed distinct community reassembly processes when the rexenic strains were established, with the emergence of rexenic-specific taxa that were not recovered over six years from the xenic strain (Fig. 4). Furthermore, the results of our high-resolution phylogenetic analyses revealed that strong shifts in the associated communities linked with significant microdiversity can nonetheless promote host functioning and health comparable to the indigenous microbial communities (Fig. 5). Taken together our results show that while shifts in the microbiome can be expected as a reflection of changing environmental conditions, the newly established microbiome can maintain host physiology.

Long-term steady-state community assembly process

One of the most striking aspects of microbiome communities is their long-term stability in a host, often termed the 'core' microbiome (Faith et al., 2013). However, many uncertainties remain about what determines the stability of these communities and furthermore, how much time it takes for the microbiome to become stabilized in response to perturbation. Our long-term study provides novel insights into the dynamics of the community assembly process, as its lengthy period of study contrasts to most other similar studies that were monitored over shorter time

Accepted Article

15507408, ja, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/jeu.12966 by IFREMER Centre Bretagne BLP, Wiley Online Library on [13/02/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Accepted Article

scales (Mönnich et al., 2020).

We found a significant number of taxa were lost in the reconstituted communities during fifteen months of cultivation. Many of them represented <1% of the total sequences on average, and therefore might not be able to establish a population within the reconstituted communities. However, several taxa contributing >10% of the total sequences were also not recovered, indicating that even highly represented taxa could be lost during the microbiome reassembly process, which is not easily identifiable through shorter-term studies. Conversely, some taxa that were not present in the source microbiome emerged later in the reconstituted communities. Our results clearly showed many taxa could be lost or develop in the reconstructed microbiomes over the long-term (> 1 yr), which cannot be resolved in short-term studies.

The remaining ASVs were found throughout the xenic and the rexenic strains, and five of them (xenic and rexenic dominant cluster, Fig. 4) were abundant throughout all samples, thus representing the core microbiome. Among them, the flavobacterial genus *Formosa* sp. was one of the major indigenous attached taxa, though its abundance in the xenic strain declined during six years of cultivation. It was well represented in the rexenic strains until nine months but barely detected at 15 months. *Formosa* sp. is known to have expanded metabolic capacity and becomes competitive among other microbes during phytoplankton blooms (Unfried et al., 2018). Our results suggest that this taxon may be less competitive under steady-state conditions and have no distinct influence on host physiology. *Marinobacter* sp. is another core microbiome taxon mostly observed from the attached communities. It has been shown to enhance iron availability to associated phytoplankton via siderophore production (Amin et al., 2009), thus has potential to directly affect host physiology. However, again our results do not show a correlation between the dynamics of this taxon and host functioning. Similarly, we do not see correlations between host physiological changes and the remaining core taxa dynamics, including two ASVs which belong to the *Roseobacter* clade. *Roseobacter* clade has the ability to metabolize dimethyl-sulfoniopropionate (DMSP), a major source of organic sulfur produced by phytoplankton (Moran et al., 2003). Among DMSP-producing phytoplankton, *Alexandrium* sp. is known to have high DMSP cell content (Caruana and Malin 2014). Interestingly, a few ASVs from the xenic strain did not persist under normal culture maintenance conditions and were not found after six years, indicating that there are members that can maintain stable long-term partnerships while others do not. Thus, the assembly process is certainly not the sole driver of the observed microbial dynamics which included both the loss and emergence of ASVs over time within the community sequence sets, even in the xenic strain where the microbiomes were not manipulated (Supplementary Data S1).

The severe loss of diversity in the reconstructed free-living communities was unexpected. For example, free-living communities in the rexenic strain 2 showed the lowest diversity despite having the highest rexenic host function, so the effect of free-living communities on host physiology may be minimal. In contrast, we did not see a distinct loss of diversity from the attached communities. Instead, we found noticeable changes in diversity patterns in attached communities that parallel host function. The attached communities from the xenic and rexenic strain 2 showed increased diversity over time while rexenic strain 1 did not (Supplementary Fig. S6). Not surprisingly, rexenic strain 1 at time 1 exhibited the lowest host function. Considering the different host physiology observed from rexenic strain 1 at time 1 compared to xenic and rexenic 2 strains, attached communities in which diversity increases over time may represent healthy microbiomes that positively impact host physiology, thus attached communities are likely to be key in terms of microbiome diversity and host functioning. Taken together, our

15507408, ja, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/jeu.12966 by IFREMER Centre Bretagne BLP, Wiley Online Library on [13/02/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Accepted Article

results suggest that the composition of attached communities appears to be driven deterministically by host traits while free-living communities comprise a flexible pool more regulated by stochastic processes.

