AMERICAN SOCIETY FOR MICROBIOLOGY | Applied and Environmental Microbiology®

ENVIRONMENTAL MICROBIOLOGY

Transcriptional Activities of the Microbial Consortium Living with the Marine Nitrogen-Fixing Cyanobacterium *Trichodesmium* Reveal Potential Roles in Community-Level Nitrogen Cycling

Michael D. Lee,[a] Eric A. Webb,[a] Nathan G. Walworth,[a] Fei-Xue Fu,[a] Noelle A. Held,[b] Mak A. Saito,[b] David A. Hutchins[a]

[a]Marine Environmental Biology, Department of Biological Sciences, University of Southern California, Los Angeles, California, USA

[b]Marine Chemistry and Geochemistry Department, Woods Hole Oceanographic Institute, Woods Hole, Massachusetts, USA

ABSTRACT *Trichodesmium* is a globally distributed cyanobacterium whose nitrogen-fixing capability fuels primary production in warm oligotrophic oceans. Like many photoautotrophs, *Trichodesmium* serves as a host to various other microorganisms, yet little is known about how this associated community modulates fluxes of environmentally relevant chemical species into and out of the supraorganismal structure. Here, we utilized metatranscriptomics to examine gene expression activities of microbial communities associated with *Trichodesmium erythraeum* (strain IMS101) using laboratory-maintained enrichment cultures that have previously been shown to harbor microbial communities similar to those of natural populations. In enrichments maintained under two distinct CO_2 concentrations for ~8 years, the community transcriptional profiles were found to be specific to the treatment, demonstrating a restructuring of overall gene expression had occurred. Some of this restructuring involved significant increases in community respiration-related transcripts under elevated CO_2, potentially facilitating the corresponding measured increases in host nitrogen fixation rates. Particularly of note, in both treatments, community transcripts involved in the reduction of nitrate, nitrite, and nitrous oxide were detected, suggesting the associated organisms may play a role in colony-level nitrogen cycling. Lastly, a taxon-specific analysis revealed distinct ecological niches of consistently cooccurring major taxa that may enable, or even encourage, the stable cohabitation of a diverse community within *Trichodesmium* consortia.

IMPORTANCE *Trichodesmium* is a genus of globally distributed, nitrogen-fixing marine cyanobacteria. As a source of new nitrogen in otherwise nitrogen-deficient systems, these organisms help fuel carbon fixation carried out by other more abundant photoautotrophs and thereby have significant roles in global nitrogen and carbon cycling. Members of the *Trichodesmium* genus tend to form large macroscopic colonies that appear to perpetually host an association of diverse interacting microbes distinct from the surrounding seawater, potentially making the entire assemblage a unique miniature ecosystem. Since its first successful cultivation in the early 1990s, there have been questions about the potential interdependencies between *Trichodesmium* and its associated microbial community and whether the host's seemingly enigmatic nitrogen fixation schema somehow involved or benefited from its epibionts. Here, we revisit these old questions with new technology and investigate gene expression activities of microbial communities living in association with *Trichodesmium*.

KEYWORDS *Trichodesmium*, bacterial consortium, gene expression, high CO_2 adapted, metatranscriptome, nitrogen fixation, proteome

Received 14 September 2017 **Accepted** 15 October 2017

Accepted manuscript posted online 20 October 2017

Citation Lee MD, Webb EA, Walworth NG, Fu F-X, Held NA, Saito MA, Hutchins DA. 2018. Transcriptional activities of the microbial consortium living with the marine nitrogen-fixing cyanobacterium *Trichodesmium* reveal potential roles in community-level nitrogen cycling. Appl Environ Microbiol 84:e02026-17. https://doi.org/10.1128/AEM.02026-17.

Editor Harold L. Drake, University of Bayreuth

Copyright © 2017 American Society for Microbiology. All Rights Reserved.

Address correspondence to David A. Hutchins, dahutch@usc.edu.

Downloaded from http://aem.asm.org/ on December 15, 2017 by USC Norris Medical Library

Trichodesmium is a globally distributed nitrogen-fixing genus of cyanobacteria common in warm oligotrophic oceans. They provide a substantial proportion of new nitrogen (N) to N-limited systems, thereby helping to fuel primary productivity (1–3). As a keystone organism in major marine elemental cycles, *Trichodesmium* has been the focus of studies probing its responses to ongoing rapid changes in the global oceans, such as changes in temperature, pH, and CO_2 (e.g., see references 4–9). One common response of *T. erythraeum* strain IMS101 to elevated CO_2 has been increased N fixation rates (8, 10, 11), yet curiously, these do not coincide with increased transcripts or proteins of the nitrogenase complex responsible for N fixation (8, 12). Understanding the controls on one of the greatest sources of new marine N is essential for the modeling of global biogeochemical cycles and the assimilation of CO_2 in the present and future oceans.

Trichodesmium often forms large macroscopic colonies comprising tens to hundreds of filaments, each composed of tens to hundreds of cells (1). Like many other aggregating or relatively large primary producers (particularly algae), these colonies act as nutrient-rich substrates that harbor a diverse microbial community (13–17). While the interactions occurring between *Trichodesmium* and its associated epibionts have long been recognized as likely being important for the host, epibionts, or both (18–20), the extent to which they modulate host physiology and N fixation remains largely unknown. Moreover, *Trichodesmium* is difficult to maintain in culture (19, 21), and it has been suggested this may be due to the existence of obligate dependencies of the host on its associated members (16, 19, 22, 23).

Attempts to establish stable axenic cultures of *Trichodesmium* have been unsuccessful, perhaps because such a stable cohabitation of organisms has led to complex, interdependent cooperative interactions (24, 25). One possible example of such an interaction within *Trichodesmium* consortia involves the host's critical role of N fixation. The nitrogenase complex is inhibited by molecular oxygen (O_2), and *Trichodesmium* has long been considered enigmatic as it carries out N fixation while contemporaneously performing O_2-evolving photosynthesis (1, 19, 22, 26). Some studies have implicated temporal and spatial segregations (27 and reviewed in reference 28), but another proposed mechanism for this capability involves a cascade of interwoven interactions between *Trichodesmium* and its associated community. This hypothesis suggests that host exudation of organic carbon supports the growth of associated heterotrophs and fuels respiration, resulting in microenvironments of decreased O_2 concentrations. These suboxic microenvironments then serve as havens of host N fixation, ultimately benefiting the colony as a whole (19, 22, 29–32). In the context of solely photosynthetic carbon fixation, aerobic bacteria living in association with algae have indeed been shown to stimulate host growth via O_2 consumption, creating conditions more conducive to photosynthesis (33). In *Trichodesmium*, microelectrode measurements of primarily "raft"-type colonies revealed decreased intracolony O_2 levels correlating with increased N fixation under steady light (30), while another recent study working with "puff"-type colonies found decreased O_2 concentrations within colony cores only in darkness (11). Though colony size seems to be one factor responsible for the generation of O_2-depleted microenvironments, it remains unknown if colony morphology is also a component, and direct evidence supporting the role of the associated community is lacking.

Characterizations of the associated microbial communities of *Trichodesmium* organisms and other photoautotrophs have revealed consistently observed groups of major taxa, and notably, communities distinct from those in the surrounding seawater (16, 34–36). At a broad taxonomic level, commonly associated organisms include members of *Alphaproteobacteria*, *Gammaproteobacteria*, and *Bacteroidetes* taxa (14–17, 34, 36–40). While there is evidence supporting host-specific associations at finer taxonomic resolutions within these large clades (23, 41–43), there are also likely underlying general lifestyle and functional characteristics responsible for these broad trends in photoautotroph-heterotroph associations (e.g., host carbon and/or nitrogen fixation, epibiont traits for particle-association, copiotrophic/opportunistic lifestyles, etc.) (37, 44). Deciphering the

Downloaded from http://aem.asm.org/ on December 15, 2017 by USC Norris Medical Library

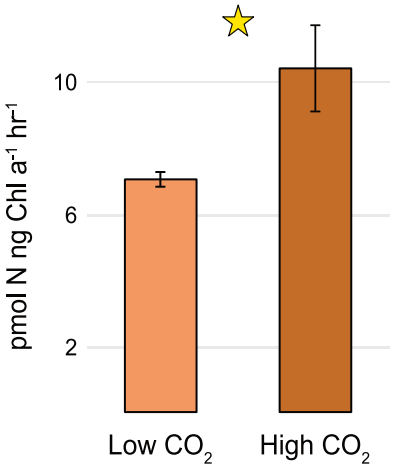

FIG 1 Host (IMS101) demonstrates significantly increased nitrogen fixation rates under elevated CO_2 concentrations ($n = 3$; star, $P < 0.05$ by *t* test).

activities occurring within these interconnected communities is essential to understanding the net biogeochemical contributions from any host that exists in perpetual association with other organisms (45).

