Tight Adherence (Tad) Pilus Genes Indicate Putative Niche Differentiation in Phytoplankton Bloom Associated *Rhodobacterales*

Ashley Isaac[1,2], Ben Francis[2], Rudolf I. Amann[2] and Shady A. Amin[1]*

[1] Marine Microbial Ecology Laboratory, Biology Program, New York University Abu Dhabi, Abu Dhabi, United Arab Emirates,
[2] Max Planck Institute for Marine Microbiology, Bremen, Germany

The multiple interactions of phytoplankton and bacterioplankton are central for our understanding of aquatic environments. A prominent example of those is the consistent association of diatoms with *Alphaproteobacteria* of the order *Rhodobacterales*. These photoheterotrophic bacteria have traditionally been described as generalists that scavenge dissolved organic matter. Many observations suggest that members of this clade are specialized in colonizing the microenvironment of diatom cells, known as the phycosphere. However, the molecular mechanisms that differentiate *Rhodobacterales* generalists and phycosphere colonizers are poorly understood. We investigated *Rhodobacterales* in the North Sea during the 2010–2012 spring blooms using a time series of 38 deeply sequenced metagenomes and 10 metaproteomes collected throughout these events. *Rhodobacterales* metagenome assembled genomes (MAGs) were recurrently abundant. They exhibited the highest gene enrichment and protein expression of small-molecule transporters, such as monosaccharides, thiamine and polyamine transporters, and anaplerotic pathways, such as ethylmalonyl and propanoyl-CoA metabolic pathways, all suggestive of a generalist lifestyle. Metaproteomes indicated that the species represented by these MAGs were the dominant suppliers of vitamin B_{12} during the blooms, concomitant with a significant enrichment of genes related to vitamin B_{12} biosynthesis suggestive of association with diatom phycospheres. A closer examination of putative generalists and colonizers showed that putative generalists had persistently higher relative abundance throughout the blooms and thus produced more than 80% of *Rhodobacterales* transport proteins, suggesting rapid growth. In contrast, putative phycosphere colonizers exhibited large fluctuation in relative abundance across the different blooms and correlated strongly with particular diatom species that were dominant during the blooms each year. The defining feature of putative phycosphere colonizers is the presence of the tight adherence (*tad*) gene cluster, which is responsible for the assembly of adhesive pili that presumably enable attachment to diatom hosts. In addition, putative phycosphere colonizers possessed higher prevalence of secondary metabolite biosynthetic gene clusters, particularly

homoserine lactones, which can regulate bacterial attachment through quorum sensing. Altogether, these findings suggest that while many members of *Rhodobacterales* are competitive during diatom blooms, only a subset form close associations with diatoms by colonizing their phycospheres.

INTRODUCTION

Marine phytoplankton blooms are an essential part of ocean and global nutrient cycles. These blooms are transient, seasonal and localized events characterized by an increased abundance of algal cells and a spike in the assimilation of CO_2 and inorganic nutrients (nitrogen and phosphorus) (Behrenfeld et al., 2005; Behrenfeld et al., 2006). This ecosystem imbalance is accompanied by an increase in abundance and activity of heterotrophic bacteria (Buchan et al., 2014). These bacteria consume dissolved and particulate organic matter (DOM/POM) derived from both living and senescent/dead phytoplankton cells (Smith et al., 1992). Both DOM and POM consumption by bacteria contribute significantly to the biogeochemical cycles of carbon, nitrogen, and silicon (Azam and Long, 2001), and lead to remineralization of organic matter back to inorganic constituents required for phytoplankton growth (Bidle et al., 2002).

Helgoland Roads, a long-term ecological research station located in the German Bight of the North Sea hosts annual spring and summer blooms. Spring phytoplankton blooms in these shallow temperate waters are often dominated by diatoms (Eisma, 1987; Giebel et al., 2011), with multiple studies reporting on the composition of bacterioplankton communities (Eilers et al., 2000, 2001; Zubkov et al., 2001; Gerdts et al., 2004; Alderkamp et al., 2006; Rink et al., 2007). During the 2010–2012 spring algal blooms, members of the family *Flavobacteriaceae* were observed to be the most abundant and recurrent bacterial clade (Teeling et al., 2016). Due to their ability to degrade algal polysaccharides, they have thus been a primary focus of research (Xing et al., 2015; Francis et al., 2019, 2021; Kappelmann et al., 2019; Krüger et al., 2019; Avcı et al., 2020). However, members of the order *Rhodobacterales* (also referred to as Roseobacter(s) hereafter), are also frequently found in the North Sea with varying relative abundance (Alderkamp et al., 2006), and have been shown to exhibit significant changes in relative abundance in response to changes in phytoplankton composition and throughout different stages of algal blooms (Rink et al., 2007; González and Moran, 1997; Zubkov et al., 2001). They have also been regularly isolated from phytoplankton cultures (Behringer et al., 2018; Johansson et al., 2019; Wang et al., 2019), indicating that these bacteria are likely to interact closely with phytoplankton. Additionally, members of the Roseobacter group have been shown to influence the microbial community surrounding phytoplankton (Majzoub et al., 2019) and modify growth dynamics of the host (Bolch et al., 2017).

Phytoplankton are expected to be dependent on marine heterotrophic bacteria for supply of various cofactors, while bacteria require phytoplankton photosynthate for growth (Amin et al., 2012; Seymour et al., 2017). Vitamin exchange is among the most important metabolite exchanges that underpin algal-bacterial interactions. Estimates indicate that over half of marine algal species require exogenous sources of vitamin B_{12} (Croft et al., 2005), with the vitamin or its precursors acquired from bacteria, likely in exchange for dissolved organic carbon (Wienhausen et al., 2017). Cooper et al. (2019) demonstrated that cobalamin and thiamine auxotrophy in phytoplankton is alleviated by co-culture with *Dinoroseobacter shibae*, a member of the order *Rhodobacterales*, which in turn receives niacin (B_3), folate (B_9) precursors and biotin from the phytoplankton. Aside from vitamin exchange, it has also been demonstrated that the Roseobacter *Sulfitobacter pseudonitzschiae* SA11 (previously *Sulfitobacter* sp. SA11) promotes diatom cell division by the secretion of the hormone indol-3-acetic acid (Amin et al., 2015), which presumably increases DOM output by the diatom. This bacterium also upregulates genes responsible for uptake and assimilation of taurine produced by the diatom.

The region around phytoplankton cells is known as the phycosphere, and molecules within it (at a distance of approximately <100 μm) are mostly transported through diffusion rather than advection (Guasto et al., 2012). The phycosphere is enriched in exuded DOM comprising carbohydrates, amino acids, sugar alcohols and organic acids, which likely serve as chemoattractants for bacteria (Seymour et al., 2010). Recruiting chemotactic bacteria is potentially an important mechanism to establish associations between phytoplankton and beneficial bacteria and recent research suggests chemotaxis plays an important role in the microbial composition of the phycosphere (Raina et al., 2019). Chemotactic bacteria may remain in the phycosphere by attaching to phytoplankton cells or to phytoplankton transparent exopolymeric particles (TEP) (Samo et al., 2018), with attachment resulting in increased nutrient incorporation for both bacteria and phytoplankton and may be a long-term adaptation of specific bacteria 'symbiotic' with phytoplankton (Arandia-Gorostidi et al., 2017). Roseobacters have been shown to be the most transcriptionally active taxa in the microbiome of a model diatom, and were able to switch between motile and sessile lifestyles in response to diatom-derived metabolites (Fei et al., 2020; Shibl et al., 2020).

Given the frequency with which Roseobacters are observed to co-occur with phytoplankton, together with their metabolic variability, we sought to differentiate niches of Roseobacters that respond to diatom bloom events through the lens of metagenomics and metaproteomics. Although several members of this group have traditionally been described as ecological generalists, many observations suggest others are specialized in colonizing diatom phycospheres (Newton et al., 2010). However, the molecular mechanisms that differentiate *Rhodobacterales*

generalists and phycosphere colonizers are poorly understood. Here, we investigate the interactions of Roseobacters during spring diatom blooms in the North Sea and tease apart the molecular characteristics that distinguish putative generalists from phycosphere colonizers.

MATERIALS AND METHODS

Sampling

Helgoland Roads is a long-term research site that hosts annual spring phytoplankton blooms. Phytoplankton data was assessed on a weekly basis as part of the Helgoland Roads LTER time series and details of data acquisition have been described previously (Teeling et al., 2012; Teeling et al., 2016) (see **Supplementary Table 1** for dates on which samples were taken). Samples of bacterioplankton (bacteria and archaea) were collected at the 'Kabeltonne' station (54° 11.3′ N, 7° 54.0′ E) near Helgoland island during the 2010, 2011, and 2012 spring algal blooms at bi-weekly to weekly intervals as described previously (Teeling et al., 2012; Teeling et al., 2016). Briefly, surface seawater collected at a depth of ∼1 m was pre-filtered through 10 µm sized filters, followed by 3 µm filtration and finally bacterioplankton were collected on 0.2 µm filters from which DNA was extracted. From this 0.2 to 3 µm size fraction, DNA was extracted for metagenome sequencing and proteins were extracted for metaproteomics (Teeling et al., 2012, 2016).

Metagenome Assembled Genomes (MAGs)

Thirty-eight surface seawater samples were subject to metagenome sequencing at the DOE Joint Genome Institute on the Illumina HiSeq2000 platform as described previously (Teeling et al., 2016). Quality filtering, raw read trimming, metagenome assembly, binning and MAG refinement was performed through the Anvi'o pipeline as described previously (Krüger et al., 2019). Metagenomes were assembled and binned independently, hence redundant MAGs were obtained. The MAGs were clustered at 95% average nucleotide identity (ANI) which yielded 492 approximate species clusters (hereafter referred to as representative MAGs). Taxonomy of representative MAGs was determined by GTDB-Tk v0.8 (GTDB v89) classification and subsequently deposited into the European Nucleotide Archive under accession PRJEB28156. Representative MAGs with >70% completion or >69% for *Rhodobacterales* MAGs were selected for further analysis (392 MAGs in total). The threshold of completion for *Rhodobacterales* MAGs was lowered to conservatively maximize the number of *Rhodobacterales* MAGs included in downstream analysis (at >70% completion there were 32 *Rhodobacterales* MAGs and at >69% there were 33 MAGs). Metagenome raw reads were mapped to these MAGs at 97% identity. Read counts were normalized against MAG length to obtain reads per kilo base per million mapped reads (RPKM) [number of reads mapped ÷ (length of the MAG in kilobase pairs × total number of mapped reads in dataset ÷ 1,000,000)] and the relative abundance of each MAG in a metagenome was calculated ($RPKM_{(MAG)} \div RPKM_{(Sum\ of\ MAGs\ in\ metagenome)}$) as a reflection of *in situ* MAG abundance. MAG relative abundance figures and heatmaps were visualized with R 3.6.1 using ggplot2.

Metaproteomes

A total of 10 metaproteome from the 2010–2012 Helgoland spring blooms roughly corresponding to pre, middle and post bloom based on chlorophyll *a* concentrations (**Supplementary Table 1**) were obtained as described in Teeling et al. (2012) and Kappelmann et al. (2019). Briefly, proteins were extracted from bacterioplankton biomass, separated, subject to tryptic digestion and fragment detection carried out using an LTQ Orbitrap Velos mass spectrometer. Spectrometric data is available at the PRIDE repository (Vizcaíno et al., 2016) with the project ID PXD008238. The mass spectrometric data was analyzed using Sequest v27r11 and searches were carried out against a database of proteins from all corresponding metagenomes using USEARCH as described in Krüger et al. (2019). Protein and peptide validation were carried out and normalized spectral abundance factors (%NSAF) were calculated for semiquantitative analysis as previously described (Teeling et al., 2012; Krüger et al., 2019).

Bioinformatic Analyses

Representative MAG Gene Enrichment

Representative MAGs were annotated using EnrichM (v0.5.0) to obtain KEGG Orthology (KO) annotations as described in Boyd et al. (2019)[1]. The *classify* and *enrichment* subcommands were used to determine KEGG module completion and enriched KEGG modules, respectively, among the top four most abundant orders (comparisons between *Rhodobacterales* and *Flavobacterales*, *Pelagibacterales* and *Pseudomonadales*). The output files for presence/absence of KOs determined by Fisher's exact test and overrepresentation within a group determined by Mann-Whitney U test were further refined based on significance (corrected p value <0.05) and presence of the KO in >50% of MAGs within an order. KO annotations of significance were searched against the KO database[2] (Kanehisa et al., 2016) to obtain the number of enriched KOs that belong to specific BRITE hierarchies. The resulting heatmap was generated with GraphPad Prism 9.

Metaproteomic Analysis

Using standalone BLAST v2.10.0 +, predicted protein sequences from representative MAGs (392) were queried against a redundant metaproteome sequence database (3,212,324 sequences). Hits were filtered for percent identity >99% and *e*-value <1e-10. Proteins having significant hits to multiple MAGs were excluded from downstream analysis to avoid incorrectly assigning a protein to a MAG which may skew relative abundance calculations. The proteins with definitive hits against MAGs were queried against the KEGG database using an implementation of KofamKOALA to obtain KO numbers (Aramaki et al., 2020). BRITE hierarchical categories of the KO numbers were obtained with R package KEGGREST 1.24.1. The top 10 most abundant protein categories in the 2010 – 2012 metaproteome were identified based on the total

---

[1]http://github.com/geronimp/enrich

[2]https://www.genome.jp/kegg/ko.html

number of unique proteins per BRITE category. The %NSAF of proteins from each category were summed and the log of relative abundance was calculated for individual years and phase of bloom for each of the four most abundant taxonomic groups. The data plot was created and visualized with RStudio (Version 1.3.1093) using ggplot2.

Phylogenomics of *Rhodobacterales* MAGs

The *Rhodobacterales* MAGs were phylogenetically placed among 62 complete or draft genomes from the *Rhodobacteraceae* family downloaded from NCBI (**Supplementary Table 2**). Functional annotation of the representative MAGs and publicly available Roseobacter group genomes was performed using Prokka v1.14.0. Concatenated sequences of 107 single-copy core genes were obtained using bcgTree (Ankenbrand and Keller, 2016) implemented with HMMER v3.1b2 (Eddy, 2009), MUSCLE v3.8.31 (Edgar, 2004) and Gblocks 0.91b (Castresana, 2000) to create and refine a multiple sequence alignment. The phylogenomic tree was generated using RAxML (Stamatakis, 2014) through the ETE3 Python package (Huerta-Cepas et al., 2016) with predefined workflow standard_trimmed_raxml_bootstrap. The unrooted tree was visualized on the Interactive Tree of Life (iTOL) v5 (Letunic and Bork, 2019).