Roseobacter microdiversity in phytoplankton microbiome

The *Roseobacter* clade represents one of the most abundant and widely distributed bacterial groups in marine ecosystems (Luo and Moran 2014). Many are associated with phytoplankton and can metabolize DMSP produced by phytoplankton, thus playing a major role in oceanic sulfur cycling (Moran et al., 2003). Indeed, over half of the ASVs identified here belong to the *Roseobacter* clade (55%, 40 of 73 ASVs), and they accounted for 17-39% of the reads in attached and free-living communities over six years in xenic strains. Our microbiome reconstitution study found that the *Roseobacter* clade can virtually dominate the microbiome, comprising up to 88% and 98% of the sequence reads in attached and free-living communities from the rexenic strain 2 at time 2. Our results also revealed that the presence of highest microdiversity occurred among the *Roseobacter* clade members associated with the rexenic strain but not from the xenic strain. Interestingly, high microdiversity was observed among the rarely occurring *Roseobacter* clade members, while the most prevalent *Roseobacter* clade taxa generally had lesser or no observed microdiversity. Collectively, these results suggest that the microbial reassembly process can promote microbiome diversity through microdiversity (Fig. 5b). We also compared our *Roseobacter* ASVs to ones from field studies (Hasegawa et al., 2007; Hold et al., 2001; Jasti et al., 2005) to see whether they are truly the representatives of the wild population (Supplementary Fig. S7). Many *Roseobacter* clade taxa identified from our study were also detected from the other studies with highly divergent sequences, suggesting that the observed microdiversity is widely distributed and important.

CONCLUSIONS

From our results, it is clear that overall community composition does not tell the whole story of the phytoplankton microbiome. Hosts with similar function, such as the xenic and the rexenic 2 strain, may harbor microbiomes that are distinct in their overall structure. The relatively long time scale of this study shows that community assembly is dynamic over months to years following disturbance, and that it continues to occur, albeit at a much slower pace, in indigenous communities. Diverse phytoplankton-associated communities appear to be governed more by deterministic processes and promote better host function. In contrast, free-living microbial communities are less diverse and more flexible in their composition. Notably, reconstitution and assembly foster the establishment of microdiversity which may contribute to microbiome function. The observed microdiversity can be the result of microevolution coupling ecological and physiological diversity. Microdiversity in host-microbe systems can also represent the genetic basis for ecological adaptations to different niches resulting from interactions with the host and other microbial partners. Considering the limited time and space employed in this study, the observed microdiversity and estimates of microbial impacts on host physiology likely represent a minimum of the true ecological situation. Our results suggest more efforts are needed using representative host-microbe systems to gain systemic understanding of these highly dynamic host-microbe interactions.

ACKNOWLEDGEMENTS

This research was made possible by a grant from the National Oceanic and Atmospheric

15507408, ja, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/jeu.12966 by IFREMER Centre Bretagne BLP, Wiley Online Library on [13/02/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Administration Center for Sponsored Coastal Ocean Research ECOHAB program under award no. NA09NOS4780166 to the University of Texas Marine Science Institute. This is ECOHAB contribution number 839.

CONFLICT OF INTERESTS

The authors declare that they no competing interests.

REFERENCES

Amin, S. A., Hmelo, L. R., Van Tol, H. M., Durham, B. P., Carlson, L. T., Heal, K. R., Morales, R. L., Berthiaume, C. T., Parker, M. S., Djunaedi, B.. Ingalls, A. E., Parsek, M. R., Moran, M. A. & Armbrust. E. V. 2015. Interaction and signalling between a cosmopolitan phytoplankton and associated bacteria. *Nature,* 522:98-101.

Amin, S. A., Green, D. H., Hart, M. C., Küpper, F. C., Sunda, W. G. & Carrano, C. J. 2009. Photolysis of iron–siderophore chelates promotes bacterial–algal mutualism. *Proceedings of the National Academy of Sciences,* 106:17071-17076.

Andrews, S. (2010) FastQC: a quality control tool for high throughput sequence data. Babraham Bioinformatics, Babraham Institute, Cambridge, United Kingdom.

Behringer, G., Ochsenkühn, M. A., Fei, C., Fanning, J., Koester, J. A. & Amin, S. A. 2018. Bacterial communities of diatoms display strong conservation across strains and time. *Frontiers in Microbiology,* 9:659.

Bell, W. & Mitchell, R. 1972. Chemotactic and growth responses of marine bacteria to algal extracellular products. *The Biological Bulletin,* 143:265-277.

Bourne, D. G., Morrow, K. M. & Webster, N. S. 2016. Insights into the coral microbiome: underpinning the health and resilience of reef ecosystems. *Annual Review of Microbiology,* 70:317-340.

Brucker, R. M. & Bordenstein, S. R. 2013. The hologenomic basis of speciation: gut bacteria cause hybrid lethality in the genus Nasonia. *Science,* 341:667-669.

Buchan, A., González, J. M. & Moran, M. A. 2005. Overview of the marine Roseobacter lineage. *Applied and Environmental Microbiology,* 71:5665-5677.

Buchan, A., LeCleir, G. R., Gulvik, C. A. & González, J. M. 2014. Master recyclers: features and functions of bacteria associated with phytoplankton blooms. *Nature Reviews Microbiology,* 12:686-698.

Capella-Gutiérrez, S., Silla-Martínez, J. M. & Gabaldón, T. 2009. trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics,* 25:1972-1973.

Caruana, A. M. & Malin, G. 2014. The variability in DMSP content and DMSP lyase activity in marine dinoflagellates. *Progress in Oceanography,* 120:410-424.

