



# Coordinated gene expression between *Trichodesmium* and its microbiome over day–night cycles in the North Pacific Subtropical Gyre

Kyle R Frischkorn<sup>1,2</sup> · Sheean T Haley<sup>2</sup> · Sonya T Dyhrman<sup>1,2</sup>

Received: 3 June 2017 / Revised: 3 October 2017 / Accepted: 16 December 2017 / Published online: 30 January 2018  
© International Society for Microbial Ecology 2018

## Abstract

*Trichodesmium* is a widespread, N<sub>2</sub> fixing marine cyanobacterium that drives inputs of newly fixed nitrogen and carbon into the oligotrophic ecosystems where it occurs. Colonies of *Trichodesmium* ubiquitously occur with heterotrophic bacteria that make up a diverse microbiome, and interactions within this *Trichodesmium* holobiont could influence the fate of fixed carbon and nitrogen. Metatranscriptome sequencing was performed on *Trichodesmium* colonies collected during high-frequency Lagrangian sampling in the North Pacific Subtropical Gyre (NPSG) to identify possible interactions between the *Trichodesmium* host and microbiome over day–night cycles. Here we show significantly coordinated patterns of gene expression between host and microbiome, many of which had significant day–night periodicity. The functions of the co-expressed genes suggested a suite of interactions within the holobiont linked to key resources including nitrogen, carbon, and iron. Evidence of microbiome reliance on *Trichodesmium*-derived vitamin B12 was also detected in co-expression patterns, highlighting a dependency that could shape holobiont community structure. Collectively, these patterns of expression suggest that biotic interactions could influence colony cycling of resources like nitrogen and vitamin B12, and decouple activities, like N<sub>2</sub> fixation, from typical abiotic drivers of *Trichodesmium* physiological ecology.

## Introduction

The cyanobacterium *Trichodesmium* is biogeochemically important in nutrient-poor oligotrophic ocean ecosystems because of its ability to supply fixed nitrogen to the water column [1]. Through biological N<sub>2</sub> fixation, *Trichodesmium* is estimated to supply approximately half of the total fixed nitrogen in the ocean [1, 2]. As such, *Trichodesmium* plays a keystone role in fueling marine primary production that is otherwise limited by the availability of nitrogen across much of the oligotrophic ocean gyres [3, 4]. The dissolved organic matter produced through this primary production is one of the

largest pools of reduced carbon on earth, and its fate in the ocean is hypothesized to be greatly influenced by microbe–microbe interactions [5]. These interactions can span entire microbial communities and are subject to environmental and evolutionary pressures that work in tandem to influence ecology and subsequent biogeochemical impact [6]. In marine microbial communities, photosynthetic and heterotrophic bacteria have been shown to exhibit synchronous periodic gene expression patterns over day–night cycles that are attributed to cascades in metabolic activity sparked by primary productivity [7–9]. The taxonomic patterns of these responses in free-living communities were similar between coastal and open ocean systems, suggesting conserved ecological interactions and synergistic activities that could be used to inform or predict biogeochemical transformations [7].

The presence of bacteria on *Trichodesmium* filaments was first noted in the 1980s [10], and it is now appreciated that *Trichodesmium* colonies ubiquitously occur with these heterotrophic bacterial epibionts [11–14]. This microbiome is distinct from the bacteria found free living in the water column [15], and varies across ocean basins [14]. Such tight associations between photosynthetic diazotroph and heterotrophic bacteria, in conjunction with detection of

---

**Electronic supplementary material** The online version of this article (https://doi.org/10.1038/s41396-017-0041-5) contains supplementary material, which is available to authorized users.

✉ Sonya T Dyhrman  
sdyhrman@ldeo.columbia.edu

<sup>1</sup> Department of Earth and Environmental Sciences, Columbia University, New York, NY 10025, USA

<sup>2</sup> Lamont-Doherty Earth Observatory, Palisades, NY 10964, USA

microenvironments of depleted oxygen within colonies, led to the hypothesis that these interactions could facilitate exchange of organic carbon, fixed N<sub>2</sub>, iron, and vitamins [10]. Metagenomic analyses of *Trichodesmium* consortia have begun to elucidate the functional underpinnings of the relationship between host and microbiome, with the abundance and diversity of microbiome genes related to phosphorus acquisition, iron cycling, carbon metabolism, and vitamin B12 transport further suggesting that the exchange of critical resources between host and microbiome could occur [11, 13].

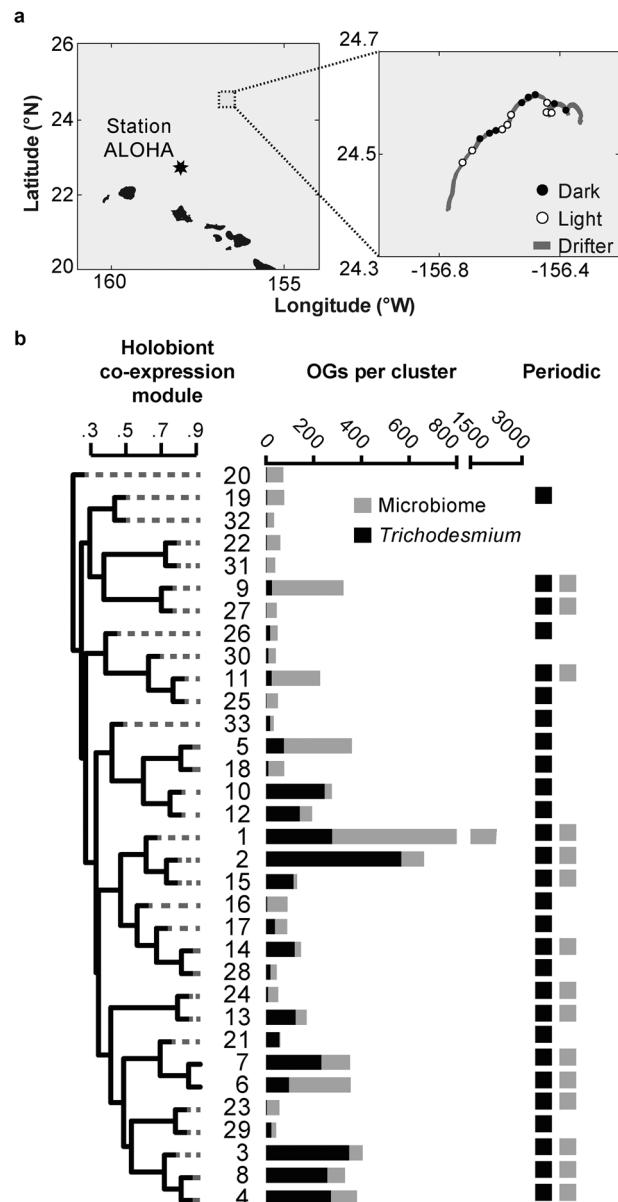
The functional potential recovered from *Trichodesmium* consortia indicates a suite of mechanisms that might underlie previously observed dynamics between host and microbiome. For example, selective modulation of the microbiome community with quorum sensing molecules, which *Trichodesmium* does not have receptors for, resulted in increased activity of the organic phosphate hydrolyzing enzyme alkaline phosphatase within colonies [16]. This finding suggests that the interactions and activities of the microbiome could modulate phosphorus cycling in colonies, and underscores gene expression analyses that indicate a phosphorus-limited microenvironment within colonies may be decoupled from the surrounding water column [17]. The dynamics of how the other metabolic functions observed in metagenomes are partitioned and expressed between *Trichodesmium* and the microbiome may yield insight into the nature of this relationship, helping to determine whether epibionts are important members of these consortia or simply stochastic colonizers.