Niche Differentiation of *Rhodobacterales* MAGs

Significant differences in MAG size among *Rhodobacterales* phylogenetic groupings were confirmed by Student's t test (p value <0.05). All *Rhodobacterales* MAGs were submitted to secondary metabolite gene cluster analysis using ARTS 2.0 (Mungan et al., 2020), which implements antiSMASH 5.0 (Blin et al., 2019). Presence of biosynthetic gene clusters (BGCs) was tabulated. Given the inability to effectively differentiate *Rhodobacterales* based solely on phylogeny and secondary metabolism, the MAGs were reorganized into two groups based on their correlation to dominant diatom species during the blooms. MAG relative abundances were centered log-ratio transformed prior to statistical analysis. Redundancy analysis (RDA) was first used to infer the underlaying relationship between diatoms and Roseobacter MAGs using the R package, *vegan* (Version 2.5-7) (Oksanen et al., 2020). RDA is a constrained multivariate ordination technique that extracts gradients of variation in dependent variables (Roseobacter MAG relative abundance) explainable by independent variables (diatom cell counts) and assumes a linear relationship. Diatom cell counts were normalized by Hellinger-transformation prior to RDA to ensure the data met the statistical assumption of linearity. The significance of variation in Roseobacter MAG abundances explained by the explanatory variables was tested using an ANOVA-like Monte Carlo permutation test as implemented in *vegan*. Following confirmation of the relationship between Roseobacter MAGs and diatom species, diatom cell counts and MAG relative abundance were used to calculate Spearman rank correlation coefficients using GraphPad Prism 9. MAGs with statistically significant (p value <0.05) positive correlation coefficients, indicative of positive relationships to specific diatom species, were identified.

The relationships between MAGs and diatoms with positive correlations was visualized with Cytoscape v3.8.2. MAGs with no significant correlations to diatoms, those that were significantly negatively correlated to diatoms and those MAGs that had consistently high relative abundance were not considered as potential phycosphere colonizers. Consequently, a subset of MAGs with positive correlations to diatoms were categorized as potential phycosphere colonizers while the remaining *Rhodobacterales* MAGs were categorized as putative generalists. The *Rhodobacterales* MAGs were analyzed using Anvi'o 6.1 (Eren et al., 2015) following the microbial pangenomics workflow[3] in order to identify core, and enriched genes and functions based on the original phylogenetic groups and importantly, niche groupings.

Taxonomic Distribution and Completeness of the *Tad* Gene Locus

The *Aggregatibacter actinomycetemcomitans* TadA protein sequence (NCBI Accession No. AF152598) was used to search for homologs with BLASTX against the non-redundant protein sequences (nr) database and the resulting top 100 hits with percent identity > 60% and e-value = 0.0 were downloaded. These 100 TadA homologs together with TadA homologs from the *Rhodobacterales* MAGs were aligned with an online implementation of MUSCLE (Madeira et al., 2019) and the resulting multiple sequence alignment in ClustalW format used to generate a hidden Markov model (HMM) profile on HMMER v3.2.1 with hmmbuild. The hmm profile was queried against the Tara Oceans Microbiome Reference Gene Catalog version 1 on the Ocean Gene Atlas[4] webserver (Villar et al., 2018) with an e-value threshold of 1e-50. Taxonomic abundances of homologs found in surface and deep chlorophyll maximum samples across all size fractions (0–3 μm) and geographical locations was visualized as a single donut plot.

The *Rhodobacterales* MAGs, *A. actinomycetemcomitans* (available on RAST), *Sulfitobacter pseudonitzschiae* F5, *Phaeobacter* sp. F10 and *Alteromonas macleodii* F12 genomes (GenBank accession numbers WKFG01000000 and CP046140-CP046144 reported in Fei et al. (2020) were annotated using RAST (Aziz et al., 2008) in order to identify all genes and homologs that constitute the *tad* gene locus. Presence and completion of the *tad* gene cluster, also known as the Widespread Colonization Island (WCI) was determined by comparison against the *A. actinomycetemcomitans tad* cluster on SEED Viewer 2.0 (Overbeek et al., 2014). GBK files were downloaded and viewed in Geneious 11.1.5 and regions containing *tad* related genes were extracted and visualized with RStudio (Version 1.3.1093) using gggenes[5].

DATA AVAILABILITY STATEMENT

The metagenomic dataset PRJEB28156 for this study can be found in the European Nucleotide Archive under accession

---

[3]https://merenlab.org/2016/11/08/pangenomics-v2/
[4]http://tara-oceans.mio.osupytheas.fr/ocean-gene-atlas/
[5]https://github.com/wilkox/gggenes

[https://www.ebi.ac.uk/ena/data/view/PRJEB28156]. The metaproteomic spectrometric data ID PXD008238 for this study can be found in the PRIDE repository (Vizcaíno et al., 2016).

RESULTS AND DISCUSSION

MAG *in situ* Abundances and Comparative Genomics

We obtained 492 representative MAGs from 38 previously assembled 2010-2012 Helgoland metagenomes, of which 392 MAGs with CheckM completion > 69% for *Rhodobacterales* and > 70% for all other taxa were selected for downstream analysis (**Supplementary Table 3**). The abundances of these MAGs in the bloom samples were assessed by metagenomic read recruitment. These representative MAGs showed a pattern of succession across the three consecutive years that corroborated 16S rRNA gene amplicon sequencing data published by Teeling et al. (2016), which reported increased relative abundance of *Bacteroidetes* at the onset of the bloom, followed by *Gammaproteobacteria* and subsequently surpassed by *Alphaproteobacteria*. Similar patterns of abundance and recurrence have been reported in response to phytoplankton blooms (Williams et al., 2013; Needham and Fuhrman, 2016; Zhang et al., 2020) with heterotrophic members of the *Bacteroidetes*, *Gammaproteobacteria* and *Alphaproteobacteria*, particularly, *Rhodobacterales*, being among those groups that respond to phytoplankton-derived polysaccharides (Buchan et al., 2014). At a finer level of taxonomic resolution, it was revealed that *Rhodobacterales* MAGs were among the most recurrent and abundant throughout the datasets, in addition to *Flavobacteriales*, *Pelagibacterales* and *Pseudomonadales* (**Supplementary Figure 1**). *Flavobacteriales* are typically first responders to phytoplankton blooms consistent with their ability to break down complex algal derived organic matter, making labile compounds available to alphaproteobacterial populations that are generally dominated by the SAR11 clade and *Rhodobacterales* (Teeling et al., 2012; Williams et al., 2013).

Investigation into the transporter gene profile and expression of these transports suggest a difference in nutritional strategies along taxonomic lines. During the North Sea spring phytoplankton bloom of 2009, Teeling et al. (2012) showed that *Gammaproteobacteria* and *Flavobacteria* display similar transporter expression profiles with a dominant expression of TonB-dependent transporter (TBDT) components that are known to facilitate uptake of large compounds such as oligosaccharides, siderophores and vitamin B_{12}. On the other hand, *Alphaproteobacteria* and particularly, Roseobacters, showed high expression levels of adenosine triphosphate (ATP)–binding cassette (ABC) and tripartite ATP-independent periplasmic (TRAP) transporters for low-molecular-weight substrates. A subsequent metagenomic survey of four consecutive North Sea spring blooms (2009 – 2012) confirmed that TBDT genes were prevalent in *Bacteroidetes* while TRAP transporters were mostly present in *Alphaproteobacteria* (Teeling et al., 2016). In light of this and given that *Rhodobacterales, Flavobacteriales, Pelagibacterales* and *Pseudomonadales* MAGs were the most abundant taxa

(**Supplementary Figure 1**), we investigated the metabolic differences between these four major orders by searching for enriched KEGG orthologies (KO) and modules within the respective groups. With the relatively low number of *Rhodobacterales* MAGs (33 MAGs), we thought it prudent to compare the *Rhodobacterales* MAGs against each order individually rather than comparing them against the other orders collectively (359 MAGs), thus reducing the risk of diluting potential nuanced differences. In all scenarios *Rhodobacterales* were enriched for a number of KOs over diverse protein categories (EnrichM analysis) (**Supplementary Figure 2** and **Supplementary File 1**). *Pelagibacterales,* like *Rhodobacterales,* is classified as an alphaproteobacterial order, however there still appears to be distinct KO enrichment in the latter, suggesting a more diverse metabolic potential when compared to other *Alphaproteobacteria*. Hogle et al., 2016 reported that *Rhodobacterales* have a more versatile array of trace metal (e.g., nickel, zinc, manganese, and iron) uptake transporters when compared to SAR11 (*Pelagibacterales*), which have more streamlined genomes. In the current study it is observed that *Rhodobacterales* are enriched for nickel, molybdate and tungstate transporters when compared to *Pelagibacterales* (**Supplementary File 1**). *Rhodobacterales* possess enriched KOs for peptidases and inhibitors, amino acid related proteins, DNA repair and recombination proteins, ribosome biogenesis, tRNA biogenesis, secretion system and membrane trafficking compared to *Pelagibacterales*. Similar categories are enriched in *Rhodobacterales* MAGs when compared to *Flavobacteriales* and *Pseudomonadales*, however, the enrichment profile is much more pronounced. Other categories that are enriched in *Rhodobacterales* when compared to *Flavobacteriales* and *Pseudomonadales* include two-component systems, lipid biosynthesis proteins, lipopolysaccharide biosynthesis proteins and photosynthesis proteins. Enriched KOs in *Rhodobacterales* among all comparisons, of particular note, include those related to transport proteins (confirming the metagenomic survey in Teeling et al., 2016) and porphyrin and chlorophyll metabolism.

Transport proteins were the most highly enriched category in *Rhodobacterales* MAGs and porphyrin and chlorophyll metabolism is of interest as vitamin B_{12} biosynthetic intermediates are produced through this metabolic network. Specific KEGG modules that were consistently enriched in *Rhodobacterales* relative to the other orders were transporters of monosaccharides, thiamine and polyamines (**Supplementary File 1**). Physiological studies have shown that Roseobacters are able to grow on a wide variety of substrates and this versatility is reflected in their high number of transport proteins (Giebel et al., 2013; Voget et al., 2015). Anaplerotic pathways, including ethylmalonyl and propanoyl-CoA metabolic pathways were also enriched in *Rhodobacterales* MAGs. These pathways offer alternate metabolic routes for carbon assimilation. They promote growth on precursor metabolites (intermediates of central carbon metabolism) that replenish tricarboxylic acid (TCA) cycle intermediates, thus feeding central metabolism (Carlson et al., 2019). Collectively, these functional and metabolic features are suggestive of a generalist lifestyle commonly associated with members of the *Rhodobacterales*.

Additionally, enzymes related to cobalamin (vitamin B_{12}) biosynthesis and a tricarboxylic acid transport two-component regulatory system (TctD-TctE) were enriched in the *Rhodobacterales* MAGs. All but one of the *Rhodobacterales* MAGs possessed a complete biosynthetic pathway for vitamin B_{12}. *Rhodobacterales* MAG 20160419_Bin_37_1 lacked most of the biosynthetic pathway but does still possess the final enzyme in the pathway to produce vitamin B_{12} coenzyme. This confirms findings that the majority of Roseobacters have evidence of functioning cobalamin biosynthesis pathways, in contrast to more than 60% of marine bacterial species that are unable to synthesize vitamin B_{12} (Luo and Moran, 2014; Sanudo-Wilhelmy et al., 2014). Phytoplankton are often auxotrophic ($\sim$50% of >300 surveyed phytoplankton) for vitamin B_{12}; the ability of Roseobacters to produce vitamin B_{12} and their co-occurrence with phytoplankton may be indicative of interactions between them. The TctD-TctE system plays a role in resistance and tolerance to aminoglycosides in *Pseudomonas aeruginosa* within biofilms and thus contributes to persistence of bacteria within these mixed communities (Taylor et al., 2019). These features bring to light an alternate lifestyle of Roseobacters, one that is suggestive of association with phytoplankton phycospheres.