Cooper, M. B. & Smith, A. G. 2015. Exploring mutualistic interactions between microalgae and bacteria in the omics age. *Current Opinion in Plant Biology,* 26:147-153.

David, L. A., Maurice, C. F., Carmody, R. N., Gootenberg, D. B., Button, J. E., Wolfe, B. E., Ling, A. V., Devlin, A. S., Varma, Y. & Fischbach, M. A. 2014. Diet rapidly and reproducibly alters the human gut microbiome. *Nature,* 505:559-563.

Edgar, R. C. 2016. SINTAX: a simple non-Bayesian taxonomy classifier for 16S and ITS sequences. *BioRxiv*:074161.

Edgar, R. C. 2010. Search and clustering orders of magnitude faster than BLAST. *Bioinformatics,* 26:2460-2461.

Accepted Article

15507408, ja, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/jeu.12966 by IFREMER Centre Bretagne BLP, Wiley Online Library on [13/02/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Accepted Article

Edgar, R. C. & Flyvbjerg, H. 2015. Error filtering, pair assembly and error correction for next-generation sequencing reads. *Bioinformatics,* 31:3476-3482.

Faith, J. J., Guruge, J. L., Charbonneau, M., Subramanian, S., Seedorf, H., Goodman, A. L., Clemente, J. C., Knight, R., Heath, A. C. & Leibel, R. L. 2013. The long-term stability of the human gut microbiota. *Science,* 341:1237439.

Fitzpatrick, C. R., Copeland, J., Wang, P. W., Guttman, D. S., Kotanen, P. M. & Johnson, M. T. 2018. Assembly and ecological function of the root microbiome across angiosperm plant species. *Proceedings of the National Academy of Sciences,* 115:E1157-E1165.

Fraune, S. & Bosch, T. C. 2007. Long-term maintenance of species-specific bacterial microbiota in the basal metazoan Hydra. *Proceedings of the National Academy of Sciences,* 104:13146-13151.

Goldford, J. E., Lu, N., Bajić, D., Estrela, S., Tikhonov, M., Sanchez-Gorostiaga, A., Segrè, D., Mehta, P. & Sanchez, A. 2018. Emergent simplicity in microbial community assembly. *Science,* 361:469-474.

Gould, A. L., Zhang, V., Lamberti, L., Jones, E. W., Obadia, B., Korasidis, N., Gavryushkin, A., Carlson, J. M., Beerenwinkel, N. & Ludington, W. B. 2018. Microbiome interactions shape host fitness. *Proceedings of the National Academy of Sciences,* 115:E11951-E11960.

Grossart, H. P., Levold, F., Allgaier, M., Simon, M. & Brinkhoff, T. 2005. Marine diatom species harbour distinct bacterial communities. *Environmental Microbiology,* 7:860-873.

Guillard, R. R. 1975. Culture of phytoplankton for feeding marine invertebrates. Culture of marine invertebrate animals. Springer.29-60.

Guindon, S., Dufayard, J.-F., Lefort, V., Anisimova, M., Hordijk, W. & Gascuel, O. 2010. New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Systematic Biology,* 59:307-321.

Hacquard, S., Garrido-Oter, R., González, A., Spaepen, S., Ackermann, G., Lebeis, S., McHardy, A. C., Dangl, J. L., Knight, R. & Ley, R. 2015. Microbiota and host nutrition across plant and animal kingdoms. *Cell Host & Microbe,* 17:603-616.

Hasegawa, Y., Martin, J. L., Giewat, M. W. & Rooney‐Varga, J. N. 2007. Microbial community diversity in the phycosphere of natural populations of the toxic alga, Alexandrium fundyense. *Environmental Microbiology,* 9:3108-3121.

Hold, G. L., Smith, E. A., Rappë, M. S., Maas, E. W., Moore, E. R., Stroempl, C., Stephen, J. R., Prosser, J. I., Birkbeck, T. H. & Gallacher, S. 2001. Characterisation of bacterial communities associated with toxic and non-toxic dinoflagellates: Alexandrium spp. and Scrippsiella trochoidea. *FEMS Microbiology Ecology,* 37:161-173.

Jasti, S., Sieracki, M. E., Poulton, N. J., Giewat, M. W. & Rooney-Varga, J. N. 2005. Phylogenetic diversity and specificity of bacteria closely associated with Alexandrium spp. and other phytoplankton. *Applied and Environmental Microbiology,* 71:3483-3494.

Jauzein, C., Evans, A. N. & Erdner, D. L. 2015. The impact of associated bacteria on morphology and physiology of the dinoflagellate Alexandrium tamarense. *Harmful Algae,* 50:65-75.

John, U., Litaker, R. W., Montresor, M., Murray, S., Brosnahan, M. L. & Anderson, D. M. 2014. Formal revision of the Alexandrium tamarense species complex (Dinophyceae) taxonomy: the introduction of five species with emphasis on molecular-based (rDNA) classification. *Protist,* 165:779-804.

15507408, ja, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/jeu.12966 by IFREMER Centre Bretagne BLP, Wiley Online Library on [13/02/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Joshi, N. & Fass, J. (2011) Sickle: A sliding-window, adaptive, quality-based trimming tool for FastQ files (Version 1.33)[Software].