New evidence of conserved relationships between *Trichodesmium* and specific epibionts across ocean basins [13, 14] suggest that these host–microbiome interactions are widely conserved and could be used to inform or predict biogeochemical transformations, as has been demonstrated for free-living communities [7]. For example, synchronization of carbon and nitrogen metabolic processes between *Trichodesmium* and the microbiome could alter the flux of limiting resources like nitrogen from the colony to the water column, and ultimately influence the primary production and carbon cycling these resources support. Here we used metatranscriptomic sequencing of *Trichodesmium* colonies collected during high-frequency Lagrangian sampling over day–night transitions in the NPSG to identify the coordinated patterns in host and microbiome gene expression that underpin interactions within the holobiont.

## Materials and methods

### Sample collection

Sampling was carried out between 27 July and 30 July 2015 in the NPSG (Fig. 1a). *Trichodesmium* samples were



**Fig. 1** Sampling location and holobiont co-expression modules. **a** Lagrangian sampling was carried out Northeast of Station ALOHA in the North Pacific Subtropical Gyre with an approximate location indicated by a dashed box in the left panel, with an expanded view in the right panel. Sampling followed a drifter track (gray line) with dots demarcating where samples were taken and the light conditions at the time of collection, either night (filled symbols) or day (open symbols). **b** A dendrogram showing the 33 co-expression modules generated from *Trichodesmium* and microbiome metatranscriptomes. Horizontal bars indicate the number of orthologous groups (OGs) per module, each of which contained representatives from *Trichodesmium* and microbiome subsets. Squares on the far right indicate the presence of OGs with significantly periodic expression as determined by RAIN analysis [29]

obtained with six hauls of a 130 µm mesh size net tow through surface seawater every 4 h starting at approximately 1:30 p.m. on 27 July. *Trichodesmium* colonies were isolated by hand and washed of non-tightly associated

microorganisms by serial transfer through 0.2 µm sterile-filtered surface seawater after Frischkorn et al. [11]. Studies using a similar approach for colony collection have shown that the dominant heterotrophic bacterial species tightly associated with *Trichodesmium* colonies are conserved within, and to some extent across, ocean basins [13, 14]. Therefore, this approach should yield the stable, tightly associated microbiome, limiting the presence of opportunistic colonizers, which would contribute to colony function in a transient way. An average of 14 *Trichodesmium* colonies were isolated and preserved per time point. Sampling was performed under red light during the time points at night. Colonies were cleaned within roughly 15 min of collection and immediately filtered by gentle vacuum onto 5 µm, 47 mm polycarbonate filters and stored in liquid nitrogen until RNA extraction.

### RNA extraction and sequencing

Prokaryotic RNA was extracted with the Qiagen RNeasy Mini Kit (Qiagen, Hildern, Germany) with a minor modification to the lysis step. Briefly, ~500 µl zirconia/silica beads (0.5 mm) were added to each sample tube after the addition of Buffer RLT, and the samples were vortexed for 5 min. The resulting lysate was processed as per the remainder of the manufacturer's instructions, including on-column DNase digestion (RNase-free DNase Kit, Qiagen). Prokaryotic RNA was enriched in the eluted total RNA with a MICROBEnrich Kit (ThermoFisher Scientific, Waltham, MA, USA) following the manufacturer's instructions. Ribosomal RNA was removed using the Ribo-Zero Magnetic Kit for bacteria (Illumina, San Diego, CA, USA) as per the manufacturer's instructions. Purified prokaryotic mRNA was concentrated using the RNeasy MinElute Cleanup Kit (Qiagen) according to the directions provided by the manufacturer. The mRNA concentration and quality was assessed with a BioAnalyzer using the RNA 6000 Nano Kit (Agilent Technologies, Santa Clara, CA, USA). The Illumina TruSeq RNA sample preparation kit was used by the JP Sulzberger Genome Center at Columbia University (CUGC). Sequencing of 60 million paired end reads from each sample was performed on an Illumina HiSeq at the CUGC. Sequences were deposited in the NCBI SRA under accession number PRJNA381915.

### Metatranscriptome analysis

Sequenced reads were trimmed, normalized, and assembled following the Eel Pond Protocol for mRNASeq [18], yielding 1,809,695 contigs total from across the 16 time points. To create one master assembly to which reads from all 16 time points could be mapped, individual assemblies were pooled together and clustered at 98% identity using

CD-Hit to combine highly similar sequences across the samples [19–22]. This clustering yielded 1,247,416 contigs total. The merged assembly was then filtered to remove sequences shorter than 210 nucleotides and translated into corresponding amino acid sequences using Prodigal's metagenomics setting [23].

Taxonomic affiliation of contigs into the *Trichodesmium* and microbiome subsets was determined using DIAMOND [24] against the NCBI nr database and analyzed using MEGAN5 software [25]. Functional annotations were obtained by DIAMOND against the UniRef90 database [26] as well as the Kyoto Encyclopedia of Genes and Genomes (KEGG) with the online Automatic Annotation Server using the single-directional best-hit method targeted to prokaryotes and with the metagenomic option selected. Consensus annotations for orthologous groups (OGs) were determined by taking the most abundant UniRef or KEGG annotation for all proteins within that group. In some instances, similar KEGG submodule categories were combined to simplify figures. The amino acid metabolism category includes arginine and proline metabolism, aromatic amino acid metabolism, branched-chain amino acid metabolism, cysteine and methionine metabolism, histidine metabolism, and serine and threonine metabolism KEGG submodules. The carbohydrate metabolism category includes central carbohydrate metabolism and other carbohydrate metabolism KEGG submodules. The sugar transport and metabolism includes saccharide and polyol transport system and sugar metabolism KEGG submodules. The transport category includes ABC-2 type and other transport systems, metallic cation, iron siderophore and vitamin B12 transport system, mineral and organic ion transport system, and peptide and nickel transport system KEGG submodules. Taxonomic identity of the most abundant proteins making up each OG is reported in the Supplementary Dataset and represents the most abundant UniRef annotation taxonomy for that group.

Read mapping to clustered and size-filtered contigs was carried out with RSEM and the default settings with the exception of using the paired end option and the bowtie2 option [27]. On average, over 98% of reads from each time point mapped to this clustered assembly. Read counts were summed across OGs separately for *Trichodesmium* and heterotrophic bacterial epibiont-identified contigs. The OGs were generated by performing a reciprocal comparison with DIAMOND followed by MCL (Markov cluster algorithm) set to an inflation parameter of 1.4, as described elsewhere [11, 28], yielding 462,229 OGs total.

To keep downstream analyses conservative, only those OGs with read coverage of >100 and 200 reads total for the *Trichodesmium* and microbiome subsets, respectively, across all time points were used in downstream analyses. This approach is consistent with previous work in other

systems [9, 29]. With this conservative approach, a total of 3,188 *Trichodesmium* OGs and 4,827 microbiome OGs recruited sufficient read coverage across the 16 samples for subsequent analyses (Supplementary Dataset). Read counts of these abundant OGs in the *Trichodesmium* and microbiome subsets were normalized using the Variance Stabilizing Transformation (VST) in DESeq [30]. Significant periodicity in normalized OG expression was determined using Rhythmicity Analysis Incorporating Non-parametric Methods (RAIN) [29] in R. OGs with  $p$  values  $<0.1$  after false-discovery rate correction [31] were considered to have significant periodicity after Wilson et al. [9].