Here, we present detected transcriptional and proteomic data of associated communities of *T. erythraeum* strain IMS101 following 8 years of selection under two distinct CO_2 concentrations. The enrichment cultures in this study harbor microbial communities similar to those found in association with natural *Trichodesmium* populations (17), suggesting that they can serve as a window into the complex network of interactions occurring between the cyanobacterium and its epibionts. Accordingly, we examined 3 primary questions. (i) What role might the community play in N cycling? (ii) Could increased community respiration be facilitating the corresponding increased rates of host N fixation under elevated CO_2? (iii) Can transcriptional and proteomic profiles help define the distinct ecological niches of the major taxa commonly associated with *Trichodesmium*?

RESULTS AND DISCUSSION

The cultures utilized in this study were previously split from one IMS101 cell line and maintained under two CO_2 concentrations (low [ambient 380 to 400 μatm] and high [800 μatm]) for ~8 years (~1,200 to 1,700 generations [8]). Cultured strain IMS101 usually grows as individual filaments—macroscopic millimeter-long chains of hundreds of cells—rather than as aggregated colonies. In the open ocean, it has been observed that most (>80%) of the total *Trichodesmium* biomass is present as filaments rather than colonies (46–49). This filamentous lifestyle may sustain a broad baseline distribution of *Trichodesmium* between the more episodic (but more frequently sampled) bloom events. Both filamentous and colonial morphologies of *Trichodesmium* host similar dominant microbial communities (17), albeit in differing organismal relative abundance, but little is known about how the ecophysiology of the entire system (host and epibionts) changes with and/or drives these structural differences.

As has been reported in numerous studies (8–12) and confirmed here, IMS101 demonstrates significantly increased N fixation rates under elevated CO_2 (Fig. 1). The data presented here are following 8 years of selection, but the same phenotype has been observed after just weeks (8, 10). Curiously, these increased rates were not accompanied by increases in relevant transcripts or proteins (8, 9), an important reminder that gene expression levels, and even levels of protein products, tell us about their respective levels only and may or may not be reflected in actual physiological and biogeochemical responses. Thus, transcriptomics and proteomics are useful tools for hypothesis generation rather than hypothesis confirmation. Deep metatranscriptomic

Downloaded from http://aem.asm.org/ on December 15, 2017 by USC Norris Medical Library

sequencing of these enrichment cultures (~95 million reads/sample) (see Table S1 in the supplemental material) enabled an investigation of the transcriptional activity of IMS101's associated microbial community. As enrichments of strain IMS101, the total RNA pools recovered were dominated by *Trichodesmium*, which was the source of ~75% of total reads (see Table S1 in the supplemental material). However, the depth of sequencing achieved recovered an average of ~15 million reads per sample originating solely from the associated community (Table S1). While the transcriptional activities and physiology of IMS101 from these experiments are integrated here in a relevant context; we focus here primarily on *Trichodesmium*'s associated community.

***De novo* reference library construction and taxonomic composition.** A *de novo* coassembly of multiple metatranscriptomes recovered from these cultures was performed and subsequently identified coding sequences (CDSs) were utilized as our reference library for recruiting metatranscriptomic reads from each individual sample and for proteomic analysis. Of ~45,000 CDSs derived solely from the IMS101-associated community, taxonomy was assigned to ~33,000 (~73%) (see Table S2). The vast majority of CDSs were identified as bacterial in origin (93%) (see Fig. S1 and Table S2). Consistent with (i) the only available environmental *Trichodesmium* metatranscriptome (36), (ii) a clone-library study of open-ocean *Trichodesmium* colonies (16), and (iii) analysis of these specific cultures and other laboratory-maintained and environmental samples (17), our *de novo* reference library was dominated by the major bacterial taxa *Bacteroidetes*, *Cyanobacteria*, *Alphaproteobacteria*, and *Gammaproteobacteria* (Fig. S1 and S2). As has been reported previously (16, 36), within these broad taxonomic clades were populations distinct from typical planktonic microbial communities, including the conspicuous absence of major taxa such as SAR11 and any *Archaea*. In our reference library, *Bacteroidetes* were predominantly composed of members of the *Saprospiraceae* family, *Phaeodactylibacter xiamenensis*, and *Lewinella cohaerens*, whereas *Cyanobacteria* CDSs were sourced almost entirely from *Synechococcus* spp., *Alphaproteobacteria* predominantly included members of the orders *Rhodobacterales* and *Rhizobiales*, and *Gammaproteobacteria* were dominated by the order *Alteromonadales*. The relatively few eukaryotic CDSs recovered were predominantly fungi and algae (Viridiplantae), consistent with observations of *Trichodesmium* in the open ocean (36). Importantly, the associated communities from these specific cultures have been shown to be environmentally relevant (17).

Utilizing this custom-built reference library, we employed two distinct approaches in the characterization of gene expression within these IMS101-associated communities. First, we functionally analyzed the global metatranscriptome around *Trichodesmium* as a collective unit irrespective of taxonomy (i.e., without consideration of "who" was doing "what"). This enabled an overall assessment of which predominant metabolic activities may be ultimately influencing the host's microenvironment. We then focused on the transcriptional activity of each major taxonomic group to investigate the potential for unique ecological niches of these consistently cooccurring and stable associations (16, 17, 36).

We additionally reanalyzed recently published proteomic data from these same samples using our *de novo* community reference library (12). The depth of metatranscriptomic sequencing employed enabled a detailed analysis of the associated community's transcriptional activities despite the host's dominant relative abundance (Table S1), but the primary focus of the proteomics work was to characterize solely the host's proteome (12). As such, the resultant proteomic coverage precluded a comprehensive profiling of the associated community's global proteome. For this reason, we do not attempt to interpret the community-related proteomics data quantitatively. However, integrating these data sets did enable the detection of many protein products of the discussed transcripts (presumably those protein products in greatest relative abundance). Accordingly, this serves as the foundation for a much-needed custom proteomics database for use with environmental samples of *Trichodesmium*.

Distinct global community transcriptional profiles according to CO_2 concentration. Hierarchical clustering and principle coordinates analysis of sample read

Downloaded from http://aem.asm.org/ on December 15, 2017 by USC Norris Medical Library