Katoh, K. & Standley, D. M. 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Molecular Biology and Evolution,* 30:772-780.

Kirchman, D. L. 2002. The ecology of Cytophaga–Flavobacteria in aquatic environments. *FEMS Microbiology Ecology,* 39:91-100.

Koskella, B., Hall, L. J. & Metcalf, C. J. E. 2017. The microbiome beyond the horizon of ecological and evolutionary theory. *Nature Ecology & Evolution,* 1:1606-1615.

Lawson, C. A., Raina, J. B., Kahlke, T., Seymour, J. R. & Suggett, D. J. 2018. Defining the core microbiome of the symbiotic dinoflagellate, Symbiodinium. *Environmental Microbiology Reports,* 10:7-11.

Louca, S., Polz, M. F., Mazel, F., Albright, M. B., Huber, J. A., O'Connor, M. I., Ackermann, M., Hahn, A. S., Srivastava, D. S. & Crowe, S. A. 2018. Function and functional redundancy in microbial systems. *Nature Ecology & Evolution,* 2:936-943.

Luo, H. & Moran, M. A. 2014. Evolutionary ecology of the marine Roseobacter clade. *Microbiology and Molecular Biology Reviews,* 78:573-587.

Martin, M. 2011. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet. journal,* 17:10-12.

Matsen, F. A., Kodner, R. B. & Armbrust, E. V. 2010. pplacer: linear time maximum-likelihood and Bayesian phylogenetic placement of sequences onto a fixed reference tree. *BMC Bioinformatics,* 11:538.

Mazel, F., Davis, K. M., Loudon, A., Kwong, W. K., Groussin, M. & Parfrey, L. W. 2018. Is host filtering the main driver of phylosymbiosis across the tree of life? *mSystems,* 3:e00097-18.

McMurdie, P. J. & Holmes, S. 2013. phyloseq: an R package for reproducible interactive analysis and graphics of microbiome census data. *PloS One,* 8:e61217.

Mönnich, J., Tebben, J., Bergemann, J., Case, R., Wohlrab, S. & Harder, T. 2020. Niche-based assembly of bacterial consortia on the diatom Thalassiosira rotula is stable and reproducible. *The ISME Journal,* 14:1614-1625.

Moran, M. A., González, J. M. & Kiene, R. P. 2003. Linking a bacterial taxon to sulfur cycling in the sea: studies of the marine Roseobacter group. *Geomicrobiology Journal,* 20:375-388.

Oksanen, J., Kindt, R., Legendre, P., O'Hara, B., Stevens, M. H. H., Oksanen, M. J. & Suggests, M. 2007. The vegan package. *Community ecology package,* 10:631-637.

Paez-Espino, D., Eloe-Fadrosh, E. A., Pavlopoulos, G. A., Thomas, A. D., Huntemann, M., Mikhailova, N., Rubin, E., Ivanova, N. N. & Kyrpides, N. C. 2016. Uncovering Earth's virome. *Nature,* 536:425-430.

Pollock, F. J., McMinds, R., Smith, S., Bourne, D. G., Willis, B. L., Medina, M., Thurber, R. V. & Zaneveld, J. R. 2018. Coral-associated bacteria demonstrate phylosymbiosis and cophylogeny. *Nature Communications,* 9:1-13.

Rambaut, A. 2007. FigTree, a graphical viewer of phylogenetic trees.

Roberts, D. W. 2016. Package 'labdsv'. *Ordination and Multivariate,* 775.

Ronquist, F., Teslenko, M., Van Der Mark, P., Ayres, D. L., Darling, A., Höhna, S., Larget, B., Liu, L., Suchard, M. A. & Huelsenbeck, J. P. 2012. MrBayes 3.2: efficient Bayesian

Accepted Article

15507408, ja, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/jeu.12966 by IFREMER Centre Bretagne BLP, Wiley Online Library on [13/02/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Accepted Article

phylogenetic inference and model choice across a large model space. *Systematic Biology,* 61:539-542.

Rooney-Varga, J. N., Giewat, M. W., Savin, M. C., Sood, S., LeGresley, M. & Martin, J. 2005. Links between phytoplankton and bacterial community dynamics in a coastal marine environment. *Microbial Ecology,* 49:163-175.

Sapp, M., Wichels, A. & Gerdts, G. 2007. Impacts of cultivation of marine diatoms on the associated bacterial community. *Applied and Environmental Microbiology,* 73:3117-3120.

Seyedsayamdost, M. R., Case, R. J., Kolter, R. & Clardy, J. 2011. The Jekyll-and-Hyde chemistry of Phaeobacter gallaeciensis. *Nature Chemistry,* 3:331-335.

Seymour, J. R., Amin, S. A., Raina, J.-B. & Stocker, R. 2017. Zooming in on the phycosphere: the ecological interface for phytoplankton–bacteria relationships. *Nature Microbiology,* 2:17065.

Smriga, S., Fernandez, V. I., Mitchell, J. G. & Stocker, R. 2016. Chemotaxis toward phytoplankton drives organic matter partitioning among marine bacteria. *Proceedings of the National Academy of Sciences,* 113:1576-1581.