To cluster consortia OGs into co-expression modules, counts from *Trichodesmium* and microbiome subsets were first pooled. The combined subsets were then normalized as a whole using VST in DESeq [30]. Normalized read counts were then clustered using the R package weighted correlation network analysis (WGCNA) [32] with a soft threshold of 6 selected after a scale-free network topology test and the “blockwiseModules” command set with a minimum module size of 30 OGs and a cut height of 0.25 after Wilson et al. [9]. A simplified visualization of the cluster dendrogram produced by WGCNA was generated by hierarchically clustering the pooled, normalized read counts in each module using the hclust command in R.

## Results and discussion

### Coordinated expression in the *Trichodesmium* holobiont

High-frequency Lagrangian sampling of *Trichodesmium* colonies generated a total of 16 metatranscriptome samples spread over 3 day–night transitions in the NPSG (Fig. 1a). After mapping to a merged de novo assembly, read counts were partitioned between *Trichodesmium* and all heterotrophic bacteria OGs as determined by MEGAN [25] and summed for each OG. The microbiome OG subset was dominated by the genera *Pseudoalteromonas* and *Alteromonas* (Supplementary Dataset). Other abundant genera included *Fulvivirga* (Bacteroidetes) and *Nisaea* (Alphaproteobacteria) (Supplementary Dataset). These genera are common members of the *Trichodesmium* microbiome based on previous 16S and metagenomic surveys of *Trichodesmium* epibiont diversity [11–14].

Expression profiles from the *Trichodesmium* and microbiome OGs were analyzed with a WGCNA [7, 9, 32] to ascertain if there was significant co-expression between *Trichodesmium* and the microbiome. Expression patterns were significantly coordinated between *Trichodesmium* and the microbiome, with 3,140 *Trichodesmium* and 4,730 epibiont OGs clustering into 33 WGCNA co-expression

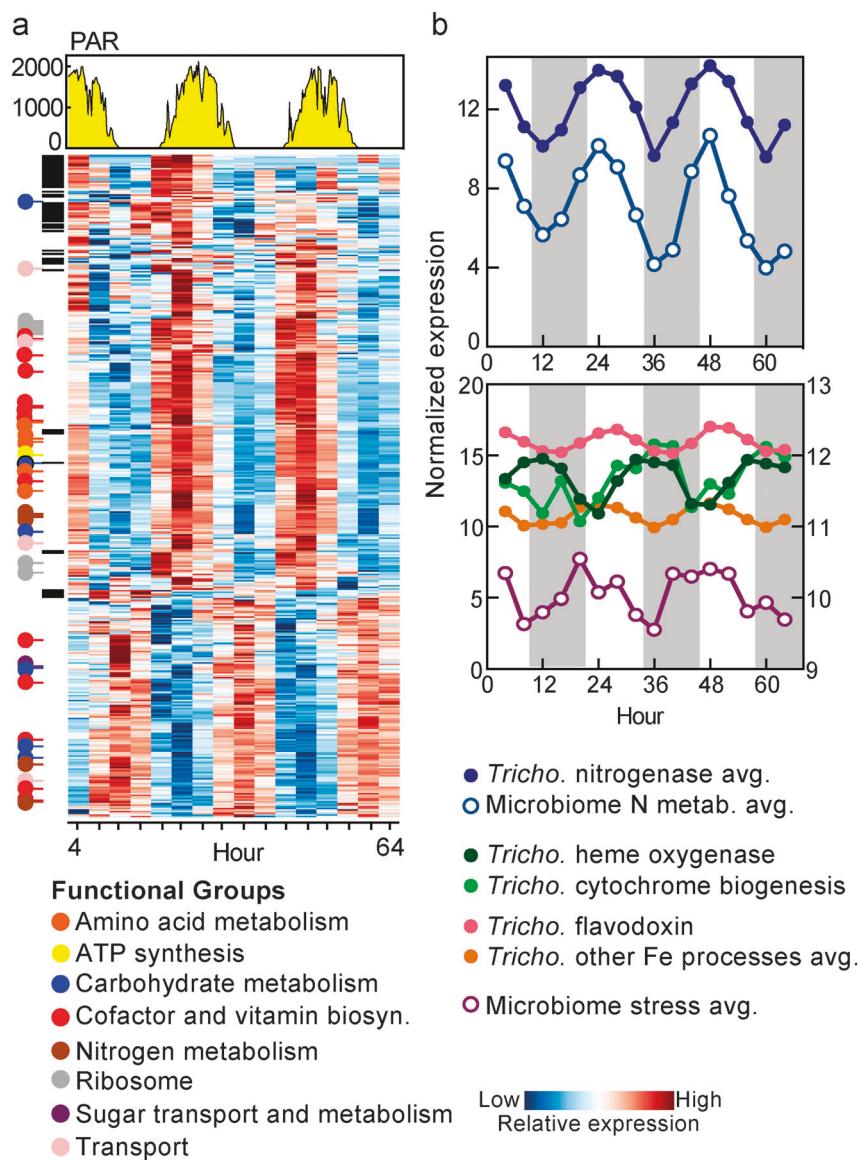
modules (Fig. 1b). The presence of OGs from both *Trichodesmium* and the microbiome within every significant co-expression module indicates that gene expression is tightly integrated within the holobiont. Many co-expression modules were also dominated by microbiome OGs (Fig. 1b), such as Module 9, which was composed of a suite of different functions including those related to ribosomes, sugar transport and metabolism, ATP synthesis, and phosphate and amino acid transport functions (Supplementary Figure 1). The presence of microbiome-dominated modules suggests that there is significant co-expression between epibionts. Together, these patterns of co-expression suggest coordination between *Trichodesmium* and the microbiome, as well as between epibionts within the microbiome.

Dynamics within the majority of these co-expression modules were defined by diel oscillations in gene expression that had significant periodicity as determined by RAIN [29] (Fig. 1b). In *Trichodesmium*, these periodic OGs included gene sets related to photosynthesis, phosphorus, iron, N<sub>2</sub>, and carbon fixation among others (Supplementary Figure 2; Supplementary Dataset). Global expression patterns in *Trichodesmium* would be expected to be dominated by diel patterns, as photosynthetic and N<sub>2</sub> fixing activities are both tightly regulated across day–night cycles [1]. Here, periodic microbiome expression patterns strikingly mirrored those observed in *Trichodesmium* (Supplementary Figure 3). In the microbiome, OGs with day–night periodicity included genes related to stress responses, nitrogen metabolism, carbon metabolism, and cofactor and vitamin biosynthesis among others (Supplementary Figure 2; Supplementary Dataset). Similar synchronous diel transcriptional oscillations have also been observed between unicellular marine cyanobacteria and free-living heterotrophic bacteria, as heterotrophic activity appears to be tied to pulses in primary production [7, 8]. Overall, microbiome activities are likely tied to both *Trichodesmium* photosynthesis and N<sub>2</sub> fixation as the release of dissolved nitrogen and organic carbon is tied to these processes. Though *Trichodesmium* fixes both carbon and N<sub>2</sub> during daylight [33], net release of organic carbon, dissolved organic nitrogen, and ammonium oscillate throughout the day and night [34]. The extensive coordination and functional characteristics in gene expression between *Trichodesmium* and the microbiome suggest that the physiology of these associated microbes are attuned to the host, and could thus influence the cycling and fate of carbon, nitrogen, and other resources in the colony microenvironment.