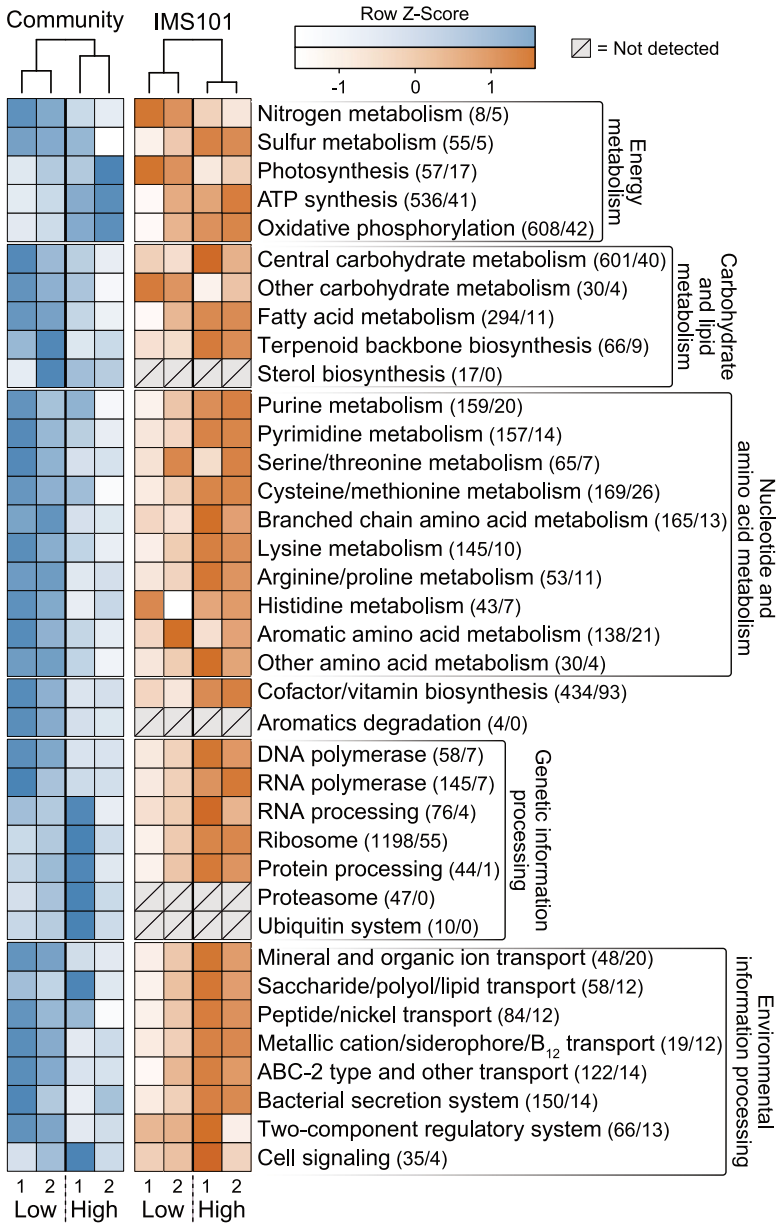

FIG 2 Relative expression levels for KEGG modules in *Trichodesmium* and its associated community under replicate low and high CO_2 treatments. Numbers in parentheses following module name represent the no. of CDSs identified in the community/no. identified in *Trichodesmium*.

recruitment to our CDSs both revealed distinct clusters of global community transcriptional activity corresponding to each treatment (see Fig. S3). Thus, the community exhibited unique transcriptional profiles with respect to CO_2 concentration. To probe these community-level transcriptional differences, CDSs were functionally annotated with Kyoto Encyclopedia of Genes and Genomes (KEGG) orthologs ([KOs] 50). This resulted in ~14,000 CDSs assigned to KOs (~30% of the total) (Table S2), which were then collapsed into KEGG modules (KEGG "modules" represent organized functional units of genes) (Table S3). To examine the associated community's activities in the context of its host, IMS101's genes were retrieved from Integrated Microbial Genomes (IMG [51]) and processed in the same manner. This perspective revealed an overall trend wherein under low CO_2, the associated community more evenly distributed its transcriptional pool across a broad range of modules, while IMS101's transcriptional allocation was primarily invested in nitrogen metabolism and photosynthesis (Fig. 2).

Downloaded from http://aem.asm.org/ on December 15, 2017 by USC Norris Medical Library

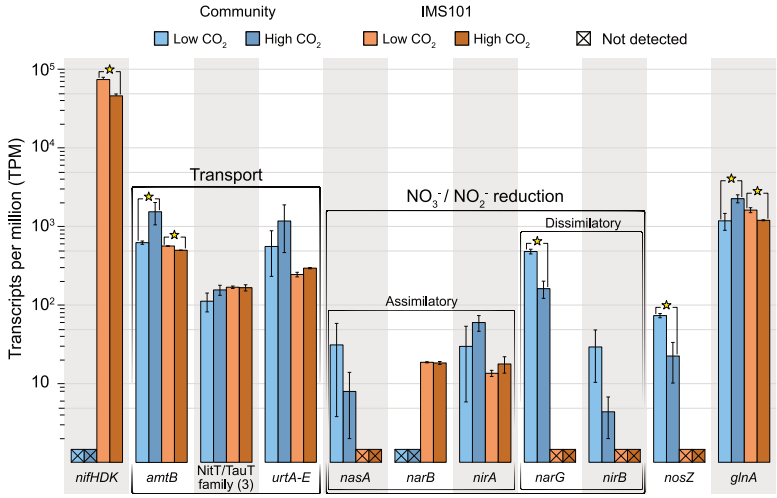

FIG 3 Nitrogen cycling transcript expression levels in *Trichodesmium* (orange and brown bars) and in its associated community considered as a whole (blue bars) under low (ambient) and high (800 μatm) CO_2. Stars indicate significance at a 0.05 threshold by 2-tailed *t* tests. *nifHDK*, nitrogenase; *amtB*, ammonium transport; NitT/TauT, nitrate/nitrite/taurine transport; *urtA-E*, urea transport; *nasA-narB*, assimilatory nitrate reductase; *nirA*, assimilatory nitrite reductase; *narG*, dissimilatory nitrate reductase; *nirB*, dissimilatory nitrite reductase; *nosZ*, nitrous oxide reductase; *glnA*, glutamine synthesis.

Contrasting with this, under high CO_2, the community as a whole allotted relatively more transcripts to a fewer number of processes, such as photosynthesis, ATP synthesis, and oxidative phosphorylation (potentially suggesting increased growth rates and/or overall community activity), whereas *Trichodesmium*'s transcript pool shifted to being more evenly distributed across a broad range of modules and less concentrated on photosynthesis and nitrogen metabolism (Fig. 2). This shift in nitrogen metabolism was directly driven by significant decreases in host nitrogenase transcripts under elevated CO_2 (*nifHDK*, $P = 0.04$, 2-tailed *t* test) (see Table S4), which actually coincides with significant increases in measured nitrogen fixation rates as noted above (Fig. 1). As these treatments had diverged under their respective CO_2 concentrations for 8 years prior to sampling, the global transcriptional clustering patterns (Fig. S3) and the presence of trends even at this coarsely resolved level of KEGG modules (Fig. 2) speak to the functional robustness of the systems.

Community nitrogen cycling. Given *Trichodesmium*'s significance as a source of fixed N to otherwise nitrogen-starved systems and the dearth of information regarding the role of its perpetually present associated community in this process, we specifically investigated the known primary genes involved in nitrogen cycling in both our host and the associated community. Though other nitrogen-fixing microbes have been seen with *Trichodesmium* in the open ocean (19, 52), the only detected N fixation genes in our enrichments were sourced from IMS101 (Fig. 3), indicating the measured N fixation rates were solely the result of the host. Interestingly, transcripts for dissimilatory nitrate/nitrite reduction were detected (*narG-nirB*), as well as the final step of denitrification (*nosZ*), which catalyzes the conversion of nitrous oxide (N_2O) into N_2 (Fig. 3); the expression of such genes is expected to be induced only under conditions where oxygen is depleted. Such anaerobic N transformations have recently been demonstrated to occur within the associated communities of another marine cyanobacterial diazotroph, *Nodularia spumigena* (53), and they have been proposed to occur within *Trichodesmium* colonies as well (54). Klawonn et al. additionally identified anoxic interior cores of the millimeter-sized colonies extending to ~5% of their total size, even when suspended in 100% air-saturated water (53), similar to observed decreases in O_2 concentration near the core of *Trichodesmium* colonies (30). Facultative anaerobes are thought to have an advantage when it comes to particle-associated lifestyles due to such microenvironments (55), and moreover, O_2 concentrations undergo rapid changes

Downloaded from http://aem.asm.org/ on December 15, 2017 by USC Norris Medical Library

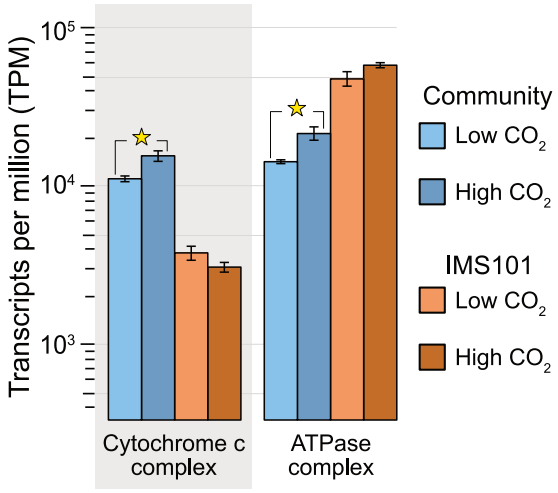

FIG 4 Expression levels of the cytochrome *c* and ATPase complexes involved in oxidative phosphorylation. The transcripts were enriched in the community under elevated CO_2. Stars indicate significance at a 0.05 threshold by 2-tailed *t* tests.