Sörenson, E., Bertos‐Fortis, M., Farnelid, H., Kremp, A., Krüger, K., Lindehoff, E. & Legrand, C. 2019. Consistency in microbiomes in cultures of Alexandrium species isolated from brackish and marine waters. *Environmental Microbiology Reports,* 11:425-433.

Stamatakis, A. 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics,* 30:1312-1313.

Stegen, J. C., Lin, X., Konopka, A. E. & Fredrickson, J. K. 2012. Stochastic and deterministic assembly processes in subsurface microbial communities. *The ISME Journal,* 6:1653-1664.

Sunagawa, S., Coelho, L. P., Chaffron, S., Kultima, J. R., Labadie, K., Salazar, G., Djahanschiri, B., Zeller, G., Mende, D. R. & Alberti, A. 2015. Structure and function of the global ocean microbiome. *Science,* 348:1261359.

Suzuki, R. & Shimodaira, H. 2006. Pvclust: an R package for assessing the uncertainty in hierarchical clustering. *Bioinformatics,* 22:1540-1542.

Taylor, M. W., Radax, R., Steger, D. & Wagner, M. 2007. Sponge-associated microorganisms: evolution, ecology, and biotechnological potential. *Microbiology and Molecular Biology Reviews,* 71:295-347.

Teeling, H., Fuchs, B. M., Becher, D., Klockow, C., Gardebrecht, A., Bennke, C. M., Kassabgy, M., Huang, S., Mann, A. J. & Waldmann, J. 2012. Substrate-controlled succession of marine bacterioplankton populations induced by a phytoplankton bloom. *Science,* 336:608-611.

Turnbaugh, P. J., Hamady, M., Yatsunenko, T., Cantarel, B. L., Duncan, A., Ley, R. E., Sogin, M. L., Jones, W. J., Roe, B. A. & Affourtit, J. P. 2009. A core gut microbiome in obese and lean twins. *Nature,* 457:480-484.

Unfried, F., Becker, S., Robb, C. S., Hehemann, J.-H., Markert, S., Heiden, S. E., Hinzke, T., Becher, D., Reintjes, G. & Krüger, K. 2018. Adaptive mechanisms that provide competitive advantages to marine bacteroidetes during microalgal blooms. *The ISME Journal,* 12:2894-2906.

Vergin, K. L., Beszteri, B., Monier, A., Thrash, J. C., Temperton, B., Treusch, A. H., Kilpert, F., Worden, A. Z. & Giovannoni, S. J. 2013. High-resolution SAR11 ecotype dynamics at

15507408, ja, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/jeu.12966 by IFREMER Centre Bretagne BLP, Wiley Online Library on [13/02/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

the Bermuda Atlantic Time-series Study site by phylogenetic placement of pyrosequences. *The ISME Journal,* 7:1322-1332.
Walter, J. & Ley, R. 2011. The human gut microbiome: ecology and recent evolutionary changes. *Annual Review of Microbiology,* 65:411-429.
Zeng, Q., Wu, S., Sukumaran, J. & Rodrigo, A. 2017. Models of microbiome evolution incorporating host and microbial selection. *Microbiome,* 5:127.
Zhou, J., Lyu, Y., Richlen, M. L., Anderson, D. M. & Cai, Z. 2016. Quorum sensing is a language of chemical signals and plays an ecological role in algal-bacterial interactions. *Critical Reviews in Plant Sciences,* 35:81-105.

FIGURE LEGENDS

FIGURE 1. Overview of the experimental setup for the phytoplankton microbiome reconstitution. A culture of xenic CCMP1493 was lysed and the lysate was filtered through a 5.0 µm Nitex mesh. The resulting bacterial filtrate was inoculated into three culture of axenic CCMP1598 to create three rexenic strains.

FIGURE 2. Changes in host physiology and the microbiome in the reconstituted conditions. Hierarchical clustering analysis with Bray-Curtis dissimilarity based on **a** the physiological parameters of the host phytoplankton, *A. tamarense* and **b** the relative abundance of associated microbial communities. The terms, 'high function', 'transition', and 'low function' are indicating the host physiological state based on the host functional signatures, such as growth rate, photosynthetic capability, carbon flux, and cell size (Table 1) (Jauzein et al., 2015). Approximately unbiased (AU) probability values based on the multiscale bootstrap resampling (10,000 replications) were calculated and expressed as *p*-values (%).

FIGURE 3. Alpha and beta diversity within the phytoplankton microbiome. **a** Principal coordinate analysis (PCoA) and **b-e** Shannon diversity index (H′) combined with multivariate analysis of variance. All samples were pooled together for the PCoA and the Shannon diversity indices were calculated based on the treatment (xenic, rexenic strain 1, and rexenic strain 2), lifestyle (free-living and attached), time-dependent effects (time 1 and time 2), and host physiology (high function, transition, and low function).

FIGURE 4. Detailed description of microbial community changes during the microbiome reconstitution using hierarchical clustering with Bray-Curtis dissimilarity, heatmap, indicator species, and the taxonomic classification based on the relative abundance of amplicons to overall microbial communities.