### Coordinated expression in nitrogen and carbon pathways

The OGs related to *Trichodesmium* N<sub>2</sub> fixation, including the genes of the nitrogenase-encoding *nif* cassette, were

**Fig. 2** Resource-related patterns in the *nif*-containing co-expression Module 3. **a** Photosynthetically active radiation (PAR) levels are shown over the duration of the Lagrangian sampling period and correspond to vertical bars in the heatmap, which depicts row averaged relative expression of the significantly co-expressed *Trichodesmium* and microbiome orthologous groups (OGs). Black bars denote rows that are microbiome OGs. Colored circles show functional annotations based on KEGG submodules. **b** Details of select OG expression in Module 3 highlighting patterns between *Trichodesmium* and the microbiome related to nitrogen and iron metabolism. Gray bars indicate dark conditions. *Tricho.*, *Trichodesmium*; nitrogenase avg., average normalized expression for all *nif* genes clustered in Module 3; microbiome N metab. avg., average normalized expression of nitrogen-related OGs in the microbiome subset in Module 3, including aminotransferases, amino acid adenylate proteins, and peptidases; *Tricho.* other Fe processes avg., average normalized expression of iron-related OGs including cytochromes, iron-sulfur cluster assembly protein, iron-requiring hydrogenase, ferrochelatase, ferredoxin, and ferredoxin-dependent bilin reductase; microbiome stress avg., average normalized expression of a chaperone and two general stress proteins

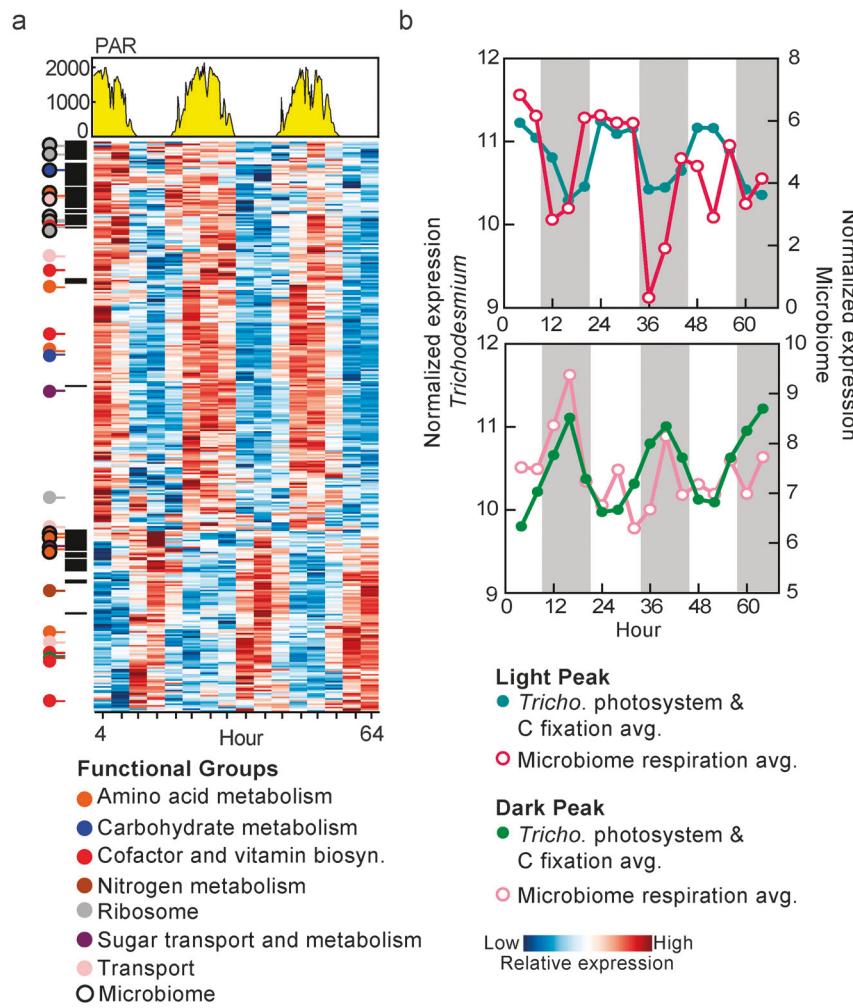


clustered into co-expression Module 3. Module 3 contained 348 *Trichodesmium* and 57 microbiome OGs (Fig. 1b) and was dominated by OGs with maximum or minimum expression just prior to peak photosynthetically active radiation (PAR) and strong day–night periodicity (Fig. 2a). Strikingly, Module 3 contained a suite of microbiome OGs related to nitrogen processes, which mirrored expression of nitrogenase subunit OGs in *Trichodesmium* (Fig. 2b). These nitrogen-related microbiome OGs include aminotransferases, amino acid adenylate proteins, and peptidases with significantly periodic patterns (Fig. 2b). *Trichodesmium* releases biologically available nitrogen in culture and the environment, even when maintained axenically [34–37]. The extent to which this release facilitates

cultivation of a symbiotic heterotrophic bacterial microbiome or whether associated epibionts are simply a product of opportunistic, stochastic colonization remains an outstanding question. However, this expression data, coupled with recent observations of ubiquitous epibionts across ocean basins [13, 14], underscores conserved relationships in these consortia wherein the transfer of *Trichodesmium*-derived fixed N<sub>2</sub> may synchronize microbiome metabolism with host physiology.

Expression patterns between host and microbiome also indicate possible synchrony between photosynthetic carbon fixation and release by *Trichodesmium* and the respiration of that organic carbon by the microbiome. Module 8 had strong day–night periodicity in both *Trichodesmium* and the

**Fig. 3** Patterns in carbon fixation and respiration in co-expression Module 8. **a** Photosynthetically active radiation (PAR) levels are shown over the duration of the Lagrangian sampling period and correspond to vertical bars in the heatmap, which depicts row averaged relative expression of the significantly co-expressed *Trichodesmium* and microbiome orthologous groups (OGs). Black bars denote rows that are microbiome OGs. Colored circles show functional annotations based on KEGG submodules. **b** Detail of expression patterns in *Trichodesmium* photosystem and carbon fixation OGs and microbiome respiration OGs. Gray bars indicate dark conditions. Microbiome respiration OGs averaged include pyruvate carboxyltransferase and a C<sub>4</sub>-dicarboxylate transporter in the top panel and two glycosyltransferases in the bottom panel. *Tricho.*, *Trichodesmium*



microbiome and contained a high proportion of *Trichodesmium* photosystem and carbon fixation OGs, including a RuBisCO subunit, three photosystem subunits, and a phycoerythrin assembly protein (Fig. 3a). *Trichodesmium* photosynthetic- related and carbon fixation-related OGs were segregated into two significantly co-expressed patterns, which had peak expression during light and dark periods (Fig. 3b; Supplementary Dataset). The microbiome subset of this module contained a number of functions that suggested synchronization of carbon metabolism with *Trichodesmium* including those related to gluconeogenesis, a sugar binding protein, sugar transporter, and carbohydrate active enzymes, as well as OGs related to ribosome functions and amino acid metabolism (Fig. 3). The expression patterns in these microbiome OGs exhibited peaks in both light and dark conditions that mirrored those of *Trichodesmium* photosystem and carbon fixation-related gene expression (Fig. 3b). The microbiome pyruvate carboxyltransferase, an enzyme from gluconeogenesis and a transporter of C<sub>4</sub>-dicarboxylates like succinate, fumarate, and malate [38] peaked during daylight. Peaking during nighttime were two microbiome glycosyltransferase enzymes,

along with *Trichodesmium* RuBisCO (Fig. 3b). Overall, these microbiome expression patterns are consistent with organic carbon metabolism in the microbiome being tuned to periodic fluxes of *Trichodesmium* organic carbon.