(within minutes) from supersaturated to subsaturated within *Trichodesmium* colonies transitioned from light to darkness (30). Though as noted above, in culture, IMS101 tends to grow as filaments not colonies. The generation of anaerobic microenvironments conducive to such N transformations may also occur within exopolysaccharide matrices (32), the production of which appears to be an active process in our system (discussed below), but it is unclear if these detected community N-reducing expression levels are simply due to their constitutive expression. Regardless, due to the similarity of the microbial communities between these cultures and those in natural colonies, and given the recent insights into these processes in *N. spumigena* (53), the potential implications remain the same. It is possible for aggregating cyanobacterial diazotrophs to harbor denitrifying facultative anaerobes, which raises interesting questions regarding N cycling within *Trichodesmium* colonies.

With respect to CO_2 treatment, as a whole, there was a restructuring of the associated community's global N-related transcriptional profiles (Fig. 3), coinciding with the increased host N fixation rate under elevated CO_2 (Fig. 1). Ammonium transport transcripts were significantly increased in the associated community but were less abundant in *Trichodesmium* under high CO_2 (Fig. 3, *amtB*). Also significantly enriched in the associated community but decreased in the host in the high CO_2 treatment were transcripts for glutamine biosynthesis (*glnA*), one of the primary pathways that incorporates new N into biomass. It is possible these community shifts reflect corresponding increases in growth rates or overall cellular metabolisms in response to the increased availability of fixed N supplies.

Increased community respiration under elevated CO_2 concentrations. As discussed above, associated microbial respiration could benefit the *Trichodesmium* host by aiding in the generation of microenvironments of lower O_2 concentrations, thereby relieving oxic inhibition of N fixation (19, 22, 29–32). In the context of our experiment, this is one possible explanation for the increased rates of host N fixation under elevated CO_2 (Fig. 1) despite the decreases in nitrogenase transcripts (Fig. 3) and proteins (8, 9). To investigate this possibility through the lens of community transcriptional expression, we examined the cytochrome *c* oxidase and ATPase complexes involved in respiration and oxidative phosphorylation.

Trichodesmium's associated community as a whole was indeed found to be dedicating a significantly increased proportion of its transcripts toward these processes under elevated CO_2 (Fig. 4). As these cell lines had diverged for ~8 years under these distinct CO_2 concentrations, these results also suggest this to be a stable state. It is possible that increased community respiration led to an increase in microenvironments

Downloaded from http://aem.asm.org/ on December 15, 2017 by USC Norris Medical Library

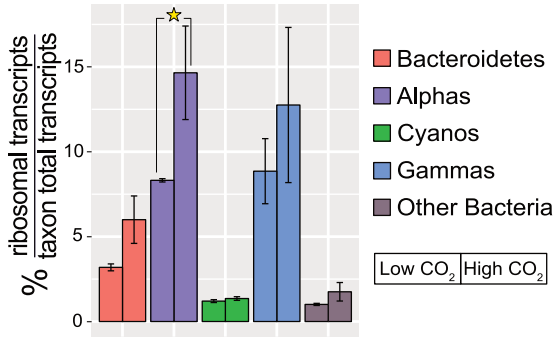

FIG 5 Taxon allocation of ribosomal protein transcripts relative to their respective transcript pools indicates possible increased growth rates under elevated CO_2. Stars indicate significance at a 0.05 threshold by 2-tailed *t* tests.

of lower O_2 concentrations, thereby lessening the effective O_2 inhibition of the nitrogenase enzyme and ultimately facilitating the increased IMS101 N fixation rates consistently observed under elevated CO_2 (8, 10, 56). To our knowledge, these are the first data specifically supporting a decades-old hypothesis of what may be a fundamental mechanism (the associated community's influence) involved with *Trichodesmium*'s seemingly enigmatic N fixation strategy.

One potential correlating trait of higher respiration rates is increased growth rates. While there is no way to directly quantify growth rates of the taxa associated with our IMS101 enrichments via the available metatranscriptomic data, the transcriptional investment in ribosomal proteins (i.e., the proportion of transcripts of the total transcript pool) has been shown to correlate with growth (57) and has been used as a quantitative metric of activity from metatranscriptomic data (58). An analysis of our major associated bacterial clades did reveal a greater proportion of ribosomal transcripts under elevated CO_2 (Fig. 5). This finding, along with an increase in glutamine synthesis transcripts (Fig. 3), suggests that increased growth rates may correlate with the enrichment of respiration transcripts under increased CO_2 (Fig. 4). To note, eukaryotes were not considered in this ribosomal protein analysis due to the overabundance of confounding ribosomal transcripts originating from chloroplasts and mitochondria (see Fig. S4).

Transcriptional support for distinct ecological niches within *Trichodesmium* consortia. Members of *Bacteroidetes*, *Cyanobacteria*, *Alphaproteobacteria*, and *Gammaproteobacteria* taxa commonly cooccur in natural *Trichodesmium* populations as well as in stable associations in all laboratory enrichments examined thus far (16, 17, 36). This sustained cohabitation spanning years in laboratory enrichments without added fixed carbon or nitrogen (17) and the fact that there are no axenic cultures of *Trichodesmium* support the idea that there is a persistent network for nutrient cycling occurring between the host and the community. Given this and the conserved association of these major taxa with *Trichodesmium* and other photoautotrophs, we examined the transcriptional profiles of these groups to identify potentially distinct ecological niches.

While investigating the relative contributions of various phylogenetic groups to the degradation of the multifarious milieus of dissolved organic matter (DOM) in aquatic environments, Cottrell and Kirchman (59) noted that while no individual major taxon can metabolize all forms of DOM, an assemblage comprising representatives from just three groups—*Bacteroidetes*, *Alphaproteobacteria*, and *Gammaproteobacteria*—likely could. This is because even at this relatively broad level of taxonomic resolution, there are still clear distinguishing characteristics of these clades (59), and it is likely these differences are what ultimately underlie the distinct ecological roles that enable and encourage the cooccurrence of these major taxa not only within *Trichodesmium* consortia but also adhered to particles (60) and algae (37).