FIGURE 5. Diversity and distribution of *Roseobacter* clade in the phytoplankton microbiome. **a** A high-taxonomic resolution of *Roseobacter* clade tree for the finer-scale analyses of *Roseobacter* amplicons. **b** Overall description of *Roseobacter* clade changes during the microbiome reconstitution using hierarchical clustering, heatmap, and indicator species analyses based on the percentage amplicons relative to the total *Rhodobacteraceae*. Sequence divergence is included with percent sequence identities to the closest cultured *Roseobacter* clade representatives.

SUPPORTING INFORMATION

Accepted Article

15507408, ja, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/jeu.12966 by IFREMER Centre Bretagne BLP, Wiley Online Library on [13/02/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Accepted Article

FIGURE S1. ASV abundance distribution based on octave plots on log2 scale as described by Edgar and Flyvbjerg 2015. The x-axis shows the bins with read counts and the y-axis shows the number of ASVs associated with each bin.

FIGURE S2. Rarefaction curves based on all ASVs analyzed in the samples. The number of post-QC amplicons is reflected by the x-axis whereas the number of ASVs is on the y-axis.

FIGURE S3. Regression plot of Shannon diversity index (H′) versus the number of sample reads showing sequencing depth is not correlated with alpha diversity.

FIGURE S4. Principal component analysis of the host phytoplankton physiological parameters represented by growth rates (μ in h^{-1}), photosynthetic capacity (F_v/F_m, photochemical quantum yield of PSII), chlorophyll content (Chl *a* in pg μm^{-3}), carbon uptake (C uptake in h^{-1}) and content (C in pg $cell^{-1}$), and cell volume ($\times$ 1,000 μm^3) during microbiome reconstitution experiments.

FIGURE S5. Hierarchical clustering analysis with Bray-Curtis dissimilarity based on the relative abundance of **a** attached-specific and **b** free-living specific microbial communities. The terms, 'high function', 'transition', and 'low function' are indicating the host physiological state based on the host functional signatures, such as growth rate, photosynthetic capability, carbon flux, and cell size (Table 1) (Jauzein et al., 2015). Approximately unbiased (AU) probability values based on the multiscale bootstrap resampling (10,000 replications) were calculated and expressed as *p*-values (%).

FIGURE S6. Shannon diversity index (H′) combined with multivariate analysis of variance for each treatment (xenic, rexenic strain 1, and rexenic strain 2) based on the lifestyle (free-living and attached) and time-dependent effects (time 1 and time 2).

FIGURE S7. Divergence of *Roseobacter* 16S rRNA gene sequences observed from the phytoplankton microbiome and comparisons to ones from other studies. Sequences from this study are labeled red, and other studies are represented by dark blue (Hasegawa et al., 2007) and light blue (Jasti et al., 2005). Sequence divergence are included with percent sequence identities to the closest cultured *Roseobacter* clade representatives.

DATA S1. ASV table showing a total of 73 ASVs including the number of sequence reads in each ASV as well as the taxonomic classification of each ASV.

15507408, ja, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/jeu.12966 by IFREMER Centre Bretagne BLP, Wiley Online Library on [13/02/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

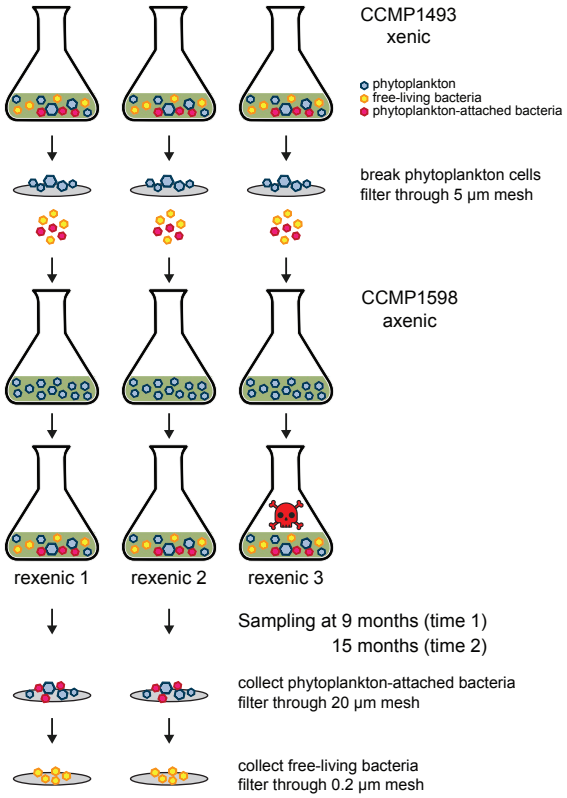
Accepted Article

Table 1 Host growth rates (μ in h^{-1}) and physiological metrics represented by photosynthetic capacity (F_v/F_m, photochemical quantum yield of PSII), chlorophyll content (Chl *a* in pg μm^{-3}), carbon uptake (C uptake in h^{-1}), cell volume (× 1,000 μm^3), and carbon content (C in pg $cell^{-1}$) during microbiome reconstitution experiments (Jauzein et al., 2015). Measurements from two rexenic strains (rexenic 1 and rexenic 2) were made at nine months (time 1) and 15 months (time 2) of culture maintenance after reintroducing bacteria from the xenic "parent" strain (CCMP1493) to the axenic strain (CCMP1598).