*Trichodesmium* releases an estimated 50% of fixed carbon [34] and up to 20% of fixed N<sub>2</sub> from colonies [33]. This loss of resources is paradoxical in the oligotrophic ocean; however, it may support microbiome colonization and beneficial syntrophic interactions with the holobiont. Tracing these potential interactions *in situ* is challenging, but they could exert important unrecognized feedbacks on the physiological ecology of *Trichodesmium* populations. For example, one benefit of the microbiome could be that microbiome respiration of *Trichodesmium* organic carbon alleviates carbon limitation of photosynthesis as well as O<sub>2</sub> inhibition of N<sub>2</sub> fixation [10, 13]. Such a feedback is consistent with the coordinated patterns of gene expression between host and microbiome seen here. If the microbiome modulates the *Trichodesmium* microenvironment in a way that decouples rates of carbon and N<sub>2</sub> fixation from water column geochemistry, or other abiotic drivers of these activities, this would exacerbate ongoing challenges in

modeling these processes in the oligotrophic ocean [39, 40]. Further, the coordinated expression of microbiome nitrogen and carbon metabolism genes with *Trichodesmium* activities, like N<sub>2</sub> fixation and carbon fixation, might allow the microbiome to influence the net flux of resources to the water column.

### Transcriptional evidence of shared phosphorus and iron demand

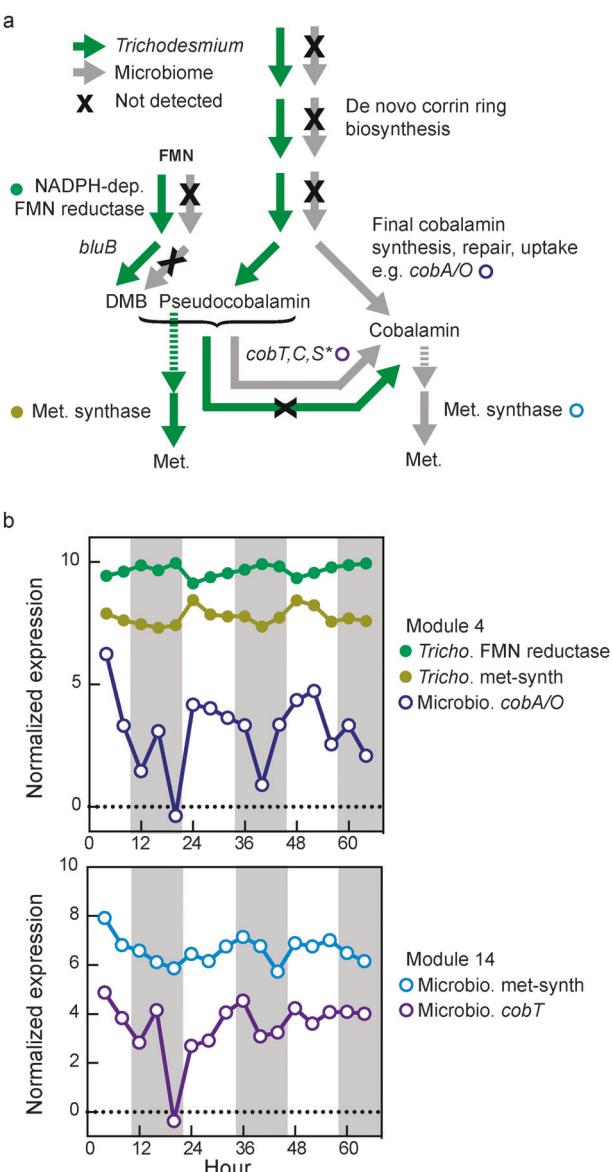
In addition to the *nif* enzyme subunits and several OGs related to photosynthesis, Module 3 also contained a number of additional *Trichodesmium* OGs encoding iron-requiring proteins (Fig. 2b), suggesting complex coordination of resources across these processes. In *Crocospaera*, a marine diazotroph that decouples oxygen evolving photosynthesis and N<sub>2</sub> fixation between day and night, respectively, the necessary iron quotas are met by degradation of metalloenzymes that liberate and subsequently repurpose iron across different processes over day–night cycles [41]. Unlike *Crocospaera*, *Trichodesmium* cannot shuttle iron between photosynthetic and N<sub>2</sub> fixing processes because both occur during the day. The majority of *Trichodesmium* iron-requiring OGs in Module 3 had peak daytime expression that mirrored *nif* and photosystem expression (Figs. 2b and 3b), and included cytochromes, ferredoxin, ferrochelatase, ferredoxin-dependent bilin reductase, an iron-sulfur cluster assembly protein, and an iron-requiring hydrogenase. Also peaking during daylight in this module was an OG identified as flavodoxin (Fig. 3b), an electron transport protein that does not require iron and can substitute for ferredoxin during low iron conditions [42]. In contrast, the expression of an OG identified as heme oxygenase, an enzyme involved in liberating iron from organic complexes [41], was opposite of *nif* expression in *Trichodesmium* (Fig. 2b). In *Crocospaera*, heme oxygenase also showed diel variability with peak expression opposite that of nitrogenase [41], suggesting that this enzyme could be used to liberate iron for cellular processes that are not related to N<sub>2</sub> fixation. In Module 3, a cytochrome biogenesis gene was the only iron-requiring OG that had expression dynamics similar to heme oxygenase (Fig. 2b), indicating that iron liberated at night could be repurposed for use in electron transport processes. Overall, the dynamics of gene expression in iron-related OGs in *Trichodesmium* underscore high daytime iron requirements that likely modulate competition for this critical resource within the holobiont over day–night cycles.

In addition to nitrogen and iron processes, *Trichodesmium* also expressed OGs for phosphate uptake in Module 3, with significantly periodic daytime expression peaks that were slightly offset from expression of the OGs for phosphonate hydrolysis, which also had significant

periodic expression but at a later peak time (Supplementary Figure 4). These patterns likely indicate increased daytime demand for phosphorus to support growth and N<sub>2</sub> fixation in *Trichodesmium*. Coincident with the expression of these *Trichodesmium* genes in Module 3 was a suite of microbiome OGs including the chaperone *htpG*, one general stress protein, and a *uspA*-like universal stress protein (Fig. 2b). In other species of bacteria, these protein families have been implicated in a variety of functions which are modulated in response to stressors such as limitation for key resources like phosphorus and iron as well as exposure to antibiotics and oxidative agents [43, 44]. These patterns suggest that relative to nitrogen, there may be an offset in the peak bioavailability of resources like phosphorus or iron to the microbiome. This offset could result in resource limitation scenarios that oscillate within the holobiont microenvironment. These data also highlight the complexity of interactions within the holobiont that span multiple resources. For example, time periods of peak N<sub>2</sub> fixation and photosynthesis, when biologically available resources should be highest for the microbiome, are concurrent with microbiome cellular stress. In conjunction with co-expression in nitrogen and carbon functions, this observation could indicate a system where modulation of *Trichodesmium*-derived resources maintains a stable, tightly regulated relationship between the host and microbiome. In contrast, at the end of a bloom *Trichodesmium* releases an abundance of organic material causing a change in community composition of the associated bacteria hypothesized to be related to the growth of opportunistic copiotrophs [45]. A similar phenomenon has also been observed in bacteria that colonize corals, where an abundance of labile carbon resulted in a shift towards more virulent activities [46]. During non-bloom conditions in *Trichodesmium*, avoiding such a resource-rich environment could help maintain a beneficial microbiome and prevent the growth of copiotrophs that could skew potential interactions towards those that are deleterious or parasitic.