We contrasted these groups, along with *Cyanobacteria*, by collapsing individual taxon CDSs into KOs and normalizing to that taxon's total number of functionally

Downloaded from http://aem.asm.org/ on December 15, 2017 by USC Norris Medical Library

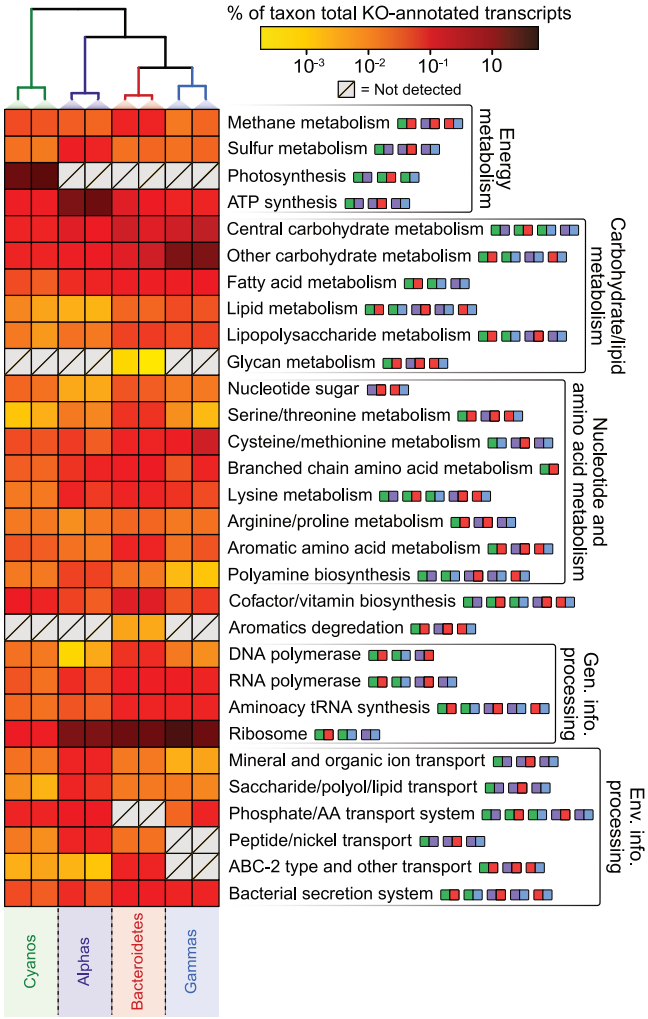

FIG 6 Relative expression of select KEGG modules for each of the 4 major taxa commonly seen in association with *Trichodesmium* under ambient CO_2. Values are percentages of total transcripts mapped to KO-annotated CDSs. Colored pairings denote statistically different expression levels between two taxa: green, *Cyanobacteria* (predominantly *Synechococcus*); purple, *Alphaproteobacteria* (*Rhodobacterales* and *Rhizobiales*); red, *Bacteroidetes* (*Phaeodactylibacter xiamenensis*); blue, *Gammaproteobacteria* (*Alteromonadales*). See Fig. S6 in the supplemental material for all treatments.

annotated transcripts. Hierarchical clustering and ordination both revealed clear groupings by taxa (see Fig. S5), and many of the primary driving forces behind this separation became apparent when CDSs were grouped into KEGG modules (Fig. 6) and upon comparing the most highly expressed genes from each taxon (Table 1; see also Table S6); for clarity, only the ambient CO_2 replicates are presented in Fig. 6, but the trends were similar irrespective of CO_2 concentration (see Fig. S6). These analyses provide a foundation to begin elucidating some of the distinct roles of these major taxa within the *Trichodesmium* consortium. While the current work is focused on *Trichodesmium*, the distinctions described here for these major taxa are potentially relevant to other photoautotroph-heterotroph interactions as well, due to the functional redundancies at the level of resolution being considered.

***Bacteroidetes*.** Marine *Bacteroidetes* are mostly known for their particle-associated lifestyles (61), high molecular-weight-compound degradation capabilities (62), and greater-than-typical abundance of genes related to exopolysaccharide production and adhesion (63). Specifically, they appear to possess 2 to 3 times more glycosyl transferase genes per million base pairs than *Alphaproteobacteria* and *Gammaproteobacteria*. Glycosyl transferases are typically outer membrane enzymes involved in the generation

Downloaded from http://aem.asm.org/ on December 15, 2017 by USC Norris Medical Library

TABLE 1 Summary of the 20 most highly expressed genes in each taxonomic group under ambient CO_2[a]

Category	Most highly expressed gene(s)			
	Bacteroidetes	*Alphaproteobacteria*	*Gammaproteobacteria*	*Cyanobacteria*
Transcription	***rpoE***, *rpoD*, ***rpoB***, ***rpoC***	*carD*	*rpoS*, *cspA*, *rpoA*, *rpoH*	
Translation	***tuf***, ***fusA***		*tuf*, *fusA*	
Ribosomes	*rpmE*, *rpsU*, *rpsP*	*rplN*, *rpsL*, *rplB*, *rpsU*	*rmf*, *rplM*	
Chaperones	*clpC*	*dnaK*	*pspA*	*osmC*
RNA methyltransferase			*trmD*	
RNA phosphotransferase				*kptA*
Glutamine/glutamate biosynthesis	*gltB*	***glnA***	*glnA*	
Fatty acid biosynthesis	*acpP*	*acpP*		
ATP biosynthesis		*atpA*, *atpD*		
Nucleotide biosynthesis				*ndk*, *nrdG*
Heme biosynthesis	*hemD*			
Glyoxylate/dicarboxylate metabolism			*aceA*, *icd*, *atoB*	
Oxidative phosphorylation	***coxA***	*nuoC*, *petB*, *nuoL*		
Photosynthesis				***cpeB***, ***apcB***, ***cpeA***, *petD*, ***petF***
Photosystem I				*psaA*, ***psaB***, *psaL*, *psaD*
Photosystem II				*psbA*, *psbC*, ***psbB***, *psbO*
Two-component regulation	*lytT*			
Nitrogen regulation			*glnK*	
General transport	***exbB***, ***tonB***			
Amino acid transport		***aapJ-bztA***		
Peptide transport		***ABC.PE.S***		
Urea transport				***urtA***
Ammonium transport				*amtB*
Chemotaxis	***motB***		*fliC*	
Glycosyltransferase	*glgA*			
Sulfur reduction		*sat-met3*		
Xenobiotics degradation		*E3.8.1.2*	*dhaA*	
Sphingolipid degradation		*aslA*		
Pyruvate dehydrogenase		*pdhB*		
Nitronate monooxygenase			*npd*	
Methionine degradation			*ahcY*	
Peroxidase			*ahpC*	

[a]Protein products for genes in boldface font were also detected. A more detailed table with TPM-normalized transcript counts and KO identifiers (Table S6) and protein spectral counts (Table S7) are presented in the supplemental material. *rpoDESH*, RNA polymerase sigma factors; *rpoABC*, RNA polymerase subunits; *carD*, transcriptional regulator, cold shock protein; *tuf-fusA*, elongation factors; *rpmE-rpsLPU-rplBMN*, ribosomal proteins; *rmf*, ribosome modulate factor; *clpC*, ATP-dependent protease; *dnaK*, RNA degradation; *pspA*, phage shock protein A; *osmC*, osmotically inducible protein; *trmD*, tRNA methyltransferase; *kptA*, putative RNA phosphotransferase; *gltB*, glutamate synthase; *glnA*, glutamine synthetase; *acpP*, acyl carrier protein; *atpAD*, F-type H^+-transporting ATPase subunits; *ndk*, nucleoside-diphosphate kinase; *nrdG*, anaerobic ribonucleoside-triphosphate reductase; *hemD*, uroporphyrinogen-III synthase; *aceA*, isocitrate lyase; *icd*, isocitrate dehydrogenase; *atoB*, acetyl-coenzyme A C-acetyltransferase; *coxA*, cytochrome *c* oxidase subunit I; *nuoCL*, NADH-quinone oxidoreductase subunits; *petB*, ubiquinol-cytochrome *c* reductase; *cpeAB*, phycoerythrin alpha/beta chains; *apcB*, allophycocyanin beta subunit; *petD*, cytochrome b_6-f complex subunit; *petF*, ferredoxin; *psaAB*, photosystem (PS)I chlorophyll *a* apoproteins; *psaDL*, PSI subunits; *psbA*, PSII reaction center protein; *psbBC*, PSII chlorophyll apoproteins; *psbO*, PSII oxygen-evolving enhancer protein; *lytT*, two-component response regulator; *glnK*, nitrogen regulatory protein; *aapJ-bztA*, amino acid transport; *ABC.PE.S*, peptide/nickel transport; *urtA*, urea transport; *amtB*, ammonium transport; *motB-fliC*, flagellar assembly proteins; *sat-met3*, sulfate adenylyltransferase; *E3.8.1.2*, 2-haloacid dehalogenase; *dhaA*, haloalkane dehalogenase; *aslA*, arylsulfatase; *pdhB*, pyruvate dehydrogenase; *npd*, nitronate monooxygenase; *ahcY*, adenosylhomocysteinase; *ahpC*, peroxiredoxin.