	xenic CCMP1493	axenic CCMP1598	rexenic1 time 1	rexenic1 time 2	rexenic2 time 1	rexenic2 time 2
μ	0.333 ± 0.035	0.260 ± 0.062	0.065 ± 0.007	0.183 ± 0.004	0.255 ± 0.007	0.280 ± 0.000
F_v/F_m	0.700 ± 0.013	0.392 ± 0.048	0.173 ± 0.016	0.605 ± 0.004	0.416 ± 0.073	0.692 ± 0.001
Chl *a*	3.330 ± 0.441	1.472 ± 0.336	0.876 ± 0.035	2.322 ± 0.109	1.772 ± 0.094	3.325 ± 0.037
C uptake	0.0220 ± 0.0028	0.0176 ± 0.0036	0.0016 ± 0.0003	0.0088 ± 0.0010	0.0151 ± 0.0003	0.0191 ± 0.0014
C content	1777.5 ± 142.9	993.1 ± 97.0	1054.2 ± 134.4	1781.5 ± 191.6	1810.3 ± 101.1	1893.2 ± 9.1
Cell volume	17.352 ± 0.258	9.996 ± 0.196	9.946 ± 0.159	13.796 ± 0.204	14.499 ± 0.348	16.427 ± 0.167

15507408, ja, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/jeu.12966 by IFREMER Centre Bretagne BLP, Wiley Online Library on [13/02/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

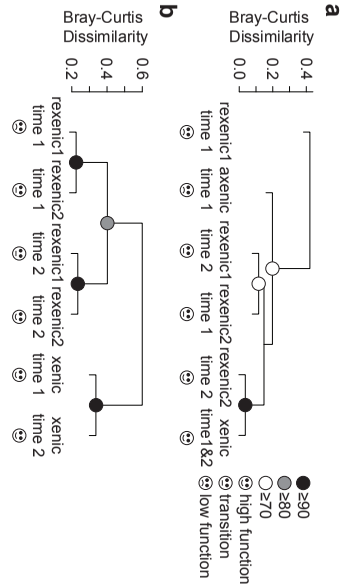
Accepted Article



15507408, ja, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/jeu.12966 by IFREMER Centre Bretagne BLP, Wiley Online Library on [13/02/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

15507408, ja, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/jeu.12966 by IFREMER Centre Bretagne BLP, Wiley Online Library on [13/02/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