### Microbiome dependence on *Trichodesmium* cobalamin

Despite the critical importance of the B12 vitamin cobalamin to all organisms, only select bacteria and archaea are capable of de novo cobalamin biosynthesis, making exchange between producers and non-producers critical [28, 47, 48]. Based on genome sequence analysis, *Trichodesmium* should possess the capability for de novo cobalamin biosynthesis [49], and OGs related to cobalamin production were detected and expressed in *Trichodesmium* across several co-expression modules largely without significant diel periodicity (Fig. 4a; Supplementary Figure 5). Only one OG in the cobalamin biosynthesis pathway



**Fig. 4** Detection of cobalamin pathway orthologous groups (OGs) and related transcriptional patterns. **a** Presence and absence of OGs from the transcriptome data in the de novo cobalamin biosynthesis, uptake and modification, and utilization pathways in *Trichodesmium* and the microbiome. \*Only *cobT* was detected in the microbiome. Colored circles indicate corresponding expression patterns that are depicted in **b**. **b** Detail of expression patterns in key cobalamin modification and utilization OGs in *Trichodesmium* and the microbiome. FMN flavin mononucleotide, DMB, 5,6-dimethylbenzimidazole, Met. methionine, *Tricho*. *Trichodesmium*. Gray bars indicate dark conditions

(*cobW*) was detected in the microbiome (Supplementary Dataset), which indicates that these genes were either not expressed at a high enough level to pass our analysis threshold or that the microbiome members do not possess the capability to synthesize cobalamin. Together with the recent finding that a conserved *Trichodesmium* microbiome member is a cobalamin auxotroph [13] and that marine diazotrophs are known to secrete large amounts of this

vitamin [50], these results suggest that *Trichodesmium* supplies cobalamin to the microbiome.

Recent findings have shown the form of cobalamin produced by cyanobacteria, pseudocobalamin, is not readily bioavailable to heterotrophic organisms [49, 51]. Prior to use by heterotrophic bacteria, an adenine ligand on cyanobacterial pseudocobalamin must be exchanged with 5,6-dimethylbenzimidazole (DMB)—a process that relies on an exogenous supply of DMB, de novo biosynthesis of DMB with the *bluB* gene or the alternative oxygen-sensitive *bzaABCDE* pathway [52]—and finally the activation of DMB with the enzyme nicotinate-nucleotide-dimethylbenzimidazole phosphoribosyltransferase (*cobT*) [49, 51, 53]. Within the microbiome, the OG for cobalamin uptake (*cobA/O*) was detected in Module 4, a module that also contained a suite of significantly co-expressed *Trichodesmium* OGs for cobalamin synthesis and cobalamin-dependent methionine synthase (Fig. 4b; Supplementary Figure 1; Supplementary Figure 5). Supporting the hypothesis that the microbiome is using *Trichodesmium*-derived cobalamin, the microbiome methionine synthase was clustered in a co-expression module that also contained microbiome *cobT* (Fig. 4b). However, despite this transcriptional evidence for microbiome conversion of *Trichodesmium*-derived pseudocobalamin and subsequent use, homologs of *bluB* and *bzaABCDE* were not detected in the microbiome subset, indicating that essential DMB is likely supplied from a source outside the microbiome.

Contrary to the overwhelming majority of cyanobacteria that do not possess the capability to synthesize or activate DMB, the *Trichodesmium* IMS101 genome encodes the *bluB* gene for DMB production [49]. Within our dataset, an OG identified as *bluB* was detected and expressed in *Trichodesmium* Module 26 (Supplementary Figure 5). Furthermore, an OG identified as flavin mononucleotide (FMN) reductase, an enzyme that generates the precursor to DMB and acts as the substrate of the BluB enzyme [53], was detected in *Trichodesmium* Module 4 (Supplementary Figure 1) along with *Trichodesmium* OGs from the cobalamin biosynthesis pathway, methionine synthase, and the microbiome *cobA/O* (Fig. 4b). Strikingly, expression patterns of *Trichodesmium* FMN reductase and cobalamin-dependent methionine synthase had significant, anti-correlated diel periodicity (Fig. 4b). One interpretation of the presence of *bluB* could be that *Trichodesmium* uses DMB to convert its own pseudocobalamin prior to downstream use of this vitamin. However, the anti-correlated expression dynamics between DMB production and cobalamin-dependent processes (Fig. 4b) seem to suggest that *Trichodesmium* is not converting its own pseudocobalamin prior to use. Instead, these data suggest that *Trichodesmium* both supplies pseudocobalamin to the microbiome and controls the timing of subsequent ligand conversion

within the microbiome, regulating this process to be out of phase with their own cobalamin-dependent enzyme gene expression. These synchronized transcriptional patterns support evidence from metagenomic data that suggests that in *Trichodesmium* consortia the direction of cobalamin transfer is opposite that of the canonical pathway [13], where heterotrophic bacterial producers typically share this vitamin with photosynthetic algae in exchange for fixed carbon [28, 54]. The processes that govern recruitment and maintenance of microbiome community structure are largely unknown, but the lack of evidence for microbiome cobalamin biosynthesis is consistent with the hypothesis from Lee et al. [13] that this vitamin could exert selective pressure on microbiome community structure. If *Trichodesmium* DMB production does control subsequent microbiome cobalamin use, then this would be an example of a microbiome dependency that drives the community structure of the *Trichodesmium* holobiont. This tightly controlled structure both differentiates *Trichodesmium* colonies from transient microbial consortia on particles that are dominated by remineralization activities [55], and reinforces the apparent microbiome influence on the flux of important resources, like cobalamin, to the oligotrophic environment.

## Conclusions

The future oceans are predicted to be warmer, higher in CO<sub>2</sub>, and have expanded oligotrophic regions [40, 56, 57]. It is expected that these conditions will increase the rate of *Trichodesmium* carbon and N<sub>2</sub> fixation [56, 58] and potentially magnify the relative importance of *Trichodesmium* activities within expanded regions of oligotrophy. In other marine systems, these future ocean conditions are known to destabilize the microbiome with resultant deleterious effects on the host [59–61]. Here we show that, like these other host–microbiome systems, the *Trichodesmium* microbiome is closely synchronized with the host, and that host–microbiome interactions likely influence the cycling of carbon, nitrogen, iron, and cobalamin. To fully understand and forecast future ocean dynamics, studies of *Trichodesmium* must consider the role of the microbiome.