of exopolysaccharides for attachment (63). As with other environments, it is likely these traits may help define their unique role in the *Trichodesmium* consortium. Our *Bacteroidetes* clade was almost entirely dominated by members of the *Saprospiraceae* family, namely, *Phaeodactylibacter xiamenensis*, which is closely related to the more commonly known *Lewinella cohaerens*. Aptly, *P. xiamenensis* was isolated from the diatom *Phaeodactylum tricornutum* (64, 65), supporting the notion that, due to underlying similar properties, the phycospheres of both algae and cyanobacteria select for similar associated organisms, at least at this level of resolution. In this analysis, *Bacteroidetes* was the only major taxon that expressed transcripts found in KEGG's aromatics degradation module, highlighting the clade's distinguishing affinity for complex carbon-compound degradation (e.g., K10217), and in the glycan metabolism module, resulting almost entirely from transcripts involved in glycosyltransferase (K12666) (Fig. 6; see also Fig. S6). The enzymes encoded by these transcripts are known to be integral to exopolysaccharide production and are believed to be largely responsible for the taxon's typical particle-associated lifestyle (66, 67). This suggests members of *Bacteroidetes* may play a key role in *Trichodesmium* consortia by contributing to the extracellular matrix to the

Downloaded from http://aem.asm.org/ on December 15, 2017 by USC Norris Medical Library

benefits of the rest of the associated community and the host. This possible benefit again involves the generation and maintenance of low-oxygen microenvironments, as the exopolysaccharide matrix restricts oxygen diffusion (32). Additionally, *Bacteroidetes*' 5th and 6th most highly expressed genes (also detected in the proteome) were components of a TonB-dependent transporter system (*exbB* [K03561] and *tonB* [K03832]) (Table 1; see also Table S6). These are typically involved in the transport of complex compounds such as siderophores, vitamin B_{12}, and carbohydrates (68). While it is unclear what these transporters are acting on in this instance, *Bacteroidetes* were expending large proportions of their transcriptional energy on these genes (~2.5% of their total transcript pool), and their corresponding protein products were detected (Tables S6 and S7).

***Alphaproteobacteria*.** The class *Alphaproteobacteria* is often the dominant major taxon responsible for the consumption of the amino acid component of marine DOM (59). In our system, *Alphaproteobacteria* (dominated by members of the orders *Rhodobacterales* and *Rhizobiales*) did indeed demonstrate significantly higher relative gene expression for amino acid transport than the other major taxa (Fig. 6); notably, *Bacteroidetes*, proposed to be more involved in higher-molecular-weight DOM decomposition, had no detectable amino acid transport expression. This enriched *Alphaproteobacteria* expression was largely driven by L-amino acid ABC transport (K09969), which comprised the taxon's 6th most highly expressed transcripts and for which protein products were detected (Table 1; see also Table S6). *Alphaproteobacteria* contributed a significantly larger proportion of their transcriptional pool to KEGG's "peptide/nickel transport" module (Fig. 6) as well, due to the expression of genes involved in di- and tripeptide transport (K02035), for which proteins were also detected (Table 1; see also Table S7). Also uniquely highly expressed were transcripts involved in chlorinated cyclic and acyclic hydrocarbon degradation (E3.8.1.2 [K01560]) (Table 1; see also Table S6). These results support the notion that *Alphaproteobacteria* may hold an environmental niche space within *Trichodesmium* consortia based on their utilization of amino acids and chlorinated hydrocarbons. The group additionally exhibited relatively greater expression of sulfur metabolism genes (Fig. 6) due primarily to transcripts for sulfur oxidation (K17222 and K17227) and sulfate adenylyltransferase (*met3* [K00958]) (Table 1), which utilizes ATP to activate sulfate yielding adenylyl sulfate. The high relative expression of these was mostly driven by members of the order *Rhodobacterales*, which have been proposed to be a key group for sulfur cycling within algal bloom events (69); as such, they may play a role in sulfur biogeochemistry within *Trichodesmium* consortia as well.

***Gammaproteobacteria*.** *Gammaproteobacteria* CDSs mostly originated from the family *Alteromonadales* (Table S2). This clade is known for its ability to utilize varied and often cyclic carbon compounds such as *N*-acetylglucosamine (59, 70). It should be noted that *rpoS* and *pspA* (2 of the 20 most highly expressed genes within the *Gammaproteobacteria*) (Table 1) are both associated with stress responses to nutrient limitation and overall suboptimal conditions (71, 72), and as such it is possible members of this clade were stressed at the time of sampling. Nonetheless, this group was significantly enriched in KEGG's "other carbohydrate metabolism" module compared to the other 3 major taxa, with this module comprising just over 10% of the taxon's total recovered transcript pool (Fig. 6). This enrichment was driven largely by transcripts encoding enzymes for isocitrate lyase (*aceA* [K01637]) and malate dehydrogenase (*icd* [K01638]) (Fig. 6; see also Table S6). These are both involved in the glyoxylate cycle and have been noted as highly enriched from *Alteromonadales* in marine microcosm experiments upon the addition of naturally occurring DOM (73). Additionally, among the 20 most highly expressed genes for *Gammaproteobacteria* were those for haloalkane dehalogenase (*dhaA*) (Table 1; see also Table S6), which falls within KEGG's xenobiotic degradation module. Taken together, this provides a potential unique ecological niche for *Gammaproteobacteria* with regard to preferred organic substrates

Downloaded from http://aem.asm.org/ on December 15, 2017 by USC Norris Medical Library

compared to the other major taxa, as has been observed beyond the *Trichodesmium* consortium (59).

***Cyanobacteria*.** The transcriptional profile of *Cyanobacteria*, primarily dominated by *Synechococcus*, was least similar among the 4 major taxa (Fig. S5). This is certainly due in large part to the substantial proportion of their transcripts involved in photosynthesis (~25% of the taxon's total) (Fig. 6). Likely related to this large transcriptional investment in photosynthesis, the taxon also had a much smaller proportion of ribosomal protein transcripts than those of the other major taxa (~2%) (Fig. 5 and 6), similar to observations of *Synechococcus* marine samples (58). *Cyanobacteria* have been commonly found within natural *Trichodesmium* consortia (16, 17, 35, 36, 74), though in the current study of enrichment cultures with a relatively low abundance of associated community members, it is difficult to speculate on a potential unique role. However, the *Cyanobacteria* did allocate more of their transcriptional pool toward ammonium (*amtB*) and urea (*urtA*) transport (Table 1), with protein products also detected for the latter, and they have been noted in the environment to have higher expression of urease transcripts than other cooccurring plankton (58). It is likely *Cyanobacteria* in part further support the community as a whole via the provision of additional fixed carbon. This could be particularly beneficial under times of nutrient stress (e.g., with phosphorus or iron limitation), when *Trichodesmium*'s ability to fix nitrogen may be inhibited, as the entire consortium would then not be dependent solely upon the host for both carbon and nitrogen fixation.

Conclusions. The marine microbial loop largely drives surface ocean global biogeochemical cycling and supports most marine food webs. An ongoing shift in research focus has led to a growing interest in understanding microbial assemblages as a whole, as opposed to individual species in isolation (45, 75). *Trichodesmium* is currently only known to exist in nature and in culture in close association with other microbes. In our efforts to understand the significance of this widespread cyanobacterium for global nitrogen and carbon cycles, it is prudent that we consider the entire *Trichodesmium* host-epibiont assemblage as a whole. This metatranscriptomic analysis of *Trichodesmium*'s associated community sheds light on the potential for the community's involvement in system-wide nitrogen cycling (Fig. 7), including identifying transcripts involved in denitrification, a process recently reported to be occurring within the consortium of another marine cyanobacterial diazotroph (53). Additionally, we report significant increases in the relative abundance of community respiration-related transcripts corresponding to increases in host nitrogen fixation rates under elevated CO_2, a finding in alignment with a decades-old hypothesis of interorganismal interactions that, alongside fine-scale spatial and temporal segregation mechanisms (27), may be integral to *Trichodesmium*'s seemingly enigmatic nitrogen fixation strategy. Finally, given the consistent association of just a few major taxonomic clades within *Trichodesmium* consortia, as well as with other photoautotrophs, we can begin to delineate the unique ecological niche space they may be occupying by coupling their transcriptional profiles and detected proteins with their known distinguishing characteristics from the literature (Fig. 7). It is likely not a coincidence that the same few major taxonomic groups that can together degrade the majority of DOM in the marine environment (59) are also the dominant groups consistently found living in association with photoautotrophs such as *Trichodesmium*. A more comprehensive understanding of the interactions occurring between photoautotrophic hosts and their consistently associated consortia will improve our understanding of environmentally relevant elemental fluxes as well as offer insights into coevolution in the marine environment.