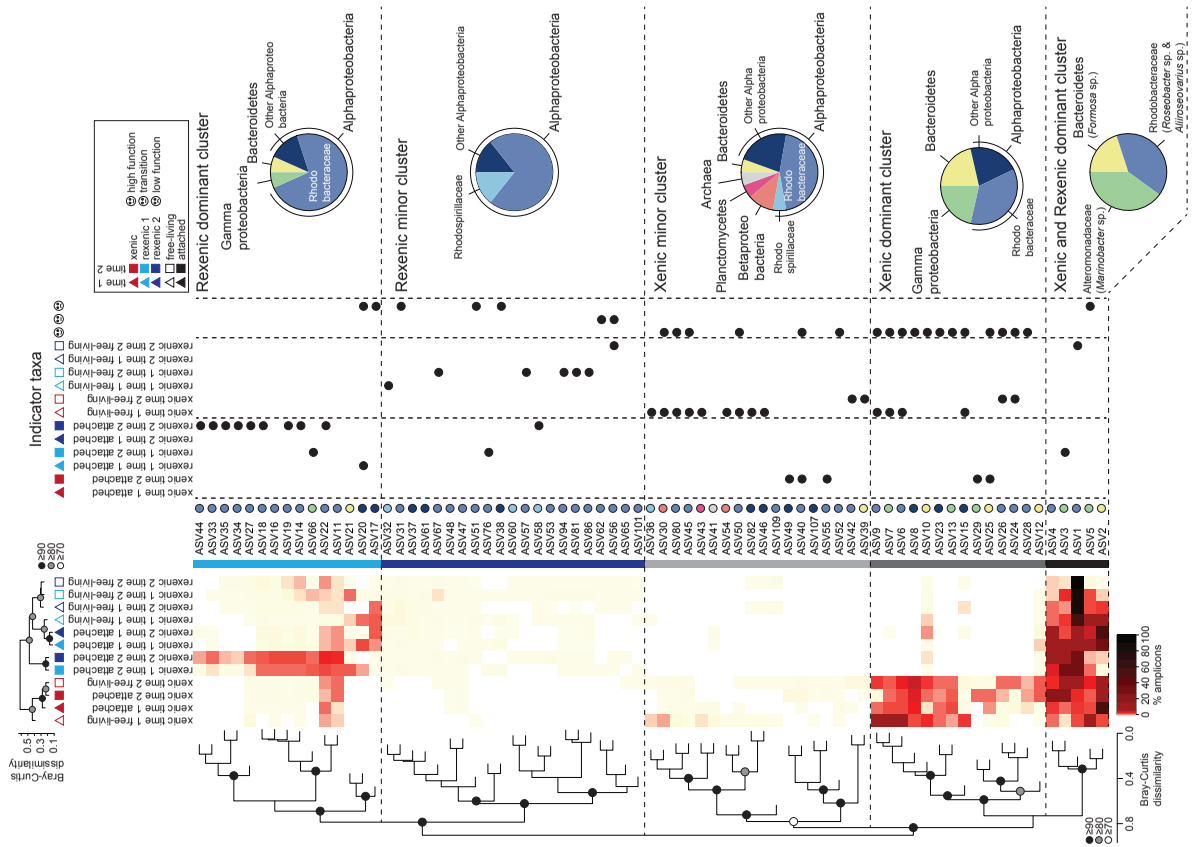
Accepted Article



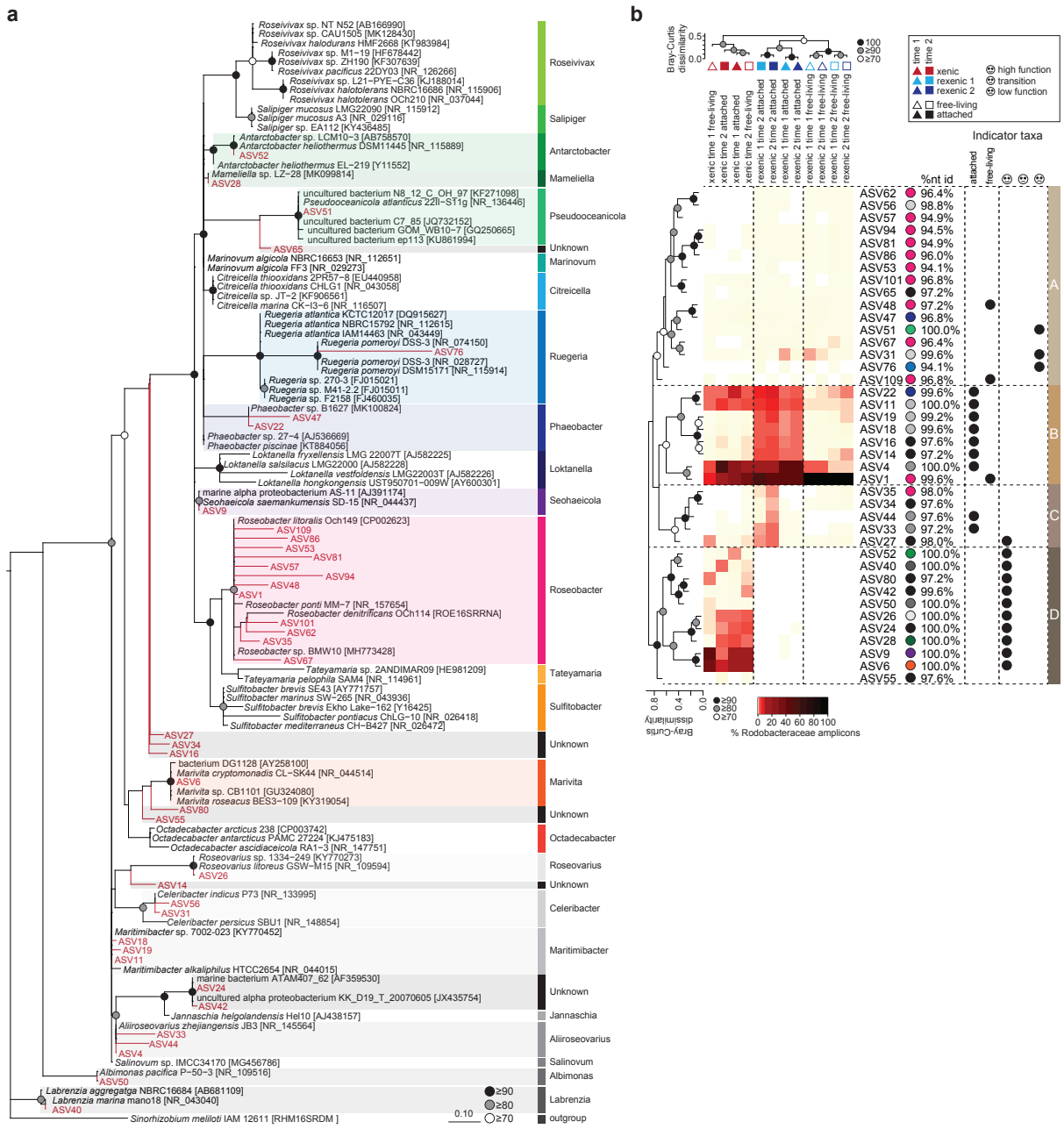
Accepted Article

15507408, ja, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/jeu.12966 by IFREMER Centre Bretagne BLP, Wiley Online Library on [13/02/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

15507408, ja, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/jeu.12966 by IFREMER Centre Bretagne BLP, Wiley Online Library on [13/02/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License



Accepted Article



Accepted Article

15507408, ja, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/jeu.12966 by IFREMER Centre Bretagne BLP, Wiley Online Library on [13/02/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License