**Acknowledgements** We thank the scientists who worked collaboratively on the 2015 research expedition as part of the Simons Collaboration on Ocean Processes and Ecology (SCOPE). For assistance at sea, we thank the captain and crew of the R/V Kilo Moana, and Samuel Wilson for serving as chief scientist. We also thank Frank Aylward for assistance with statistical analyses and Matthew Harke, Mónica Rouco, Katherine Heal, and Harriet Alexander for helpful discussions of the data. Finally, we thank the three anonymous reviewers whose comments and suggestions strengthened the manuscript. This research was funded by the Simons Foundation (SCOPE award ID 329108 to STD), and is a contribution of SCOPE. KRF was

partially supported by a National Science Foundation Graduate Research Fellowship (DGE-16-44869).

**Author contributions** KRF and STD designed the study. All authors carried out sampling. KRF performed lab work and analyzed the data. KRF and STD wrote the manuscript with input from STH.

## Compliance with ethical standards

**Conflict of interest** The authors declare that they have no conflict of interest.

## References

1. Bergman B, Sandh G, Lin S, Larsson J, Carpenter EJ. *Trichodesmium*—a widespread marine cyanobacterium with unusual nitrogen fixation properties. *FEMS Microbiol Rev*. 2013;37:286–2.
2. Sohm JA, Webb EA, Capone DG. Emerging patterns of marine nitrogen fixation. *Nat Rev Microbiol*. 2011;9:499–8.
3. Falkowski PG, Barber RT, Smetacek V. Biogeochemical controls and feedbacks on ocean primary production. *Science*. 1998;281:200–6.
4. Zehr JP. Nitrogen fixation by marine cyanobacteria. *Trends Microbiol*. 2011;19:162–73.
5. Moran MA, Kujawinski EB, Stubbins A, Fatland R, Aluwihare LI, Buchan A, et al. Deciphering ocean carbon in a changing world. *Proc Natl Acad Sci USA*. 2016;113:3143–51.
6. Brussaard CPD, Bidle KD, Pedrós-alió C, Legrand C. The interactive microbial ocean. *Nat Microbiol*. 2016;2:1–2.
7. Aylward FO, Eppley JM, Smith JM, Chavez FP, Scholin CA, DeLong EF. Microbial community transcriptional networks are conserved in three domains at ocean basin scales. *Proc Natl Acad Sci USA*. 2015;112:5443–48.
8. Ottesen EA, Young CR, Gifford SM, Eppley JM, Marin R, Schuster SC, et al. Multispecies diel transcriptional oscillations in open ocean heterotrophic bacterial assemblages. *Science*. 2014;345:207–12.
9. Wilson ST, Aylward FO, Ribalet F, Barone B, Casey JR, Connell PE, et al. Coordinated regulation of growth, activity and transcription in natural populations of the unicellular nitrogen-fixing cyanobacterium *Crocospaera*. *Nat Microbiol*. 2017;2:1–9. <https://doi.org/10.1038/nmicrobiol.2017.118>.
10. Paerl HW, Bebout BM, Prufert LE. Bacterial associations with marine *Oscillatoria* sp. (*Trichodesmium* sp.) populations: eco-physiological implications. *J Phycol*. 1989;25:773–84.
11. Frischkorn KR, Rouco M, Van Mooy BAS, Dyhrman ST. Epibionts dominate metabolic functional potential of *Trichodesmium* colonies from the oligotrophic ocean. *ISME J*. 2017;11:2090–101.
12. Gradoville MR, Crump BC, Letelier RM, Church MJ, White AE. Microbiome of *Trichodesmium* colonies from the North Pacific Subtropical Gyre. *Front Microbiol*. 2017;8:1–16. <https://doi.org/10.3389/fmicb.2017.01122>.
13. Lee MD, Walworth NG, Mcparland EL, Fu F, Mincer TJ, Levine NM, et al. The *Trichodesmium* consortium: conserved heterotrophic co-occurrence and genomic signatures of potential interactions. *ISME J*. 2017;11:1813–24.
14. Rouco M, Haley ST, Dyhrman ST. Microbial diversity within the *Trichodesmium* holobiont. *Environ Microbiol*. 2016;18:5151–60.
15. Hmelo L, Van Mooy B, Mincer T. Characterization of bacterial epibionts on the cyanobacterium *Trichodesmium*. *Aquat Microb Ecol*. 2012;67:1–14.
16. Van Mooy BAS, Hmelo LR, Sofen LE, Campagna SR, May AL, Dyhrman ST, et al. Quorum sensing control of phosphorus

acquisition in *Trichodesmium* consortia. *ISME J.* 2012;6:422–429.