MATERIALS AND METHODS

Culturing conditions, physiology, and sampling. Detailed culturing conditions and sampling methods are presented elsewhere (8, 9) and in supplemental material. Cultures were continuously bubbled with prepared air-CO_2 mixtures (Praxair) to maintain stable CO_2 concentrations of ambient (380 to 400 μatm) or 800 μatm for ~8 years prior to sampling.

Trichodesmium growth rates were calculated from microscopic cell counts, and N fixation rates were measured using acetylene reduction as described previously (8). Growth and N fixation rates were

Downloaded from http://aem.asm.org/ on December 15, 2017 by USC Norris Medical Library

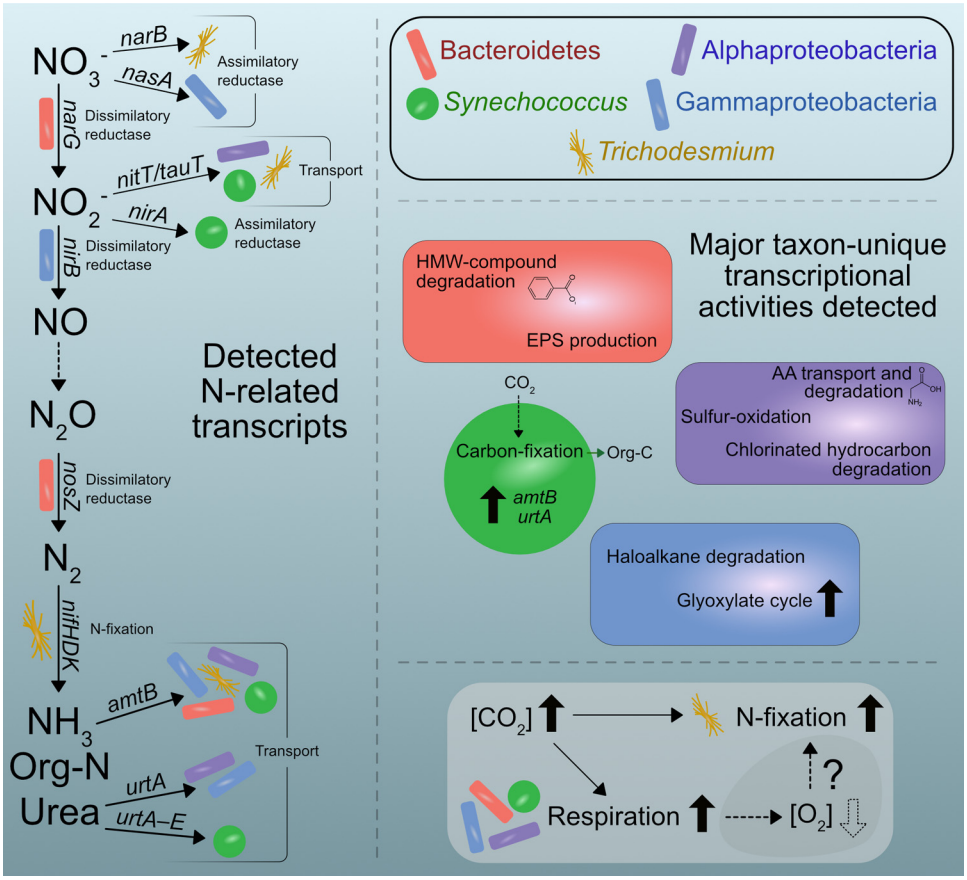

FIG 7 Summary schematic of nitrogen cycling-related transcripts and unique transcriptional activities detected in the major taxa commonly associated with *Trichodesmium* highlighting potential niche spaces. Gene names are as presented in the Fig. 3 legend; HMW, high molecular weight; EPS, exopolysaccharide; AA, amino acid.

measured during the middle of the photoperiod (76), and 200 ml of each sample was filtered (5-μm polycarbonate; Whatman) at the same time and then immediately flash-frozen and stored in liquid nitrogen until RNA extraction.

RNA extraction and sequencing. RNA was extracted from 2 randomly chosen samples of triplicate biological replicates, with steps performed to remove DNA and rRNA. Illumina Hi-Seq sequencing was performed yielding single-end 50-bp reads. Detailed information is presented in the supplemental material.

Proteomics. Protein extraction, tryptic digestion, and global proteome analysis were performed as previously described (8) and specified in the supplemental material. The search database contained translated CDSs from this study's coassembly (see Table S2 in the supplemental material) and *T. erythraeum* IMS101's genes retrieved from IMG (51). We identified 1,988 IMS101 proteins and 357 proteins from the associated community (see Table S7) based on 1,014 identified unique tryptic peptides (see Table S8).

Bioinformatics. Detailed information is presented in the supplemental material. Briefly, quality filtered reads were recruited to IMG's *T. erythraeum* IMS101 reference genome, and reads not mapped were considered to be derived from the associated community and processed further (Table S1). Following additional rRNA removal *in silico*, a coassembly of all samples was performed with Trinity v.2.4.0 (77), and Prodigal v.2.6.2 (78) was used to identify CDSs from the resultant assembled transcripts. These CDSs were utilized as our reference library for recruiting the reads from individual samples, and read counts were converted to transcripts per million (TPM) by first normalizing by CDS length and then by the size of the sample library corresponding to either the entire community or to individual taxonomic groups as noted. CDSs that failed to recruit greater than 1 TPM from any sample (as normalized to the entire community) were filtered out, leaving ~45,000 CDSs in all downstream analyses (Table S2). Amino acid sequences of CDSs were functionally annotated with Kyoto Encyclopedia of Genes and Genomes (KEGG) orthologs ([KOs] 50) and taxonomies were assigned via a two-step process comprising stand-alone BLASTp v.2.2.30 (79) run against NCBI's RefSeq protein database (80) and parsing of the outputs with MEGAN (81) v.6.7.6 using their default lowest common ancestor (LCA) algorithm.

All data visualizations were generated with RStudio v.1.0.136 (82) and R v3.4.1. Individual tests for statistical significance (i.e., those noted in Fig. 3, 4, and 5) were performed with standard two-tailed Student's *t* tests using a *P* value cutoff of 0.05, while tests for significance at the KEGG module level

Downloaded from http://aem.asm.org/ on December 15, 2017 by USC Norris Medical Library

(Fig. 6) were performed with analyses of variance (ANOVAs) followed by Tukey's honestly significant difference tests (HSDs) with an adjusted *P* value cutoff of 0.05.

Accession number(s). Raw RNA-seq fastq files are in NCBI's Gene Expression Omnibus (83), accessible through GEO Series accession number GSE94951. The sample accession numbers corresponding to the low and high CO_2 samples from this work are GSM2492342, GSM2492343, GSM2492344, and GSM2492345.

SUPPLEMENTAL MATERIAL

Supplemental material for this article may be found at https://doi.org/10.1128/AEM.02026-17.

SUPPLEMENTAL FILE 1, PDF file, 0.3 MB.

SUPPLEMENTAL FILE 2, XLSX file, 8.3 MB.

ACKNOWLEDGMENTS

M.D.L. thanks Chris Dupont for his advice and guidance.

Funding was provided by U.S. National Science Foundation grants BO-OCE 1260490 and OCE 1657757 to D.A.H., E.A.W., and F.-X.F., and NSF-OCE 1657766 and a Gordon and Betty Moore Foundation grant 3782 to M.A.S. Funding sources had no role in study design, data collection and interpretation, or the decision to submit the work for publication.