Metaproteomics

We analyzed 10 metaproteomes from the 2010 to 2012 spring algal blooms to investigate the *in situ* expression profiles of MAGs from the four most abundant orders and whether protein expression correlates with gene enrichment analyses. The proteins were compared via BLAST against all 392 representative MAGs and significant protein hits were subsequently placed into functional categories based on BRITE hierarchies. There were 5175 proteins that mapped to the four most abundant orders (**Supplementary Table 4**) with 66.9% of these proteins having KO annotations and accounting for 47% of all proteins that were mapped to representative MAGs. The 10 most abundant protein categories throughout the entire metaproteomics dataset and across all four abundant orders were related to central metabolism and cellular processes such as proteins found in glycolysis, the TCA cycle, ribosomes and RNA polymerases (**Figure 1**). In each category and in each of the four most abundant order there was a noticeable increase in protein relative abundance (%NSAF) as blooms progressed each year. Salazar et al., 2019 reported that community turn over and changes in gene expression could influence transcript abundance. At first glance the protein expression profile of the four most abundant orders appears to increase with their relative abundance as the bloom progresses (**Figure 1** and **Supplementary Figure 1**), suggesting gene copy number may be responsible for the perceived change in protein expression; however, there are specific instances, such as mid/post-bloom expression of transport proteins (2010, 2011 and 2012), where even at lower relative abundance than other orders, *Rhodobacterales* display higher levels of relative protein expression.
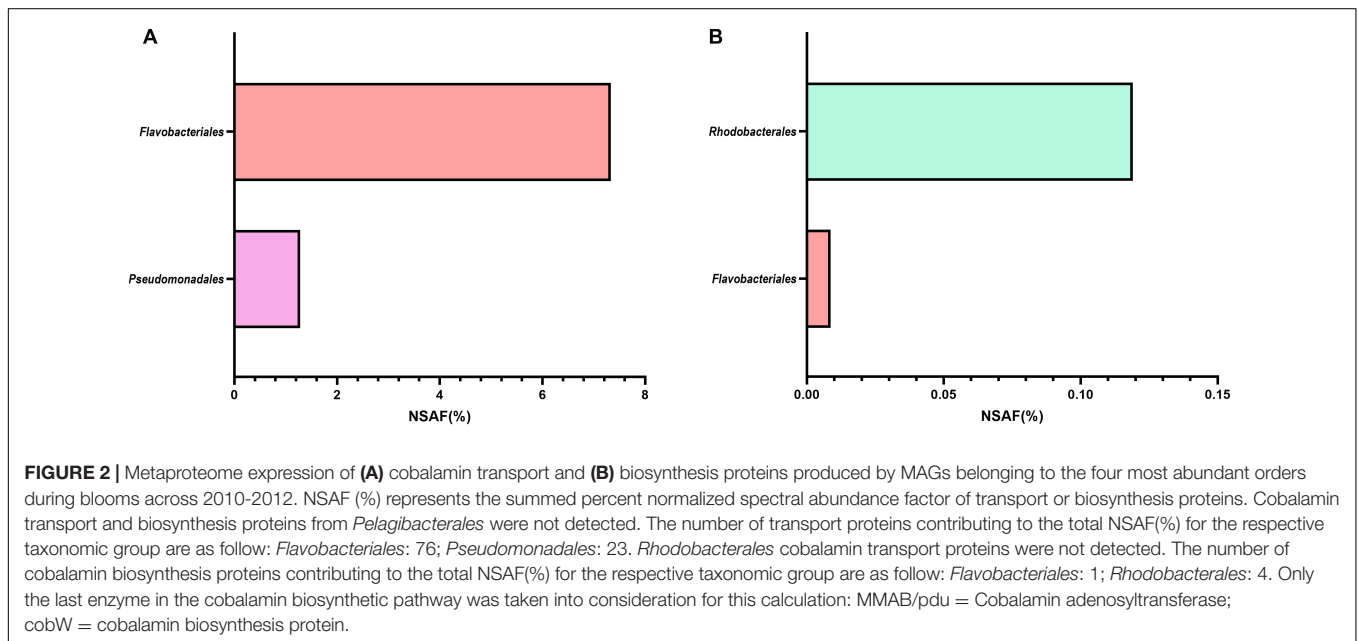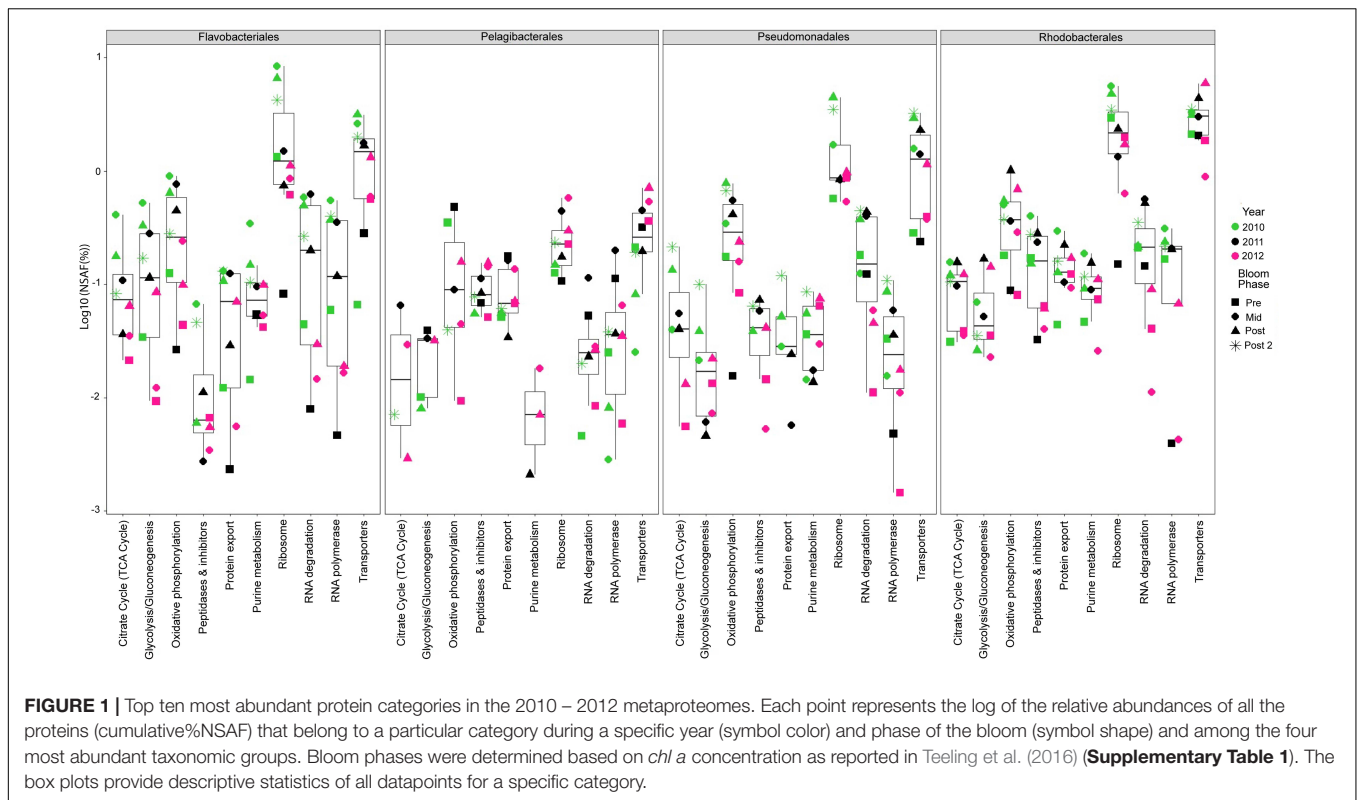
Consistent with the EnrichM KO enrichment analysis, *Rhodobacterales* MAGs consistently expressed the highest relative abundance (%NSAF) of transport proteins. Most of these transport proteins were annotated as amino acid transporters, low molecular weight sugar transporters (xylose, sorbitol/mannitol, fructose), putrescine transporters and microcin C transporters (**Supplementary Table 4**). Primary production due to phytoplankton blooms are assumed to be responsible for the availability of dissolved free carbohydrates such as glucose and other monosaccharides. It has been reported that Roseobacters were highly active during spring phytoplankton blooms in the North Sea and largely responsible for the uptake of glucose (Ittekkot et al., 1981; Alonso and Pernthaler, 2006).

Additionally, we looked specifically for the presence of proteins related to cobalamin transport and biosynthesis (**Figures 2A,B**). Representative MAGs of *Flavobacteriales* and *Pseudomonadales,* but not *Pelagibacterales* or *Rhodobacterales,* only expressed a number of cobalamin receptor proteins (**Figure 2A**). These receptor proteins bind extracellular vitamin B_{12} with high affinity before subsequent uptake (Pienko and Trylska, 2020). Confirming EnrichM gene enrichment analysis, proteins involved in the biosynthesis of vitamin B_{12} were produced almost exclusively by *Rhodobacterales* MAGs (**Figure 2B**). Only five proteins related to the direct production of vitamin B_{12} were found in the metaproteome, with only one produced by a *Flavobacteriales* MAG and expressed solely during the 2010 bloom while the remaining four proteins were produced by *Rhodobacterales* MAGs across multiple blooms. These findings corroborate the enrichment analysis that Roseobacters are likely the dominant suppliers of vitamin B_{12} during these blooms. Vitamin B_{12} is limiting in the ocean and bacteria in surface waters will likely compete with algae for vitamin B_{12} (Grant et al., 2014; Sanudo-Wilhelmy et al., 2014). Many phytoplankton-bacteria interactions hinge on the exchange of B vitamins and studies have shown that phytoplankton species meet their vitamin B_{12} requirement by entering into symbiotic relationships with specific bacteria, often times with Roseobacters, that provide vitamin B_{12} in exchange for fixed carbon and other metabolites (Croft et al., 2005; Grant et al., 2014; Durham et al., 2015; Cooper et al., 2019). In order to minimize competitive vitamin B_{12} uptake by *Flavobacteriales* and *Pseudomonadales* MAGs, the exchange of metabolites between *Rhodobacterales* MAGs and phytoplankton is likely to occur within the phycosphere, which may increase uptake of the vitamin by phytoplankton relative to vitamin production by free-living bacteria. Thus, we sought to identify specific *Rhodobacterales* MAGs that interact directly with diatoms by colonizing the phycosphere.

Niche Differentiation of *Rhodobacterales* MAGs

A phylogenetic survey of the 33 *Rhodobacterales* MAGs together with 62 publicly available *Rhodobacterales* genomes based on 107 conserved single copy genes (Ankenbrand and Keller, 2016) revealed that the MAGs fell into three distinct groups (**Figure 3**). There were 16, 9 and 8 MAGs that belonged to Groups I, II and III, respectively. Some of these MAGs were closely related to Roseobacters that were typically found in high abundance in North Sea waters, such as *Planktomarina temperata* RCA23 (20120510_Bin_64_1; 20160512_Bin_7_10; 20120503_Bin_5_4;

FIGURE 1 | Top ten most abundant protein categories in the 2010 – 2012 metaproteomes. Each point represents the log of the relative abundances of all the proteins (cumulative%NSAF) that belong to a particular category during a specific year (symbol color) and phase of the bloom (symbol shape) and among the four most abundant taxonomic groups. Bloom phases were determined based on *chl a* concentration as reported in Teeling et al. (2016) (**Supplementary Table 1**). The box plots provide descriptive statistics of all datapoints for a specific category.

FIGURE 2 | Metaproteome expression of **(A)** cobalamin transport and **(B)** biosynthesis proteins produced by MAGs belonging to the four most abundant orders during blooms across 2010-2012. NSAF (%) represents the summed percent normalized spectral abundance factor of transport or biosynthesis proteins. Cobalamin transport and biosynthesis proteins from *Pelagibacterales* were not detected. The number of transport proteins contributing to the total NSAF(%) for the respective taxonomic group are as follow: *Flavobacteriales*: 76; *Pseudomonadales*: 23. *Rhodobacterales* cobalamin transport proteins were not detected. The number of cobalamin biosynthesis proteins contributing to the total NSAF(%) for the respective taxonomic group are as follow: *Flavobacteriales*: 1; *Rhodobacterales*: 4. Only the last enzyme in the cobalamin biosynthetic pathway was taken into consideration for this calculation: MMAB/pdu = Cobalamin adenosyltransferase; cobW = cobalamin biosynthesis protein.

20100511_Bin_77) (Giebel et al., 2013) and *Planktotalea frisia* (20160426_Bin_44_1; 20110526_Bin_97_1) (Hahnke et al., 2012). Other MAGs were closely related to *Amylibacter* (20160512_Bin_25_4; 20100408_Bin_75_1; 20110530_Bin_77_1; 20160316_Bin_26_1; 20110506_Bin_63_1) and *Sulfitobacter* species (20110404_Bin_29_1; 20160419_Bin_37_1; 20160412_Bin_67; 20110523_Bin_85_1; 20120412_Bin_12;

20100423_Bin_54_1). Members of the Roseobacter group are highly diverse and span over 70 genera and hundreds of species (Pujalte et al., 2014; Simon et al., 2017). Most members of the Roseobacters have large genomes and display a high degree of genome plasticity that affords the Roseobacters their adaptability and diversity, while other pelagic Roseobacters have streamlined genomes and are possibly adapted to oligotrophic lifestyles

FIGURE 3 | Maximum likelihood unrooted phylogenomic tree of 107 conserved single copy genes from spring bloom associated *Rhodobacterales* MAGs (bold-faced) and publicly available *Rhodobacterales* genomes. Three branch points were identified and served as the basis for grouping the MAGs and genomes. Group I: Green, Group II: Red and Group III: Blue. The maximum likelihood phylogenomic tree was generated using RAxML and bootstrapping was performed at 100 iteration. Bootstrap values and sequence alignment can be found in the raw RAxML output file (**Supplementary File 2**).

(Luo and Moran, 2014; Luo et al., 2014; Voget et al., 2015; Simon et al., 2017). The MAGs in Group I (3.32 Mb ± 0.61 Mb) were found to have larger genomes than MAGs in Groups II (2.57 Mb ± 0.69 Mb) and III (2.71 Mb ± 0.31 Mb) (*t*-test, *p* value < 0.01 and < 0.05, respectively) (**Supplementary Figure 3**), suggesting a more diverse metabolic potential. Konstantinidis and Tiedje (2004) reported that larger genomes were found to be disproportionately enriched in genes related to secondary metabolism. Therefore, we postulated that the success of specific *Rhodobacterales* MAGs in colonizing diatom phycospheres might be linked to a suite of unique secondary metabolites that may enable or enhance phycosphere colonization. The MAGs were thus interrogated for the presence of potential biosynthetic gene clusters related to secondary metabolism (**Table 1**).

Group I MAGs had a more diverse genetic repertoire of secondary metabolites with the majority of them presenting with elevated counts of homoserine lactone biosynthetic gene clusters, which are responsible for the production of quorum sensing (QS) molecules. Quorum sensing systems utilize small diffusible molecules to synchronize gene expression of a bacterial population in a density-dependent manner. *N*-acyl-homoserine lactone (AHLs) are the most commonly described QS signaling molecules in *Proteobacteria*, which are synthesized by an autoinducer synthase (LuxI) and bind to an autoinducer

regulator (LuxR) to regulate a variety of genes (Case et al., 2008; Cude and Buchan, 2013). QS systems have been described in Roseobacters and have been implicated in the colonization of particulate matter (marine snow) in the ocean (Gram et al., 2002). Roseobacters isolated from diatom microbiomes are able to attach to and colonize diatom cells in response to particular AHL signals they produce (Fei et al., 2020). Group I MAGs also displayed a high prevalence for the production of putative antimicrobial compounds, such as bacteriocins, β-lactones and polyketides. These secondary metabolites have been shown to have potent bioactivity against bacteria and fungi (Ridley et al., 2008; Cotter et al., 2013; Robinson et al., 2019) and may play a role in chemical warfare within the phycosphere by eliminating competitors and protecting phytoplankton cells from potential pathogens. Several members from Groups II and III appear to share a similar biosynthetic gene profile and thus the identification of potential colonizers cannot be based solely on phylogeny. Indeed, Newton et al. (2010) reported that phylogenetic tree topology was not the best model for organizing genome characteristics in Roseobacters and that it was a lifestyle framework that better suited grouping of Roseobacter isolates. Thus, we organized and grouped the *Rhodobacterales* MAGs by their abundance relative to different diatom species occurring in the different blooms.