17. Hewson I, Poretsky RS, Dyhrman ST, Zielinski B, White AE, Tripp HJ, et al. Microbial community gene expression within colonies of the diazotroph, *Trichodesmium*, from the Southwest Pacific Ocean. *ISME J.* 2009;3:1286–300.
18. Brown CT, Sheneman LS, Camille CM, Rosenthal JH, Adina C. khmer-protocols documentation. 2013. <https://khmer-protocols.readthedocs.io/en/latest/mrnaseq/index.html>
19. Alexander H, Rouco M, Haley ST, Wilson ST, Karl DM, Dyhrman ST. Functional group-specific traits drive phytoplankton dynamics in the oligotrophic ocean. *Proc Natl Acad Sci USA.* 2015;112:E5972–79.
20. Frischkorn KR, Harke MJ, Gobler CJ, Dyhrman ST. De novo assembly of *Aureococcus anophagefferens* transcriptomes reveals diverse responses to the low nutrient and low light conditions present during blooms. *Front Microbiol.* 2014;5:1–16.
21. Li W, Godzik A. Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences. *Bioinformatics.* 2006;22:1658–59.
22. Varaljay VA, Howard EC, Sun S, Moran MA. Deep sequencing of a dimethylsulfoniopropionate-degrading gene (dmDA) by using PCR primer pairs designed on the basis of marine metagenomic data. *Appl Environ Microbiol.* 2010;76:609–17.
23. Hyatt D, Chen G-L, Locascio PF, Land ML, Larimer FW, Hauser LJ. Prodigal: prokaryotic gene recognition and translation initiation site identification. *BMC Bioinform.* 2010;11:1–11. <https://doi.org/10.1186/1471-2105-11-119>.
24. Buchfink B, Xie C, Huson DH. Fast and sensitive protein alignment using DIAMOND. *Nat Methods.* 2015;12:59–60.
25. Huson DH, El Hadidi M, Ruscheweyh H, Huson DH, Meta S. Improved metagenome analysis using MEGAN5. 2013. <http://ab.inf.uni-tuebingen.de/software/megan5/>
26. Suzek BE, Huang H, McGarvey P, Mazumder R, Wu CH. UniRef: comprehensive and non-redundant UniProt reference clusters. *Bioinformatics.* 2007;23:1282–88.
27. Li B, Dewey CN. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinform.* 2011;12:1–16. <https://doi.org/10.1186/1471-2105-12-323>.
28. Bertrand EM, McCrow JP, Moustafa A, Zheng H, McQuaid JB, Delmont TO, et al. Phytoplankton–bacterial interactions mediate micronutrient colimitation at the coastal Antarctic sea ice edge. *Proc Natl Acad Sci USA.* 2015;112:9938–43.
29. Thaben PF, Westermark PO. Detecting rhythms in time series with RAIN. *J Biol Rhythms.* 2014;29:391–400.
30. Anders S, Huber W. Differential expression analysis for sequence count data. *Genome Biol.* 2010;11:R106 <https://doi.org/10.1186/gb-2010-11-10-r106>.
31. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J R Stat Soc Ser B.* 1995;57:289–300.
32. Langfelder P, Horvath S. WGCNA: an R package for weighted correlation network analysis. *BMC Bioinforma.* 2008;9:1–13. <https://doi.org/10.1186/1471-2105-9-559>.
33. Berthelot H, Bonnet S, Camps M, Gross O. Assessment of the dinitrogen released as ammonium and dissolved organic nitrogen by unicellular and filamentous marine diazotrophic cyanobacteria grown in culture. *Front Mar Sci.* 2015;2:1–14.
34. Wannicke N, Koch BP, Voss M. Release of fixed N<sub>2</sub> and C as dissolved compounds by *Trichodesmium erythraeum* and *Nodularia spumigena* under the influence of high light and high nutrient (P). *Aquat Microb Ecol.* 2009;57:175–89.
35. Glibert PM, Bronk DA. Release of dissolved organic nitrogen by marine diazotrophic cyanobacteria, *Trichodesmium* spp. *Appl Environ Microbiol.* 1994;60:3996–4000.
36. Mulholland MR, Bernhardt PW, Heil CA, Bronk DA, Neil JMO, Boneillo G, et al. Nitrogen fixation and release of fixed nitrogen by *Trichodesmium* spp. in the Gulf of Mexico. *Limnol Oceanogr.* 2006;51:1762–76.
37. Mulholland MR, Bronk DA, Capone DG. Dinitrogen fixation and release of ammonium and dissolved organic nitrogen by *Trichodesmium*. *Aquat Microb Ecol.* 2004;37:85–94.
38. Janausch IG, Zientz E, Tran QH, Kro A. C<sub>4</sub>-dicarboxylate carriers and sensors in bacteria. *Biochim Biophys Acta.* 2002;1553:39–56.
39. Capone DG, Burns JA, Montoya JP, Subramaniam A, Mahaffey C, Gunderson T, et al. Nitrogen fixation by *Trichodesmium* spp.: an important source of new nitrogen to the tropical and subtropical North Atlantic Ocean. *Glob Biogeochem Cycles.* 2005;19: GB2024. <https://doi.org/10.1029/2004GB002331>.
40. McGillicuddy DJ. Do *Trichodesmium* spp. populations in the North Atlantic export most of the nitrogen they fix? *Glob Biogeochem Cycles.* 2014;28:103–14.
41. Saito MA, Bertrand EM, Dutkiewicz S, Bulygin VV, Moran DM, Monteiro FM, et al. Iron conservation by reduction of metalloenzyme inventories in the marine diazotroph *Crocospaera watsonii*. *Proc Natl Acad Sci USA.* 2011;108:2184–89.
42. Erdner DL, Anderson DM. Ferredoxin and flavodoxin as biochemical indicators of iron limitation during open-ocean iron enrichment. *Limnol Oceanogr.* 1999;44:1609–15.
43. Liu W, Karavolos MH, Bulmer DM, Allaoui A, Demarco R, Hormaeche CE, et al. Role of the universal stress protein UspA of *Salmonella* in growth arrest, stress and virulence. *Microb Pathog.* 2007;42:2–10.
44. Nachin L, Nannmark U, Nyström T, Nystro T. Differential roles of the universal stress proteins of *Escherichia coli* in oxidative stress resistance, adhesion, and motility. *J Bacteriol.* 2005;187:6265–72.
45. Spungin D, Pfreundt U, Berthelot H, Bonnet S, AlRoumi D, Natale F, et al. Mechanisms of *Trichodesmium* bloom demise within the New Caledonian Lagoon during the VAHINE mesocosm experiment. *Biogeoscience.* 2016;13:4187–203.
46. Cárdenas A, Neave MJ, Haroon MF, Pogoreutz C, Rädecker N, Wild C, et al. Excess labile carbon promotes the expression of virulence factors in coral reef bacterioplankton. *ISME J.* 2017;00:1–18. <https://doi.org/10.1038/ismej.2017.142>.
47. Sañudo-Wilhelmy SA, Gómez-Consarnau L, Suffridge C, Webb EA. The role of B vitamins in marine biogeochemistry. *Annu Rev Mar Sci.* 2014;6:339–67.
48. Warren MJ, Raux E, Schubert HL, Escalante-Semerena JC. The biosynthesis of adenosylcobalamin (vitamin B12). *Nat Prod Rep.* 2002;19:390–12.
49. Hellwell KE, Lawrence AD, Holzer A, Scanlan DJ, Warren MJ, Smith AG, et al. Cyanobacteria and eukaryotic algae use different chemical variants of vitamin B12. *Curr Biol.* 2016;26:999–08.
50. Bonnet S, Webb EA, Panzeca C, Karl DM, Capone DG, Sañudo-Wilhelmy SA. Vitamin B12 excretion by cultures of the marine cyanobacteria *Crocospaera* and *Synechococcus*. *Limnol Oceanoogr.* 2010;55:1959–64.
51. Heal KR, Qin W, Ribalet F, Bertagnoli AD, Coyote-Maestas W, Hmelo LR, et al. Two distinct pools of B12 analogs reveal community interdependencies in the ocean. *Proc Natl Acad Sci USA.* 2016;114:364–69.
52. Hazra AB, Han AW, Mehta AP, Mok KC, Osadchy V, Begley TP, et al. Anaerobic biosynthesis of the lower ligand of vitamin B12. *Proc Natl Acad Sci USA.* 2015;112:10792–97.
53. Taga ME, Larsen NA, Howard-Jones AR, Walsh CT, Walker GC. Blub cannibalizes flavin to form the lower ligand of vitamin B12. *Nature.* 2007;446:449–53.
54. Croft MT, Lawrence AD, Raux-deery E, Warren MJ, Smith AG. Algae acquire vitamin B12 through a symbiotic relationship with bacteria. *Nature.* 2005;438:90–93.

55. Fontanez KM, Eppley JM, Samo TJ, Karl DM, DeLong EF. Microbial community structure and function on sinking particles in the North Pacific Subtropical Gyre. *Front Microbiol*. 2015;6:1–14.
56. Hutchins DA, Walworth NG, Webb EA, Saito MA, Moran D, McIlvin MR, et al. Irreversibly increased nitrogen fixation in *Trichodesmium* experimentally adapted to elevated carbon dioxide. *Nat Commun*. 2015;6:1–7.
57. Riebesell U, Körtzinger A, Oschlies A. Sensitivities of marine carbon fluxes to ocean change. *Proc Natl Acad Sci USA*. 2009;106:20602–9.
58. Hutchins DA, Fu F-X, Zhang Y, Warner ME, Feng Y, Portune K, et al. CO<sub>2</sub> control of *Trichodesmium* N<sub>2</sub> fixation, photosynthesis, growth rates, and elemental ratios: implications for past, present, and future ocean biogeochemistry. *Limnol Oceanogr*. 2007;52:1293–304.
59. Ainsworth TD, Thurber RV, Gates RD. The future of coral reefs: a microbial perspective. *Trends Ecol Evol*. 2010;25:233–40.
60. Bourne DG, Morrow KM, Webster NS. Insights into the coral microbiome: underpinning the health and resilience of reef ecosystems. *Annu Rev Microbiol*. 2016;70:317–40.
61. Lesser MP, Fiore C, Slattery M, Zaneveld J. Climate change stressors destabilize the microbiome of the Caribbean barrel sponge *Xestospongia muta*. *J Exp Mar Bio Ecol*. 2016;475:11–18.