TABLE 1 | Putative biosynthetic gene clusters required for the production of secondary metabolites in *Rhodobacterales* MAGs.
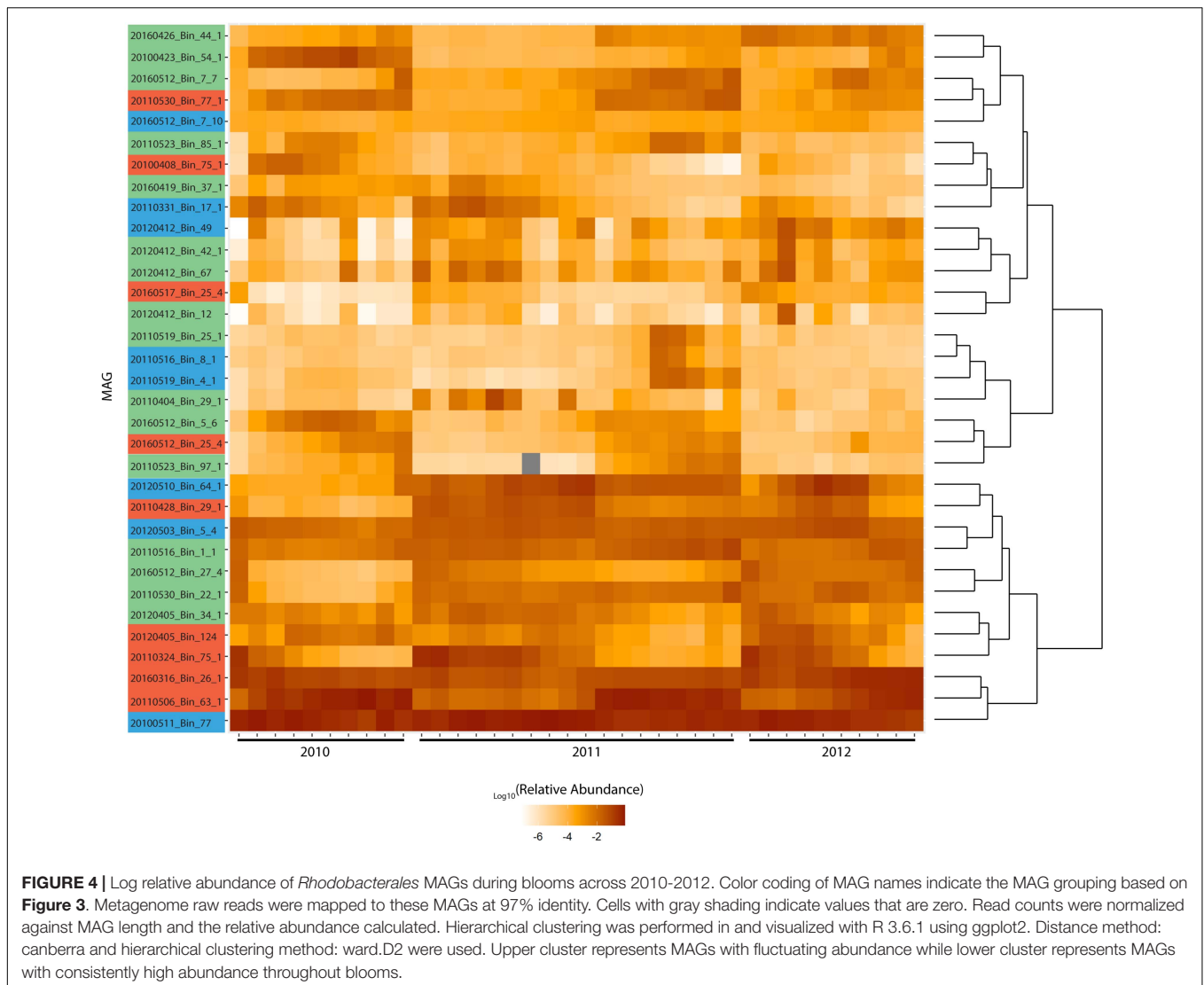
Group	MAG	Terpene	β-lactone	Hserlactone	T1PKS	T3PKS	Ectoine	NRPS	NRPS-like	TPA	Bacteriocin
Group I	20100423_Bin_54_1		•	•			•				•
	20110404_Bin_29_1			•	•	•			•		•
	20110523_Bin_85_1	•	•	•	•		•		•		
	20110523_Bin_97_1	•	•	••					•		
	20120405_Bin_34_1	•		••			•				•
	20120412_Bin_12			••		•		•	•		•
	20120412_Bin_67			•							
	20160419_Bin_37_1		••	••	•					•	•
	20160426_Bin_44_1		•	•		•					
	20160512_Bin_27_4	•		•							•
	20160512_Bin_5_6		•	••			•				•
	20160512_Bin_7_7	•	•	•		•					•
	20110519_Bin_25_1	••		••							•
	20120412_Bin_42_1	•				•			•		•
	20110516_Bin_1_1	•	•		•				••		
	20110530_Bin_22_1	•	•								
Group II	20100408_Bin_75_1	•	•	••							
	20160512_Bin_25_4			•		•	•				•
	20160517_Bin_25_4	•		•	••				•		
	20110324_Bin_75_1										
	20110428_Bin_29_1	••									
	20110506_Bin_63_1	••									
	20110530_Bin_77_1	••									
	20120405_Bin_124	•									
	20160316_Bin_26_1	••									
Group III	20110331_Bin_17_1			•••							•
	20110516_Bin_8_1	•		•	•				•		
	20110519_Bin_4_1			•							•
	20120412_Bin_49			••	•	•	•		••		
	20120503_Bin_5_4	•									•
	20160512_Bin_7_10	•									
	20120510_Bin_64_1	•									
	20100511_Bin_77	•									•

Each ● symbol represents a single complete biosynthetic gene cluster, hence ●● or ● ● ● means two or three complete gene clusters, respectively. Hserlactone: Homoserine lactone; T1PKS: Type I Polyketide synthase, T3PKS: Type III Polyketide synthase; NRPS: Non-ribosomal Peptide Synthase; TPA: Tropodithietic-acid.

The MAGs were clustered based on their relative abundance during the blooms, which yielded a distinct bifurcation in their hierarchical clustering (**Figure 4**). The two clusters represent those that are consistently present in high abundance throughout the blooms and those that show clear patterns of fluctuation within and across different blooms. There was no pattern that indicates broad phylogenetic groupings has any bearing on the response to phytoplankton succession. In marine ecosystems it has been shown that finer levels of taxonomy can display stronger patterns of correlation which can unravel specific interactions, such as those between microbial hosts and their viruses (Needham et al., 2017). Interestingly, Chafee et al. (2018) reported both broad (those that occur throughout the year) and narrow (those that are enriched during spring and summer) recurrent patterns of *Rhodobacterales* in response to seasonal bloom dynamics. Roseobacter group oligotypes of *Amylibacter* and *Planktomarina* were among five broadly recurrent taxa which collectively contributed to ∼20% of the total community throughout the year while oligotypes of *Sulfitobacter*, *Roseobacter* and *Loktanella* among others, were recurrent and enriched during spring and summer when phytoplankton blooms occur (Chafee et al., 2018). These findings further support the distinction between Roseobacter generalists and close algal associates.

Redundancy analysis (RDA) was carried out to assess the degree to which the Roseobacter MAGs were associated with different diatom species during the individual blooms (**Supplementary Figures 4A-C**). This analysis revealed that 69, 37, and 64% (adjusted-R-squared value of the model) of the variation observed in the Roseobacter MAG abundances can be explained by diatom abundance in 2010, 2011 and 2012, respectively. The relationship between Roseobacter abundance and diatoms were significant in 2010 (Monte Carlo Permutation test, $n = 999$; $F = 6.05$; $p = 0.001$) (**Supplementary Figure 4A**),
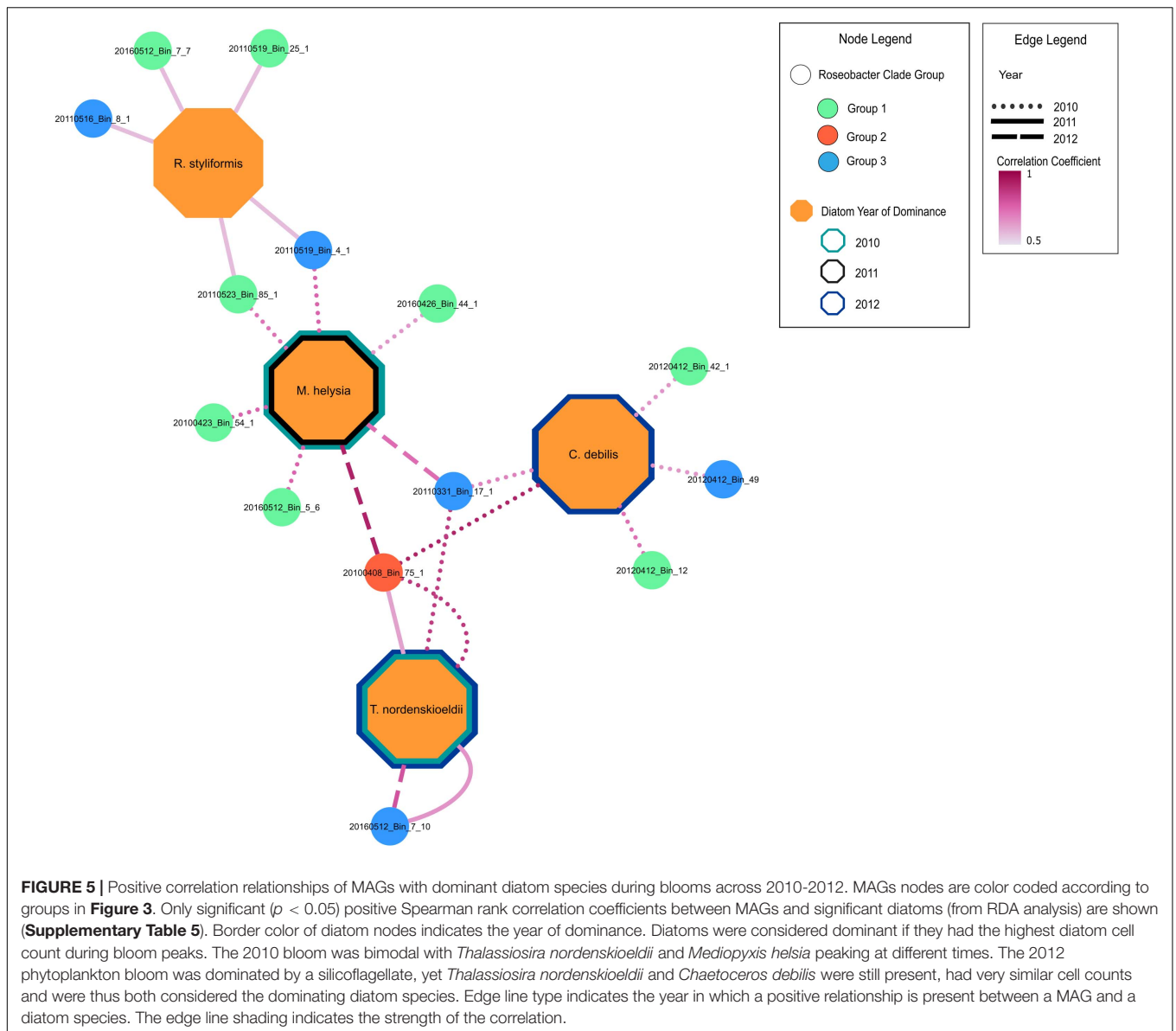
FIGURE 4 | Log relative abundance of *Rhodobacterales* MAGs during blooms across 2010-2012. Color coding of MAG names indicate the MAG grouping based on **Figure 3**. Metagenome raw reads were mapped to these MAGs at 97% identity. Cells with gray shading indicate values that are zero. Read counts were normalized against MAG length and the relative abundance calculated. Hierarchical clustering was performed in and visualized with R 3.6.1 using ggplot2. Distance method: canberra and hierarchical clustering method: ward.D2 were used. Upper cluster represents MAGs with fluctuating abundance while lower cluster represents MAGs with consistently high abundance throughout blooms.

2011 (Monte Carlo Permutation test, $n = 999$; $F = 2.75$; $p = 0.029$) (**Supplementary Figure 4B**) and 2012 (Monte Carlo Permutation test, $n = 999$; $F = 4.99$; $p = 0.001$) (**Supplementary Figure 4C**). In order to refine the distinction between putative generalists and phycosphere colonizers, MAG relative abundances (**Figure 4**) and diatom cell counts (**Supplementary Figure 5**) were used to compute Spearman rank correlation coefficients, which provide information on the strength of the relationship between specific MAGs and diatom species. Twenty-two MAGs had statistically significant (p value < 0.05) positive correlation coefficients that were all greater than 0.5, indicating moderate to strong positive relationships to specific diatom species (**Supplementary Table 5**). MAGs with no significant correlations to diatoms, those that were only negatively correlated to diatoms and those MAGs that had consistently high relative abundance (from **Figure 4**) were not considered as potential phycosphere colonizers. Consequently, 14 MAGs with positive correlations to diatoms were identified as potential colonizers (**Figure 5**) which were then compared against the remaining 19

Roseobacter MAGs that were categorized as putative generalists. It should be noted that due to the low number of independent samples ($n = 3$, one per year), these correlations are not as robust as we would like it to be.

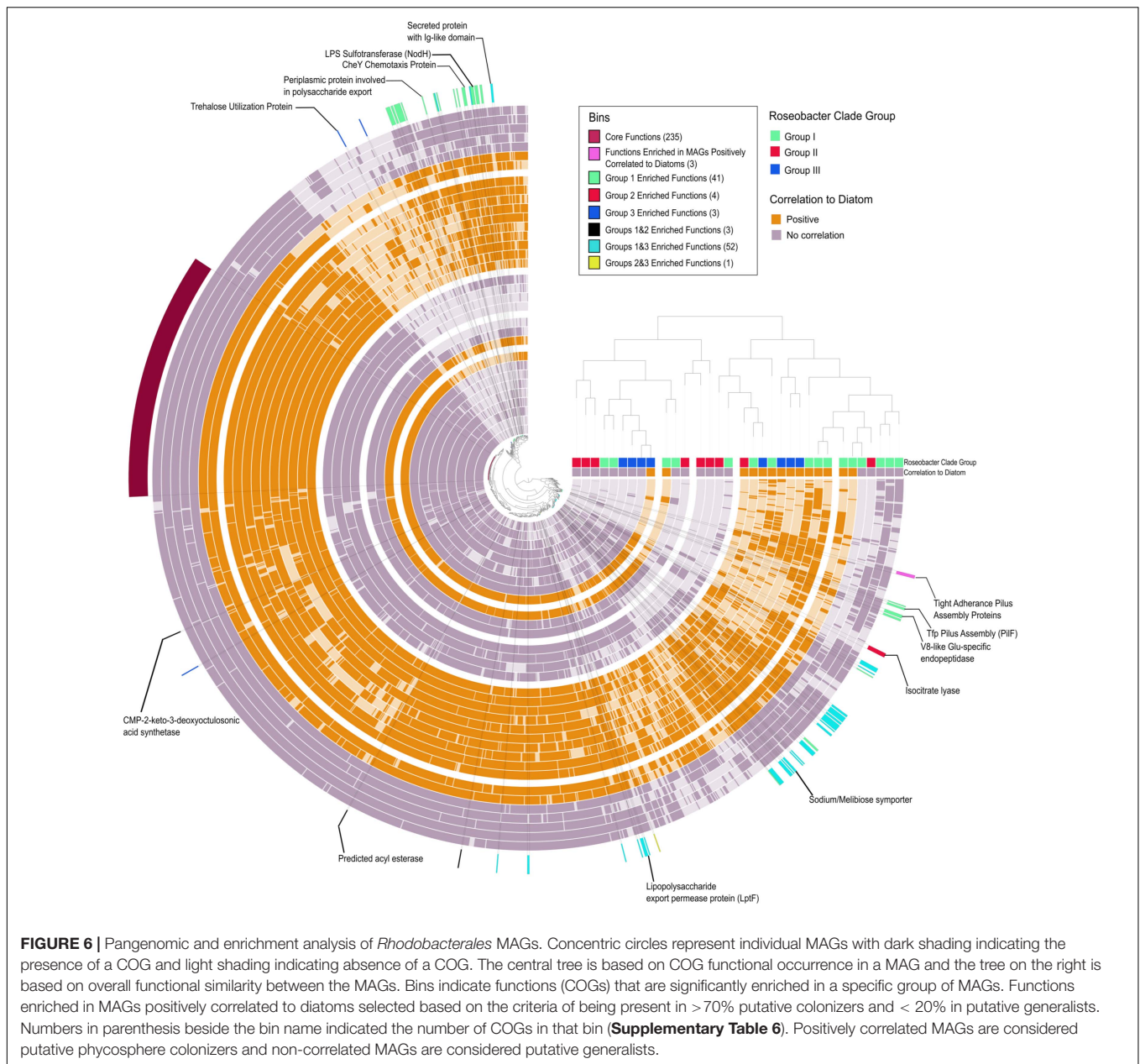## Putative Generalists vs. Phycosphere Colonizers

There are a number of studies that provide an overview of Roseobacter genome content (Moran et al., 2004; Buchan et al., 2005; Moran et al., 2007; Brinkhoff et al., 2008; Newton et al., 2010; Luo and Moran, 2014), however, metagenomics surveys have demonstrated that our view of the ecology of Roseobacters are skewed and biased toward cultured Roseobacters (Luo et al., 2012; Luo and Moran, 2014). Thus, we attempted to tease apart the genomic differences between bloom associated Roseobacters to understand if only a subpopulation of Roseobacters is able to colonize the phycosphere (Anvi'o enrichment analysis) (**Figure 6** and **Supplementary Table 6**). There were 235 core

FIGURE 5 | Positive correlation relationships of MAGs with dominant diatom species during blooms across 2010-2012. MAGs nodes are color coded according to groups in **Figure 3**. Only significant (*p* < 0.05) positive Spearman rank correlation coefficients between MAGs and significant diatoms (from RDA analysis) are shown (**Supplementary Table 5**). Border color of diatom nodes indicates the year of dominance. Diatoms were considered dominant if they had the highest diatom cell count during bloom peaks. The 2010 bloom was bimodal with *Thalassiosira nordenskioeldii* and *Mediopyxis helsia* peaking at different times. The 2012 phytoplankton bloom was dominated by a silicoflagellate, yet *Thalassiosira nordenskioeldii* and *Chaetoceros debilis* were still present, had very similar cell counts and were thus both considered the dominating diatom species. Edge line type indicates the year in which a positive relationship is present between a MAG and a diatom species. The edge line shading indicates the strength of the correlation.

protein functions [clusters of orthologous groups (COGs)] shared amongst the 33 *Rhodobacterales* MAGs that included general cellular processes, central metabolism and transport proteins that correspond to the enriched transporters discussed earlier. Members of Group I displayed the greatest number of enriched functions (41 enriched COGs), with many of them involved in chemotaxis, biofilm formation and attachment. This suggests that Group I members have a greater potential to successfully colonize the phycosphere. In contrast, Group II was only enriched in 6 COGs, one of which is isocitrate lyase, which is important in the glyoxylate pathway and serves as an alternate to the TCA cycle to enable bacteria to convert fatty acids into carbohydrates (Maharjan et al., 2005). Polyunsaturated fatty acids are produced by a number of microalgae and have a considerable contribution to the marine food web and may serve as a carbon source for marine microbes and eukaryotes

(del Rosario González-Baró and Pollero, 1998; Sasso et al., 2012). Group III MAGs displayed 3 enriched functions, one of them being a trehalose utilization protein. Some microalgae are known to produce high amounts of trehalose at night (Hirth et al., 2017), suggesting interactions between Group III members and phytoplankton could be diurnally regulated. Cumulatively, these findings suggest that distinct diatom metabolites may serve as unique currencies to a select few groups of Roseobacters. Although Group I members were significantly enriched for functions associated with interactions with phytoplankton, as observed with secondary metabolism, specific members from Groups II and III possess some, or all these features as well.

Based on the correlation analysis with diatom cell counts (**Figures 4**, **5**), putative generalist and phycosphere colonizer MAGs were investigated for gene enrichments. The enrichment analysis did not identify COGs that were significantly enriched,

FIGURE 6 | Pangenomic and enrichment analysis of *Rhodobacterales* MAGs. Concentric circles represent individual MAGs with dark shading indicating the presence of a COG and light shading indicating absence of a COG. The central tree is based on COG functional occurrence in a MAG and the tree on the right is based on overall functional similarity between the MAGs. Bins indicate functions (COGs) that are significantly enriched in a specific group of MAGs. Functions enriched in MAGs positively correlated to diatoms selected based on the criteria of being present in >70% putative colonizers and < 20% in putative generalists. Numbers in parenthesis beside the bin name indicated the number of COGs in that bin (**Supplementary Table 6**). Positively correlated MAGs are considered putative phycosphere colonizers and non-correlated MAGs are considered putative generalists.

however, specific COGs which met a strict selection criterion (> 70% occurrence in putative colonizers and < 20% occurrence in generalists) were of particular interest. Surprisingly, three COGs (CpaF, TadB and TadC) that contribute to the production of tight adherence (Tad) pilus proteins were the only features that set putative phycosphere colonizers apart from the putative generalists (**Figure 6** and **Supplementary Table 6**). A potential caveat exists in that the *Rhodobacterales* MAGs are not complete and may influence gene presence/absence and enrichment. However, given that potential phycosphere colonizers are less complete on average (88%) than generalists (91%), there would be bias against phycosphere colonizers and thus, this gives us more confidence in our results. These proteins make up a secretion system (a subtype of type IV secretion systems)

that is required for the assembly and secretion of adhesive fimbrial low-molecular-weight proteins, otherwise known as Flp pili (Tomich et al., 2007). Flp pili have been shown to contribute to the attachment of *Pseudomonas aeruginosa* to both biotic and abiotic surfaces (de Bentzmann et al., 2006) and *tad* genes were proven to be essential through transposon mutagenesis for biofilm formation, colonization and pathogenesis in a number of genera in diverse niches (Kachlany et al., 2001; Clock et al., 2008; Pu et al., 2018). In contrast to other type IV pilus systems, Tad clusters seem to lack dedicated retraction machinery and most Tad pili do not seem to be dynamic, precluding twitching motility (Tomich et al., 2007). However, recent findings by Sangermani et al. (2019) show that the *Caulobacter crescentus* Tad pili mediates initial attachment to surfaces and undergoes repeated

cycles of extension and retracting which helps reorient cells ahead of adhesive exopolysaccharide (EPS) (holdfast) production which establishes long-term attachment. The presence of adhesive pili and holdfasts have previously been reported in marine Roseobacters (Gonzalez et al., 1997; Slightom and Buchan, 2009; Bartling et al., 2018), yet this is not a feature that is universal to Roseobacters as described by Voget et al. (2015). The prevalence and importance of the *tad* gene locus in Roseobacters in association with diatoms has not been adequately investigated. The *tad* locus was first described in *Aggregatibacter (Actinobacillus) actinomycetemcomitans*. It is composed of 14 genes (*flp1–flp2–tadV–rcpCAB–tadZABCDEFG*) of which 12 are essential for fibril secretion (all except *flp2*, which does not seem to be expressed and *rcpB* which could not be fully assessed due to methodological limitations) and, therefore, the adherence phenotype (Kachlany et al., 2000, 2001; Planet et al., 2003). The *tad* gene locus appears to be present in many bacterial and archaeal genera and can be both chromosome and plasmid borne, indicating a propensity for horizontal gene transfer and is thus often termed the Widespread Colonization Island (Planet et al., 2003; Torres-Monroy and Ullrich, 2018). The TadA protein functions as an ATPase and provides the energy for Flp pilus assembly. Mutations in conserved regions of TadA abolished Flp-pilus production and disrupted the adherence

phenotype in *A. actinomycetemcomitans* (Bhattacharjee et al., 2001). We selected the *tadA* gene to examine its taxonomic distribution in the Tara Oceans dataset as a proxy for the distribution of the entire cluster in marine microbes. An HMM profile (**Supplementary File 3**) constructed from 114 *tadA* genes (including those found in our *Rhodobacterales* MAGs) was searched against the Tara Oceans database. Apart from Archaea, Roseobacters were the largest group of prokaryotes to possess the *tadA* gene and by extension, the *tad* locus, illustrative of the likely importance of these genes in marine Roseobacters and their ability to colonize surfaces (**Supplementary Figure 6**).

The *Rhodobacterales* MAGs were subsequently mined for the presence of a complete *tad* system and other genes that may be important for phycosphere colonization (**Figure 7** and **Supplementary Figure 7**). When compared against the *A. actinomycetemcomitans tad* cluster, more than 70% of the putative phycosphere colonizer MAGs possessed pili genes and complete or nearly complete *tad* gene loci, significantly more (t-test, p value < 0.01) than putative generalists of which there were only two MAGs from Group I that possessed the complete locus. In general, all the MAGs possessed genes related to motility and 69% of *Rhodobacterales* MAGs possessed genes related to chemotaxis and were significantly present in colonizer MAGs (t-test, p value < 0.05). These features enable bacteria to sense local
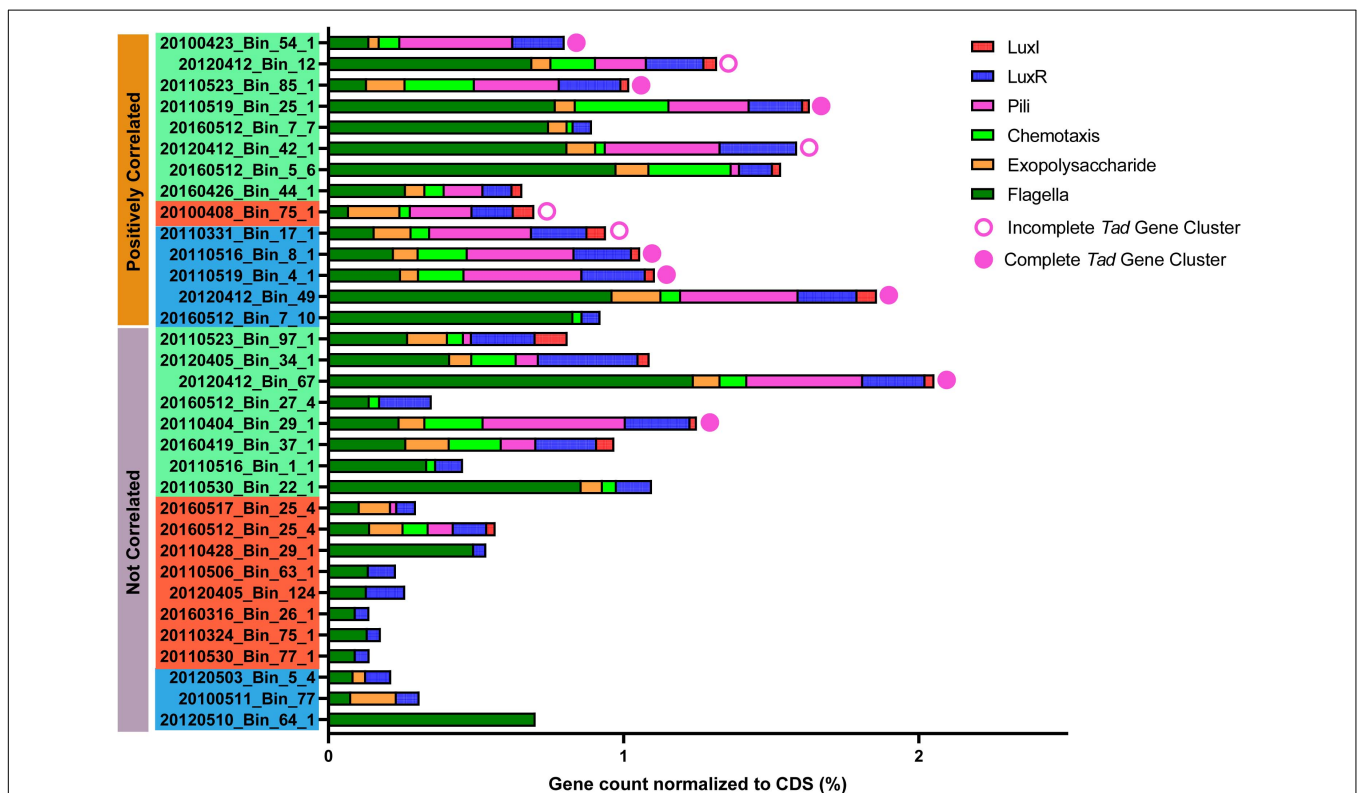

FIGURE 7 | Presence of genes/loci considered important for phycosphere colonization in *Rhodobacterales* MAGs. Gene counts were normalized against the number of coding sequences (CDS) in the respective MAGs. Solid circles indicate a complete *tad* gene locus, open circles indicate near-complete *tad* gene locus (Bins: 20120412_Bin_12, 20120412_Bin_42_1 and 20110331_Bin_17_1 are each missing 1 gene while 2010408_Bin_75_1 is missing 5 genes when compared to MAGs with a complete complement of *tad* genes), while the absence of circles indicate the complete absence of *tad* genes. Positively correlated MAGs are considered putative phycosphere colonizers and non-correlated MAGs are considered putative generalists as established in **Figure 5**. Color coding of MAG names indicate the MAG grouping based in **Figure 3**.

concentrations of metabolites and mobilize accordingly in the direction of the source (Miller et al., 2004; Stocker, 2012; Smriga et al., 2016), in this case, phytoplankton derived DOM. Motility and chemotaxis appear to be prerequisites for attachment and biofilm development in a number of Roseobacters (Miller and Belas, 2006; Bruhn et al., 2007). Thus, chemotaxis and motility enable bacteria to enter the phycosphere in response to algal exudates with some individuals switching from a motile lifestyle to a sessile one by attaching to phytoplankton cells (Geng and Belas, 2010). Given the distribution of chemotaxis proteins among the *Rhodobacterales* MAGs, it is likely that a subset of these MAGs may be able to mobilize toward phytoplankton exudates. The molecular mechanism that underpins the switch from a motile to sessile lifestyle has been attributed to the production of, and response to, QS molecules (Fei et al., 2020). QS molecules, typically AHLs produced by Roseobacters, are known to regulate production of antimicrobials, biofilm formation, motility, virulence and nutrient acquisition (Gram et al., 2002; Hmelo et al., 2011; Buchan et al., 2014; Gardiner et al., 2015). Interestingly, there is a higher prevalence of *luxI* homologs in putative phycosphere colonizers (71.4%) than in generalists (31.6%), while *luxR* homologs are more widespread with many MAGs possessing multiple *luxR* genes. This multiplicity in *luxR* homologs allow for bacteria to respond to a wide range of AHLs from other bacteria (Yao et al., 2006; Janssens et al., 2007), while some studies have noted that these so-called solo-*luxR* genes, that are missing their corresponding *luxI* genes, exhibit loss of QS function (Subramoni and Venturi, 2009). Gardiner et al. (2015) reported that a loss of QS function in Roseobacter group member, *Nautella italica* R11 led to an inability to form biofilms or attach to red macroalga. Thus, a subset of MAGs that present with chemotactic, QS and adherence genes may be well suited to colonizing the phycosphere.

Interestingly, Roseobacter isolates lacking *flp* pilus genes were also reported to lack genes encoding proteins for quorum sensing and secondary metabolism (Voget et al., 2015). It is, therefore, likely that putative generalists that lack both *tad* loci and QS genes are unable to colonize the phycosphere and will presumably use constant motility and chemotaxis to acquire DOM from algal cells on an 'as needed basis'. In contrast, putative phycosphere colonizers should be able to form tight, consistent associations with phytoplankton in the phycosphere using motility and chemotaxis to find phytoplankton cells, then QS, the *tad* gene locus and biofilm formation genes to colonize the phycosphere.

Few flagella related proteins of putative generalist origin were found in the metaproteome, while proteins related to adherence and Flp pili were not detected. This was expected assuming the low relative abundance of putative phycosphere colonizers (**Figures 4**, **5**), the low molecular weight and fragility of pili as well as the general propensity of adhesion related proteins to be in lower abundance than those for motility (Christie-Oleza et al., 2012). There is likely a regulatory link between QS and the *tad* locus, which is presumably a feature that enables specific Roseobacters to attach to algal cells. *Mesorhizobium tianshanense* was shown to possess a LuxR-LuxI type QS system that played a role in adherence to plant root hairs and nodule formation and thus has a critical role in symbiosis (Zheng

et al., 2006), while several studies report that in addition to virulence genes, QS circuits modulate *tad* gene expression in *P. aeruginosa* (Schuster et al., 2003; Wagner et al., 2003; Juhas et al., 2005). Additionally, *in vitro* expression studies have also reported that *Marinobacter adhaerens* HP15, which preferentially attaches to the diatom *Thalassiosira weissflogii* expresses several genes related to chemotaxis and attachment, including Flp pili when co-cultured with the diatom (Torres-Monroy and Ullrich, 2018). The model Roseobacter, *Dinoroseobacter shibae* possesses a QS regulatory network in which *tad* pilus genes are modulated by the perception of AHLs (Koppenhofer et al., 2019). Finally, a survey of the genomes of *Sulfitobacter pseudonitzschiae* F5, *Phaeobacter* sp. F10 and *Alteromonas macleodii* F12 which were isolated from the microbiome of diatom *Asterionellopsis glacialis* revealed that *tad* loci are only present in the two Roseobacter genomes (**Supplementary Figure 7**) but completely absent in the putative opportunist, *A. macleodii* F12, which correlated with the ability of the two Roseobacter species to colonize the phycosphere of this diatom, but not *A. macleodii* F12. Despite the presence of other attachment genes in *A. macleodii* F12, this isolate lacks a complete QS system and thus only the Roseobacters were capable of AHL production, which was shown to modulate their motility and attachment (Fei et al., 2020).

All evidence thus far consistently points toward niche differentiations among algal-associated Roseobacters. The trophic strategy of secreting a higher abundance and diversity of transporter systems can provide an advantage during phytoplankton blooms by enabling utilization of many DOM metabolites (Christie-Oleza et al., 2012). The distribution of Roseobacter transport proteins in the metaproteome supports this as more than 80% of the Roseobacter transport proteins are produced by the putative Roseobacter generalists (**Supplementary Figure 8**), which were MAGs that had consistent higher relative abundance (**Figure 4**). Putative phycosphere colonizers do not exhibit this type of growth pattern as their relative abundances rise and fall with particular diatom species. Given their ability to attach this indicates that physical associations could be an alternate trophic strategy to rapid and abundant growth. Shifts in low abundance MAGs relative abundance patterns points toward potential interactions changing over the course of algal bloom cycles. Indeed, this was observed during dinoflagellate blooms where distinct bacterial taxa colonized algal cells during different phases of the bloom (Mayali et al., 2011). Altogether, these findings suggest that while *Rhodobacterales* generally respond to diatom blooms, only a subset of members are able to form close associations and colonize diatom phycospheres.

CONCLUSION

The dominance of *Rhodobacterales* in marine environments and their apparent interactions with phytoplankton suggest they possess unique metabolic machinery to access phytoplankton exudates. Metagenomics and metaproteomics provide a unique, culture-independent opportunity to investigate microbial interactions by identifying known genes of importance for

interactions with phytoplankton. These tools enable the identification of genes related to biochemical pathways that could be used to compare metabolic and functional gene distribution in microbial communities that occupy different environments or niches (DeLong and Karl, 2005). Investigating these genes and biochemical pathways provide us with a glimpse of how widely distributed these genes and pathways are and how different bacterial lineages may rely on them to interact with phytoplankton. Here we report on comprehensive comparative genomics and metaproteomics of diatom bloom associated members of the order *Rhodobacterales* and provide insight into the functional and metabolic potential of these individuals to colonize diatom phycospheres. Our findings confirm that Roseobacters are among the most important taxa that respond to phytoplankton blooms. More importantly, we show that only a subset of *Rhodobacterales* display strong positive correlations to dominant diatom species during these blooms and that they may use a unique gene cluster, *tad*, in combination with QS, motility and chemotaxis to find and colonize the phycosphere, in strong contrast to putative *Rhodobacterales* generalists that lack the *tad* genes. Future culture-independent work on larger planktonic size fractions, and with Roseobacter isolates, is thus required to investigate the role of these *tad* genes in phycosphere colonization.

DATA AVAILABILITY STATEMENT

Publicly available datasets were analyzed in this study. This data can be found here: The metagenomic dataset PRJEB28156 for this study can be found in the European Nucleotide Archive under accession [https://www.ebi.ac.uk/ena/data/view/PRJEB28156]. The metaproteomic spectrometric data ID PXD008238 for this study can be found in the PRIDE.

AUTHOR CONTRIBUTIONS

AI, RA, and SA conceived the study. BF assembled the metagenomes and binned the MAGs. AI developed and carried out supporting algorithms, bioinformatic analyses and computational pipelines, analyzed the data, and generated figures. AI and SA wrote the manuscript with input from all authors.

FUNDING

ACKNOWLEDGMENTS

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmicb.2021.718297/full#supplementary-material

REFERENCES

Alderkamp, A.-C., Sintes, E., and Herndl, G. J. (2006). Abundance and activity of major groups of prokaryotic plankton in the coastal North Sea during spring and summer. *Aquat. Microb. Ecol.* 45, 237–246. doi: 10.3354/ame045237

Alonso, C., and Pernthaler, J. (2006). *Roseobacter* and SAR11 dominate microbial glucose uptake in coastal North Sea waters. *Environ. Microbiol.* 8, 2022–2030. doi: 10.1111/j.1462-2920.2006.01082.x

Amin, S. A., Hmelo, L. R., van Tol, H. M., Durham, B. P., Carlson, L. T., Heal, K. R., et al. (2015). Interaction and signalling between a cosmopolitan phytoplankton and associated bacteria. *Nature* 522, 98–101. doi: 10.1038/nature14488

Amin, S. A., Parker, M. S., and Armbrust, E. V. (2012). Interactions between diatoms and bacteria. *Microbiol. Mol. Biol. Rev.* 76, 667–684. doi: 10.1128/MMBR.00007-12

Ankenbrand, M. J., and Keller, A. (2016). bcgTree: automatized phylogenetic tree building from bacterial core genomes. *Genome* 59, 783–791. doi: 10.1139/gen-2015-0175

Aramaki, T., Blanc-Mathieu, R., Endo, H., Ohkubo, K., Kanehisa, M., Goto, S., et al. (2020). KofamKOALA: KEGG ortholog assignment based on profile HMM and adaptive score threshold. *Bioinformatics* 36, 2251–2252. doi: 10.1093/bioinformatics/btz859

Arandia-Gorostidi, N., Weber, P. K., Alonso-Saez, L., Moran, X. A., and Mayali, X. (2017). Elevated temperature increases carbon and nitrogen fluxes between phytoplankton and heterotrophic bacteria through physical attachment. *ISME J.* 11, 641–650. doi: 10.1038/ismej.2016.156

Avcı, B., Kruger, K., Fuchs, B. M., Teeling, H., and Amann, R. I. (2020). Polysaccharide niche partitioning of distinct *Polaribacter* clades during North

Sea spring algal blooms. *ISME J.* 14, 1369–1383. doi: 10.1038/s41396-020-0601-y

Azam, F., and Long, R. A. (2001). Oceanography: sea snow microcosms. *Nature* 414, 495–498. doi: 10.1038/35107174

Aziz, R. K., Bartels, D., Best, A. A., DeJongh, M., Disz, T., Edwards, R. A., et al. (2008). The rast server: rapid annotations using subsystems technology. *BMC Genomics* 9:75. doi: 10.1186/1471-2164-9-75

Bartling, P., Vollmers, J., and Petersen, J. (2018). The first world swimming championships of roseobacters-phylogenomic insights into an exceptional motility phenotype. *Syst. Appl. Microbiol.* 41, 544–554. doi: 10.1016/j.syapm.2018.08.012

Behrenfeld, M. J., Boss, E., Siegel, D. A., and Shea, D. M. (2005). Carbon-based ocean productivity and phytoplankton physiology from space. *Glob. Biogeochem. Cycles* 19:GB1006. doi: 10.1029/2004GB002299

Behrenfeld, M. J., O'Malley, R. T., Siegel, D. A., McClain, C. R., Sarmiento, J. L., Feldman, G. C., et al. (2006). Climate-driven trends in contemporary ocean productivity. *Nature* 444, 752–755. doi: 10.1038/nature05317

Behringer, G., Ochsenkuhn, M. A., Fei, C., Fanning, J., Koester, J. A., and Amin, S. A. (2018). Bacterial communities of diatoms display strong conservation across strains and time. *Front. Microbiol.* 9:659. doi: 10.3389/fmicb.2018.00659

Bhattacharjee, M. K., Kachlany, S. C., Fine, D. H., and Figurski, D. H. (2001). Nonspecific adherence and fibril biogenesis by *Actinobacillus actinomycetemcomitans*: TadA protein is an ATPase. *J. Bacteriol.* 183, 5927–5936. doi: 10.1128/JB.183.20.5927-5936.2001

Bidle, K. D., Manganelli, M., and Azam, F. (2002). Regulation of oceanic silicon and carbon preservation by temperature control on bacteria. *Science* 298, 1980–1984. doi: 10.1126/science.1076076

Blin, K., Shaw, S., Steinke, K., Villebro, R., Ziemert, N., Lee, S. Y., et al. (2019). antiSMASH 5.0: updates to the secondary metabolite genome mining pipeline. *Nucleic Acids Res.* 47, W81–W87. doi: 10.1093/nar/gkz310

Bolch, C. J. S., Bejoy, T. A., and Green, D. H. (2017). Bacterial associates modify growth dynamics of the dinoflagellate *Gymnodinium catenatum*. *Front. Microbiol.* 8:670. doi: 10.3389/fmicb.2017.00670

Boyd, J. A., Jungbluth, S. P., Leu, A. O., Evans, P. N., Woodcroft, B. J., Chadwick, G. L., et al. (2019). Divergent methyl-coenzyme M reductase genes in a deep-subseafloor Archaeoglobi. *ISME J.* 13, 1269–1279. doi: 10.1038/s41396-018-0343-2

Brinkhoff, T., Giebel, H. A., and Simon, M. (2008). Diversity, ecology, and genomics of the *Roseobacter* clade: a short overview. *Arch. Microbiol.* 189, 531–539. doi: 10.1007/s00203-008-0353-y

Bruhn, J. B., Gram, L., and Belas, R. (2007). Production of antibacterial compounds and biofilm formation by *Roseobacter* species are influenced by culture conditions. *Appl. Environ. Microbiol.* 73, 442–450. doi: 10.1128/AEM.02238-06

Buchan, A., Gonzalez, J. M., and Moran, M. A. (2005). Overview of the marine *Roseobacter* lineage. *Appl. Environ. Microbiol.* 71, 5665–5677. doi: 10.1128/AEM.71.10.5665-5677.2005

Buchan, A., LeCleir, G. R., Gulvik, C. A., and Gonzalez, J. M. (2014). Master recyclers: features and functions of bacteria associated with phytoplankton blooms. *Nat. Rev. Microbiol.* 12, 686–698. doi: 10.1038/nrmicro3326

Carlson, S. J., Fleig, A., Baron, M. K., Berg, I. A., and Alber, B. E. (2019). Barriers to 3-Hydroxypropionate-dependent growth of *Rhodobacter sphaeroides* by distinct disruptions of the ethylmalonyl coenzyme A pathway. *J. Bacteriol.* 201, e00556–18. doi: 10.1128/JB.00556-18

Case, R. J., Labbate, M., and Kjelleberg, S. (2008). AHL-driven quorum-sensing circuits: their frequency and function among the *Proteobacteria*. *ISME J.* 2, 345–349. doi: 10.1038/ismej.2008.13

Castresana, J. (2000). Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. *Mol. Biol. Evol.* 17, 540–552. doi: 10.1093/oxfordjournals.molbev.a026334

Chafee, M., Fernandez-Guerra, A., Buttigieg, P. L., Gerdts, G., Eren, A. M., Teeling, H., et al. (2018). Recurrent patterns of microdiversity in a temperate coastal marine environment. *ISME J.* 12, 237–252. doi: 10.1038/ismej.2017.165

Christie-Oleza, J. A., Pina-Villalonga, J. M., Bosch, R., Nogales, B., and Armengaud, J. (2012). Comparative proteogenomics of twelve *Roseobacter* exoproteomes reveals different adaptive strategies among these marine bacteria. *Mol. Cell. Proteomics* 11:M111013110. doi: 10.1074/mcp.M111.013110

Clock, S. A., Planet, P. J., Perez, B. A., and Figurski, D. H. (2008). Outer membrane components of the Tad (tight adherence) secreton of *Aggregatibacter actinomycetemcomitans*. *J. Bacteriol.* 190, 980–990. doi: 10.1128/JB.01347-07

Cooper, M. B., Kazamia, E., Helliwell, K. E., Kudahl, U. J., Sayer, A., Wheeler, G. L., et al. (2019). Cross-exchange of B-vitamins underpins a mutualistic interaction between *Ostreococcus tauri* and *Dinoroseobacter shibae*. *ISME J.* 13, 334–345. doi: 10.1038/s41396-018-0274-y

Cotter, P. D., Ross, R. P., and Hill, C. (2013). Bacteriocins—a viable alternative to antibiotics? *Nat. Rev. Microb.* 11, 95–105. doi: 10.1038/nrmicro2937

Croft, M. T., Lawrence, A. D., Raux-Deery, E., Warren, M. J., and Smith, A. G. (2005). Algae acquire vitamin B12 through a symbiotic relationship with bacteria. *Nature* 438, 90–93. doi: 10.1038/nature04056

Cude, W. N., and Buchan, A. (2013). Acyl-homoserine lactone-based quorum sensing in the *Roseobacter* clade: complex cell-to-cell communication controls multiple physiologies. *Front. Microbiol.* 4:336. doi: 10.3389/fmicb.2013.00336

de Bentzmann, S., Aurouze, M., Ball, G., and Filloux, A. (2006). FppA, a novel *Pseudomonas aeruginosa* prepilin peptidase involved in assembly of type IVb pili. *J. Bacteriol.* 188, 4851–4860. doi: 10.1128/JB.00345-06

del Rosario González-Baró, M. A., and Pollero, R. J. (1998). Fatty acid metabolism of *Macrobrachium borellii*: dietary origin of arachidonic and eicosapentaenoic acids. *Comp. Biochem. Physiol. A Mol. Integr. Physiol.* 119, 747–752. doi: 10.1016/S1095-6433(98)01028-9

DeLong, E. F., and Karl, D. M. (2005). Genomic perspectives in microbial oceanography. *Nature* 437, 336–342. doi: 10.1038/nature04157

Durham, B. P., Sharma, S., Luo, H., Smith, C. B., Amin, S. A., Bender, S. J., et al. (2015). Cryptic carbon and sulfur cycling between surface ocean plankton. *Proc. Natl. Acad. Sci. U.S.A.* 112, 453–457. doi: 10.1073/pnas.1413137112

Eddy, S. R. (2009). A new generation of homology search tools based on probabilistic inference. *Genome Inform.* 23, 205–211.

Edgar, R. C. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* 32, 1792–1797. doi: 10.1093/nar/gkh340

Eilers, H., Pernthaler, J., Glöckner, F. O., and Amann, R. (2000). Culturability and in situ abundance of pelagic bacteria from the North Sea. *Appl. Environ. Microbiol.* 66, 3044–3051. doi: 10.1128/AEM.66.7.3044-3051.2000

Eilers, H., Pernthaler, J., Peplies, J., Glöckner, F. O., Gerdts, G., and Amann, R. (2001). Isolation of novel pelagic bacteria from the german bight and their seasonal contributions to surface picoplankton. *Appl. Environ. Microbiol.* 67, 5134–5142. doi: 10.1128/AEM.67.11.5134-5142.2001

Eisma, D. (1987). The North Sea: an overview. *Philos. Trans. R. Soc. B* 316, 461–485. doi: 10.1098/rstb.1987.0032

Eren, A. M., Esen, ÖC., Quince, C., Vineis, J. H., Morrison, H. G., Sogin, M. L., et al. (2015). Anvi'o: an advanced analysis and visualization platform for 'omics data. *PeerJ* 3:e1319. doi: 10.7717/peerj.1319

Fei, C., Ochsenkuhn, M. A., Shibl, A. A., Isaac, A., Wang, C., and Amin, S. A. (2020). Quorum sensing regulates 'swim-or-stick' lifestyle in the phycosphere. *Environ. Microbiol.* 22, 4761–4778. doi: 10.1111/1462-2920.15228

Francis, T. B., Bartosik, D., Sura, T., Sichert, A., Hehemann, J. H., Markert, S., et al. (2021). Changing expression patterns of TonB-dependent transporters suggest shifts in polysaccharide consumption over the course of a spring phytoplankton bloom. *ISME J.* doi: 10.1038/s41396-021-00928-8

Francis, T. B., Krüger, K., Fuchs, B. M., Teeling, H., and Amann, R. I. (2019). Candidatus Prosiliicoccus vernus, a spring phytoplankton bloom associated member of the *Flavobacteriaceae*. *Syst. Appl. Microbiol.* 42, 41–53. doi: 10.1016/j.syapm.2018.08.007

Gardiner, M., Fernandes, N. D., Nowakowski, D., Raftery, M., Kjelleberg, S., Zhong, L., et al. (2015). VarR controls colonization and virulence in the marine macroalgal pathogen *Nautella italica* R11. *Front. Microbiol.* 6:1130. doi: 10.3389/fmicb.2015.01130

Geng, H., and Belas, R. (2010). Molecular mechanisms underlying roseobacter-phytoplankton symbioses. *Curr. Opin. Biotechnol.* 21, 332–338. doi: 10.1016/j.copbio.2010.03.013

Gerdts, G., Wichels, A., Döpke, H., Klings, K.-W., Gunkel, W., and Schütt, C. (2004). 40-year long-term study of microbial parameters near Helgoland (German Bight, North Sea): historical view and future perspectives. *Helgol. Mar. Res.* 58, 230–242. doi: 10.1007/s10152-004-0189-z

Giebel, H.-A., Kalhoefer, D., Gahl-Janssen, R., Choo, Y.-J., Lee, K., Cho, J.-C., et al. (2013). *Planktomarina temperata* gen. nov., sp. nov., belonging to the globally distributed RCA cluster of the marine Roseobacter clade, isolated from the German Wadden Sea. *Int. J. Syst. Evol. Microbiol.* 63, 4207–4217. doi: 10.1099/ijs.0.053249-0

Giebel, H. A., Kalhoefer, D., Lemke, A., Thole, S., Gahl-Janssen, R., Simon, M., et al. (2011). Distribution of Roseobacter RCA and SAR11 lineages in the North Sea and characteristics of an abundant RCA isolate. *ISME J.* 5, 8–19. doi: 10.1038/ismej.2010.87

González, J. M., and Moran, M. A. (1997). Numerical dominance of a group of marine bacteria in the alpha-subclass of the class *Proteobacteria* in coastal seawater. *Appl. Environ. Microbiol.* 63, 4237–4242.

Gonzalez, J., Mayer, F., Moran, M., Hodson, R., and Whitman, W. (1997). *Sagittula stellata* gen. nov., sp. nov., a lignin-transforming bacterium from a coastal environment. *Int. J. Syst. Evol. Microbiol.* 47, 773–780. doi: 10.1099/00207713-47-3-773

Gram, L., Grossart, H.-P., Schlingloff, A., and Kiørboe, T. (2002). Possible quorum sensing in marine snow bacteria: production of acylated homoserine lactones by *Roseobacter* strains isolated from marine snow. *Appl. Environ. Microbiol.* 68, 4111–4116. doi: 10.1128/AEM.68.8.4111-4116.2002

Grant, M. A., Kazamia, E., Cicuta, P., and Smith, A. G. (2014). Direct exchange of vitamin B 12 is demonstrated by modelling the growth dynamics of algal–bacterial cocultures. *ISME J.* 8, 1418–1427. doi: 10.1038/ismej.2014.9

Guasto, J. S., Rusconi, R., and Stocker, R. (2012). Fluid mechanics of planktonic microorganisms. *Annu. Rev. Fluid Mech.* 44, 373–400. doi: 10.1146/annurev-fluid-120710-101156

Hahnke, S., Tindall, B. J., Schumann, P., Sperling, M., Brinkhoff, T., and Simon, M. (2012). *Planktotalea frisia* gen. nov., sp. nov., isolated from the southern North Sea. *Int. J. Syst. Evol. Microbiol.* 62, 1619–1624. doi: 10.1099/ijs.0.033563-0

Hirth, M., Liverani, S., Mahlow, S., Bouget, F.-Y., Pohnert, G., and Sasso, S. (2017). Metabolic profiling identifies trehalose as an abundant and diurnally fluctuating metabolite in the microalga *Ostreococcus tauri*. *Metabolomics* 13:68. doi: 10.1007/s11306-017-1203-1

Hmelo, L. R., Mincer, T. J., and Van Mooy, B. A. (2011). Possible influence of bacterial quorum sensing on the hydrolysis of sinking particulate organic carbon in marine environments. *Environ. Microbiol. Rep.* 3, 682–688. doi: 10.1111/j.1758-2229.2011.00281.x

Hogle, S. L., Thrash, J. C., Dupont, C. L., and Barbeau, K. A. (2016). Trace metal acquisition by marine heterotrophic bacterioplankton with contrasting trophic strategies. *Appl. Environ. Microbiol.* 82, 1613–1624. doi: 10.1128/AEM.03128-15

Huerta-Cepas, J., Serra, F., and Bork, P. (2016). ETE 3: reconstruction, analysis, and visualization of phylogenomic data. *Mol. Biol. Evol.* 33, 1635–1638. doi: 10.1093/molbev/msw046

Ittekkot, V., Brockmann, U., Michaelis, W., and Degens, E. T. (1981). Dissolved free and combined carbohydrates during a phytoplankton bloom in the northern North Sea. *Mar. Ecol. Prog. Ser.* 4, 299–305.

Janssens, J. C., Metzger, K., Daniels, R., Ptacek, D., Verhoeven, T., Habel, L. W., et al. (2007). Synthesis of N-acyl homoserine lactone analogues reveals strong activators of SdiA, the *Salmonella enterica* serovar *Typhimurium* LuxR homologue. *Appl. Environ. Microbiol.* 73, 535–544. doi: 10.1128/AEM.01451-06

Johansson, O. N., Pinder, M. I. M., Ohlsson, F., Egardt, J., Topel, M., and Clarke, A. K. (2019). Friends with benefits: exploring the phycosphere of the marine diatom *Skeletonema marinoi*. *Front. Microbiol.* 10:1828. doi: 10.3389/fmicb.2019.01828

Juhas, M., Wiehlmann, L., Salunkhe, P., Lauber, J., Buer, J., and Tümmler, B. (2005). GeneChip expression analysis of the VqsR regulon of *Pseudomonas aeruginosa* TB. *FEMS Microbiol. Lett.* 242, 287–295. doi: 10.1016/j.femsle.2004.11.020

Kachlany, S. C., Planet, P. J., Bhattacharjee, M. K., Kollia, E., DeSalle, R., Fine, D. H., et al. (2000). Nonspecific adherence by *Actinobacillus actinomycetemcomitans* requires genes widespread in bacteria and archaea. *J. Bacteriol.* 182, 6169–6176. doi: 10.1128/JB.182.21.6169-6176.2000

Kachlany, S. C., Planet, P. J., DeSalle, R., Fine, D. H., and Figurski, D. H. (2001). Genes for tight adherence of *Actinobacillus actinomycetemcomitans*: from plague to pond scum. *Trends Microbiol.* 9, 429–437. doi: 10.1016/S0966-842X(01)02161-8

Kanehisa, M., Sato, Y., Kawashima, M., Furumichi, M., and Tanabe, M. (2016). KEGG as a reference resource for gene and protein annotation. *Nucleic Acids Res.* 44, D457–D462. doi: 10.1093/nar/gkv1070

Kappelmann, L., Krüger, K., Hehemann, J. H., Harder, J., Markert, S., Unfried, F., et al. (2019). Polysaccharide utilization loci of North Sea *Flavobacteriia* as basis for using SusC/D-protein expression for predicting major phytoplankton glycans. *ISME J.* 13, 76–91. doi: 10.1038/s41396-018-0242-6

Konstantinidis, K. T., and Tiedje, J. M. (2004). Trends between gene content and genome size in prokaryotic species with larger genomes. *Proc. Natl. Acad. Sci. U.S.A.* 101, 3160–3165. doi: 10.1073/pnas.0308653100

Koppenhofer, S., Wang, H., Scharfe, M., Kaever, V., Wagner-Dobler, I., and Tomasch, J. (2019). Integrated transcriptional regulatory network of quorum sensing, replication control, and SOS response in *Dinoroseobacter shibae*. *Front. Microbiol.* 10:803. doi: 10.3389/fmicb.2019.00803

Krüger, K., Chafee, M., Ben Francis, T., Glavina Del Rio, T., Becher, D., Schweder, T., et al. (2019). In marine *Bacteroidetes* the bulk of glycan degradation during algae blooms is mediated by few clades using a restricted set of genes. *ISME J.* 13, 2800–2816. doi: 10.1038/s41396-019-0476-y

Letunic, I., and Bork, P. (2019). Interactive Tree Of Life (iTOL) v4: recent updates and new developments. *Nucleic Acids Res.* 47, W256–W259. doi: 10.1093/nar/gkz239

Luo, H., and Moran, M. A. (2014). Evolutionary ecology of the marine *Roseobacter* clade. *Microbiol. Mol. Biol. Rev.* 78, 573–587. doi: 10.1128/MMBR.00020-14

Luo, H., Loytynoja, A., and Moran, M. A. (2012). Genome content of uncultivated marine *Roseobacters* in the surface ocean. *Environ. Microbiol.* 14, 41–51. doi: 10.1111/j.1462-2920.2011.02528.x

Luo, H., Swan, B. K., Stepanauskas, R., Hughes, A. L., and Moran, M. A. (2014). Evolutionary analysis of a streamlined lineage of surface ocean Roseobacters. *ISME J.* 8, 1428–1439. doi: 10.1038/ismej.2013.248

Madeira, F., Park, Y. M., Lee, J., Buso, N., Gur, T., Madhusoodanan, N., et al. (2019). The EMBL-EBI search and sequence analysis tools APIs in 2019. *Nucleic Acids Res.* 47, W636–W641. doi: 10.1093/nar/gkz268

Maharjan, R. P., Yu, P.-L., Seeto, S., and Ferenci, T. (2005). The role of isocitrate lyase and the glyoxylate cycle in *Escherichia coli* growing under glucose limitation. *Res. Microbiol.* 156, 178–183. doi: 10.1016/j.resmic.2004.09.004

Majzoub, M. E., Beyersmann, P. G., Simon, M., Thomas, T., Brinkhoff, T., and Egan, S. (2019). *Phaeobacter inhibens* controls bacterial community assembly on a marine diatom. *FEMS Microbiol. Ecol.* 95:fiz060. doi: 10.1093/femsec/fiz060

Mayali, X., Franks, P. J. S., and Burton, R. S. (2011). Temporal attachment dynamics by distinct bacterial taxa during a dinoflagellate bloom. *Aquat. Microb. Ecol.* 63, 111–122. doi: 10.3354/ame01483

Miller, T. R., and Belas, R. (2006). Motility is involved in *Silicibacter* sp. TM1040 interaction with dinoflagellates. *Environ. Microbiol.* 8, 1648–1659. doi: 10.1111/j.1462-2920.2006.01071.x

Miller, T. R., Hnilicka, K., Dziedzic, A., Desplats, P., and Belas, R. (2004). Chemotaxis of *Silicibacter* sp. strain TM1040 toward dinoflagellate products. *Appl. Environ. Microbiol.* 70, 4692–4701. doi: 10.1128/AEM.70.8.4692-4701.2004

Moran, M. A., Belas, R., Schell, M. A., Gonzalez, J. M., Sun, F., Sun, S., et al. (2007). Ecological genomics of marine Roseobacters. *Appl. Environ. Microbiol.* 73, 4559–4569. doi: 10.1128/AEM.02580-06

Moran, M. A., Buchan, A., González, J. M., Heidelberg, J. F., Whitman, W. B., Kiene, R. P., et al. (2004). Genome sequence of *Silicibacter pomeroyi* reveals adaptations to the marine environment. *Nature* 432:910. doi: 10.1038/nature03170

Mungan, M. D., Alanjary, M., Blin, K., Weber, T., Medema, M. H., and Ziemert, N. (2020). ARTS 2.0: feature updates and expansion of the antibiotic resistant target seeker for comparative genome mining. *Nucleic Acids Res.* 48, W546–W552. doi: 10.1093/nar/gkaa374

Needham, D. M., and Fuhrman, J. A. (2016). Pronounced daily succession of phytoplankton, archaea and bacteria following a spring bloom. *Nat. Microbiol.* 1:16005. doi: 10.1038/nmicrobiol.2016.5

Needham, D. M., Sachdeva, R., and Fuhrman, J. A. (2017). Ecological dynamics and co-occurrence among marine phytoplankton, bacteria and myoviruses shows microdiversity matters. *ISME J.* 11, 1614–1629. doi: 10.1038/ismej.2017.29

Newton, R. J., Griffin, L. E., Bowles, K. M., Meile, C., Gifford, S., Givens, C. E., et al. (2010). Genome characteristics of a generalist marine bacterial lineage. *ISME J.* 4, 784–798. doi: 10.1038/ismej.2009.150

Oksanen, J., Blanchet, F. G., Friendly, M., Kindt, R., Legendre, P., McGlinn, D., et al. (2020). *Vegan: Community Ecology Package. R Package Version 2.5-7*. Available online at: https://CRAN.R-project.org/package=vegan (accessed May 13, 2021).

Overbeek, R., Olson, R., Pusch, G. D., Olsen, G. J., Davis, J. J., Disz, T., et al. (2014). The SEED and the rapid annotation of microbial genomes using subsystems technology (RAST). *Nucleic Acids Res.* 42, D206–D214. doi: 10.1093/nar/gkt1226

Pienko, T., and Trylska, J. (2020). Extracellular loops of BtuB facilitate transport of vitamin B12 through the outer membrane of *E. coli*. *PLoS Comput. Biol.* 16:e1008024. doi: 10.1371/journal.pcbi.1008024

Planet, P. J., Kachlany, S. C., Fine, D. H., DeSalle, R., and Figurski, D. H. (2003). The widespread colonization island of *Actinobacillus actinomycetemcomitans*. *Nat. Genet.* 34, 193–198. doi: 10.1038/ng1154

Pu, M., Duriez, P., Arazi, M., and Rowe-Magnus, D. A. (2018). A conserved tad pilus promotes *Vibrio vulnificus* oyster colonization. *Environ. Microbiol.* 20, 828–841. doi: 10.1111/1462-2920.14025

Pujalte, M. J., Lucena, T., Ruvira, M. A., Arahal, D. R., and Macián, M. C. (2014). "The family Rhodobacteraceae," in *The Prokaryotes*, eds E. Rosenberg, E. F. DeLong, S. Lory, E. Stackebrandt, and F. Thompson (Berlin: Springer), 439–512.

Raina, J.-B., Fernandez, V., Lambert, B., Stocker, R., and Seymour, J. R. (2019). The role of microbial motility and chemotaxis in symbiosis. *Nat. Rev. Microbiol.* 17, 284–294. doi: 10.1038/s41579-019-0182-9

Ridley, C. P., Lee, H. Y., and Khosla, C. (2008). Evolution of polyketide synthases in bacteria. *Proc. Natl. Acad. Sci. U.S.A.* 105, 4595–4600. doi: 10.1073/pnas.0710107105

Rink, B., Seeberger, S., Martens, T., Duerselen, C.-D., Simon, M., and Brinkhoff, T. (2007). Effects of phytoplankton bloom in a coastal ecosystem on the composition of bacterial communities. *Aquat. Microb. Ecol.* 48, 47–60. doi: 10.3354/ame048047

Robinson, S. L., Christenson, J. K., and Wackett, L. P. (2019). Biosynthesis and chemical diversity of β-lactone natural products. *Nat. Prod. Rep.* 36, 458–475. doi: 10.1039/C8NP00052B

Salazar, G., Paoli, L., Alberti, A., Huerta-Cepas, J., Ruscheweyh, H. J., Cuenca, M., et al. (2019). Gene expression changes and community turnover differentially

shape the global ocean metatranscriptome. *Cell* 179, 1068–1083.e21. doi: 10.1016/j.cell.2019.10.014

Samo, T. J., Kimbrel, J. A., Nilson, D. J., Pett-Ridge, J., Weber, P. K., and Mayali, X. (2018). Attachment between heterotrophic bacteria and microalgae influences symbiotic microscale interactions. *Environ. Microbiol.* 20, 4385–4400. doi: 10.1111/1462-2920.14357

Sangermani, M., Hug, I., Sauter, N., Pfohl, T., and Jenal, U. (2019). Tad pili play a dynamic role in *Caulobacter crescentus* surface colonization. *mBio* 10:e01237–19. doi: 10.1128/mBio.01237-19

Sanudo-Wilhelmy, S. A., Gomez-Consarnau, L., Suffridge, C., and Webb, E. A. (2014). The role of B vitamins in marine biogeochemistry. *Ann. Rev. Mar. Sci.* 6, 339–367. doi: 10.1146/annurev-marine-120710-100912

Sasso, S., Pohnert, G., Lohr, M., Mittag, M., and Hertweck, C. (2012). Microalgae in the postgenomic era: a blooming reservoir for new natural products. *FEMS Microbiol. Rev.* 36, 761–785. doi: 10.1111/j.1574-6976.2011.00304.x

Schuster, M., Lostroh, C. P., Ogi, T., and Greenberg, E. P. (2003). Identification, timing, and signal specificity of *Pseudomonas aeruginosa* quorum-controlled genes: a transcriptome analysis. *J. Bacteriol.* 185, 2066–2079. doi: 10.1128/JB.185.7.2066-2079.2003

Seymour, J. R., Amin, S. A., Raina, J. B., and Stocker, R. (2017). Zooming in on the phycosphere: the ecological interface for phytoplankton-bacteria relationships. *Nat. Microbiol.* 2:17065. doi: 10.1038/nmicrobiol.2017.65

Seymour, J., Ahmed, T., Durham, W., and Stocker, R. (2010). Chemotactic response of marine bacteria to the extracellular products of *Synechococcus* and *Prochlorococcus*. *Aquat. Microb. Ecol.* 59, 161–168. doi: 10.3354/ame01400

Shibl, A. A., Isaac, A., Ochsenkühn, M. A., Cárdenas, A., Fei, C., Behringer, G., et al. (2020). Diatom modulation of select bacteria through use of two unique secondary metabolites. *Proc. Natl. Acad. Sci. U.S.A.* 117, 27445–27455. doi: 10.1073/pnas.2012088117

Simon, M., Scheuner, C., Meier-Kolthoff, J. P., Brinkhoff, T., Wagner-Dobler, I., Ulbrich, M., et al. (2017). Phylogenomics of *Rhodobacteraceae* reveals evolutionary adaptation to marine and non-marine habitats. *ISME J.* 11, 1483–1499. doi: 10.1038/ismej.2016.198

Slightom, R. N., and Buchan, A. (2009). Surface colonization by marine roseobacters: integrating genotype and phenotype. *Appl. Environ. Microbiol.* 75, 6027–6037. doi: 10.1128/AEM.01508-09

Smith, D. C., Simon, M., Alldredge, A. L., and Azam, F. (1992). Intense hydrolytic enzyme activity on marine aggregates and implications for rapid particle dissolution. *Nature* 359:139. doi: 10.1038/359139a0

Smriga, S., Fernandez, V. I., Mitchell, J. G., and Stocker, R. (2016). Chemotaxis toward phytoplankton drives organic matter partitioning among marine bacteria. *Proc. Natl. Acad. Sci. U.S.A.* 113, 1576–1581. doi: 10.1073/pnas.1512307113

Stamatakis, A. (2014). RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* 30, 1312–1313. doi: 10.1093/bioinformatics/btu033

Stocker, R. (2012). Marine microbes see a sea of gradients. *Science* 338, 628–633. doi: 10.1126/science.1208929

Subramoni, S., and Venturi, V. (2009). LuxR-family 'solos': bachelor sensors/regulators of signalling molecules. *Microbiology* 155, 1377–1385. doi: 10.1099/mic.0.026849-0

Taylor, P. K., Zhang, L., and Mah, T. F. (2019). Loss of the two-component system TctD-TctE in *Pseudomonas aeruginosa* affects biofilm formation and aminoglycoside susceptibility in response to citric acid. *mSphere* 4:e00102–19. doi: 10.1128/mSphere.00102-19

Teeling, H., Fuchs, B. M., Becher, D., Klockow, C., Gardebrecht, A., Bennke, C. M., et al. (2012). Substrate-controlled succession of marine bacterioplankton populations induced by a phytoplankton bloom. *Science* 336, 608–611. doi: 10.1126/science.1218344

Teeling, H., Fuchs, B. M., Bennke, C. M., Krüger, K., Chafee, M., Kappelmann, L., et al. (2016). Recurring patterns in bacterioplankton dynamics during coastal spring algae blooms. *Elife* 5:e11888. doi: 10.7554/eLife.11888

Tomich, M., Planet, P. J., and Figurski, D. H. (2007). The tad locus: postcards from the widespread colonization island. *Nat. Rev. Microbiol.* 5, 363–375. doi: 10.1038/nrmicro1636

Torres-Monroy, I., and Ullrich, M. S. (2018). Identification of bacterial genes expressed during diatom-bacteria interactions using an in vivo expression technology approach. *Front. Mar. Sci.* 5:200. doi: 10.3389/fmars.2018.00200

Villar, E., Vannier, T., Vernette, C., Lescot, M., Cuenca, M., Alexandre, A., et al. (2018). The ocean gene atlas: exploring the biogeography of plankton genes online. *Nucleic Acids Res.* 46, W289–W295. doi: 10.1093/nar/gky376

Vizcaíno, J. A., Csordas, A., Del-Toro, N., Dianes, J. A., Griss, J., Lavidas, I., et al. (2016). 2016 update of the PRIDE database and its related tools. *Nucleic Acids Res.* 44, D447–D456. doi: 10.1093/nar/gkv1145

Voget, S., Wemheuer, B., Brinkhoff, T., Vollmers, J., Dietrich, S., Giebel, H. A., et al. (2015). Adaptation of an abundant *Roseobacter* RCA organism to pelagic systems revealed by genomic and transcriptomic analyses. *ISME J.* 9, 371–384. doi: 10.1038/ismej.2014.134

Wagner, V. E., Bushnell, D., Passador, L., Brooks, A. I., and Iglewski, B. H. (2003). Microarray analysis of *Pseudomonas aeruginosa* quorum-sensing regulons: effects of growth phase and environment. *J. Bacteriol.* 185, 2080–2095. doi: 10.1128/JB.185.7.2080-2095.2003

Wang, X., Zhu, J., Feng, J., Sardar, A., Hu, Z., and Wang, H. (2019). Tropicibacter alexandrii sp. nov., a novel marine bacterium isolated from the phycosphere of a dinoflagellate, *Alexandrium minutum*. *Antonie Van Leeuwenhoek* 113, 311–320. doi: 10.1007/s10482-019-01339-8

Wienhausen, G., Noriega-Ortega, B. E., Niggemann, J., Dittmar, T., and Simon, M. (2017). The exometabolome of two model strains of the *Roseobacter* group: a marketplace of microbial metabolites. *Front. Microbiol.* 8:1985. doi: 10.3389/fmicb.2017.01985

Williams, T. J., Wilkins, D., Long, E., Evans, F., DeMaere, M. Z., Raftery, M. J., et al. (2013). The role of planktonic *Flavobacteria* in processing algal organic matter in coastal East Antarctica revealed using metagenomics and metaproteomics. *Environ. Microbiol.* 15, 1302–1317. doi: 10.1111/1462-2920.12017

Xing, P., Hahnke, R.L., Unfried, F., Markert, S., Huang, S., Barbeyron, T., et al. (2015). Niches of two polysaccharide-degrading polaribacter isolates from the North Sea during a spring diatom bloom. *ISME J.* 9, 1410–1422. doi: 10.1038/ismej.2014.225

Yao, Y., Martinez-Yamout, M. A., Dickerson, T. J., Brogan, A. P., Wright, P. E., and Dyson, H. J. (2006). Structure of the *Escherichia coli* quorum sensing protein SdiA: activation of the folding switch by acyl homoserine lactones. *J. Mol. Biol.* 355, 262–273. doi: 10.1016/j.jmb.2005.10.041

Zhang, H., Hou, F., Xie, W., Wang, K., Zhou, X., Zhang, D., et al. (2020). Interaction and assembly processes of abundant and rare microbial communities during a diatom bloom process. *Environ. Microbiol.* 22, 1707–1719. doi: 10.1111/1462-2920.14820

Zheng, H., Zhong, Z., Lai, X., Chen, W. X., Li, S., and Zhu, J. (2006). A LuxR/LuxI-type quorum-sensing system in a plant bacterium, *Mesorhizobium tianshanense*, controls symbiotic nodulation. *J. Bacteriol.* 188, 1943–1949. doi: 10.1128/JB.188.5.1943-1949.2006

Zubkov, M. V., Fuchs, B. M., Archer, S. D., Kiene, R. P., Amann, R., and Burkill, P. H. (2001). Linking the composition of bacterioplankton to rapid turnover of dissolved dimethylsulphoniopropionate in an algal bloom in the North Sea. *Environ. Microbiol.* 3, 304–311. doi: 10.1046/j.1462-2920.2001.00196.x