REFERENCES

1. Capone DG, Zehr JP, Paerl HW, Bergman B, Carpenter EJ. 1997. *Trichodesmium*, a globally significant marine cyanobacterium. Science 276:1221–1229. https://doi.org/10.1126/science.276.5316.1221.
2. Karl D, Michaels A, Bergman B, Capone D, Carpenter E, Letelier R, Lipschultz F, Paerl H, Sigman D, Stal L. 2002. Dinitrogen fixation in the world's oceans. Biogeochemistry 57:47–98. https://doi.org/10.1023/A:1015798105851.
3. Sohm JA, Webb EA, Capone DG. 2011. Emerging patterns of marine nitrogen fixation. Nat Rev Microbiol 9:499–508. https://doi.org/10.1038/nrmicro2594.
4. Sañudo-Wilhelmy SA, Kustka AB, Gobler CJ, Hutchins DA, Yang M, Lwiza K, Burns J, Capone DG, Raven JA, Carpenter EJ. 2001. Phosphorus limitation of nitrogen fixation by *Trichodesmium* in the central Atlantic Ocean. Nature 411:66–69. https://doi.org/10.1038/35075041.
5. Kustka AB, Sañudo-Wilhelmy SA, Carpenter EJ, Capone D, Burns J, Sunda WG. 2003. Iron requirements for dinitrogen- and ammonium-supported growth in cultures of *Trichodesmium* (IMS 101): comparison with nitrogen fixation rates and iron:carbon ratios of field populations. Limnol Oceanogr 48:1869–1884. https://doi.org/10.4319/lo.2003.48.5.1869.
6. Fu F-X, Zhang Y, Leblanc K, Sañudo-Wilhelmy SA, Hutchins DA. 2005. The biological and biogeochemical consequences of phosphate scavenging onto phytoplankton cell surfaces. Limnol Oceanogr 50:1459–1472. https://doi.org/10.4319/lo.2005.50.5.1459.
7. Webb EA, Jakuba RW, Moffett JW, Dyhrman ST. 2007. Molecular assessment of phosphorus and iron physiology in *Trichodesmium* populations from the western Central and western South Atlantic. Limnol Oceanogr 52:2221–2232. https://doi.org/10.4319/lo.2007.52.5.2221.
8. Hutchins DA, Walworth NG, Webb EA, Saito MA, Moran D, McIlvin MR, Gale J, Fu F. 2015. Irreversibly increased nitrogen fixation in *Trichodesmium* experimentally adapted to elevated carbon dioxide. Nat Commun 6:8155. https://doi.org/10.1038/ncomms9155.
9. Walworth NG, Lee MD, Fu F, Hutchins DA, Webb EA. 2016. Molecular and physiological evidence of genetic assimilation to high CO_2 in the marine nitrogen fixer *Trichodesmium*. Proc Natl Acad Sci U S A 113:E7367–E7374. https://doi.org/10.1073/pnas.1605202113.
10. Hutchins DA, Fu F-X, Zhang Y, Warner ME, Feng Y, Portune K, Bernhardt PW, Mulholland MR. 2007. CO_2 control of *Trichodesmium* N_2 fixation, photosynthesis, growth rates, and elemental ratios: implications for past, present, and future ocean biogeochemistry. Limnol Oceanogr 52:1293–1304. https://doi.org/10.4319/lo.2007.52.4.1293.
11. Eichner MJ, Klawonn I, Wilson ST, Littmann S, Whitehouse MJ, Church MJ, Kuypers MM, Karl DM, Ploug H. 2017. Chemical microenvironments and single-cell carbon and nitrogen uptake in field-collected colonies of *Trichodesmium* under different pCO_2. ISME J 11:1305–1317. https://doi.org/10.1038/ismej.2017.15.
12. Walworth NG, Fu F-X, Webb EA, Saito MA, Moran D, McIlvin MR, Lee MD, Hutchins DA. 2016. Mechanisms of increased *Trichodesmium* fitness under iron and phosphorus co-limitation in the present and future ocean. Nat Commun 7:12081. https://doi.org/10.1038/ncomms12081.
13. Nausch M. 1996. Microbial activities on *Trichodesmium* colonies. Mar Ecol Prog Ser 141:173–181. https://doi.org/10.3354/meps141173.
14. Fisher MM, Wilcox LW, Graham LE. 1998. Molecular characterization of epiphytic bacterial communities on charophycean green algae. Appl Environ Microbiol 64:4384–4389.
15. Sapp M, Schwaderer AS, Wiltshire KH, Hoppe HG, Gerdts G, Wichels A. 2007. Species-specific bacterial communities in the phycosphere of microalgae? Microb Ecol 53:683–699. https://doi.org/10.1007/s00248-006-9162-5.
16. Hmelo LR, Van Mooy BAS, Mincer TJ. 2012. Characterization of bacterial epibionts on the cyanobacterium *Trichodesmium*. Aquat Microb Ecol 67:1–14. https://doi.org/10.3354/ame01571.
17. Lee MD, Walworth NG, McParland EL, Fu F-X, Mincer TJ, Levine NM, Hutchins DA, Webb EA. 2017. The *Trichodesmium* consortium: conserved heterotrophic co-occurrence and genomic signatures of potential interactions. ISME J 11:1813–1824. https://doi.org/10.1038/ismej.2017.49.
18. Borstad GA. 1978. Some aspects of the occurrence and biology of *Trichodesmium* near Barbados. PhD dissertation. McGill University, Montreal, Quebec, Canada.
19. Paerl H, Bebout B, Prufert L. 1989. Bacterial associations with marine *Oscillatoria* sp. (*Trichodesmium* sp.) populations: ecophysiological implications. J Phycol 25:773–784. https://doi.org/10.1111/j.0022-3646.1989.00773.x.
20. O'Neil JM, Roman MR. 1992. Grazers and associated organisms of *Trichodesmium*, p 61–73. *In* Carpenter EJ, Capone DG, Rueter JG (ed), Marine pelagic *Cyanobacteria*: *Trichodesmium* and other diazotrophs, vol 362. Springer, Dordrecht, Netherlands.
21. Waterbury JB. 2006. The *Cyanobacteria*—isolation, purification and identification, p 1053–1073. *In* Dworkin M, Falkow S, Rosenberg E, Schleifer K-H, Stackebrandt E. The prokaryotes, 3rd ed. Springer-Verlag, New York, NY.
22. Zehr JP. 1995. Nitrogen fixation in the sea: why only *Trichodesmium*? p 333–364. *In* Joint I (ed), Molecular ecology of aquatic microbes, vol 38. Springer, Berlin, Germany.
23. Stevenson BS, Waterbury JB. 2006. Isolation and identification of an epibiotic bacterium associated with heterocystous *Anabaena* cells. Biol Bull 210:73–77. https://doi.org/10.2307/4134596.
24. Sachs JL, Hollowell AC. 2012. The origins of cooperative bacterial communities. mBio 3:e00099-12. https://doi.org/10.1128/mBio.00099-12.
25. Morris JJ, Lenski RE, Zinser ER. 2012. The black queen hypothesis:

Downloaded from http://aem.asm.org/ on December 15, 2017 by USC Norris Medical Library

evolution of dependencies through adaptive gene loss. mBio 3:e00036-12. https://doi.org/10.1128/mBio.00036-12.
26. Carpenter EJ, Price CC. 1976. Marine *Oscillatoria* (*Trichodesmium*): explanation for aerobic nitrogen fixation without heterocysts. Science 191:1278–1280. https://doi.org/10.1126/science.1257749.
27. Berman-Frank I, Lundgren P, Chen YB, Kolber Z, Bergman B, Falkowski P, Küpper H. 2001. Segregation of nitrogen fixation and oxygenic photosynthesis in the marine cyanobacterium *Trichodesmium*. Science 294:1534–1537. https://doi.org/10.1126/science.1064082.
28. Bergman B, Sandh G, Lin S, Larsson J, Carpenter EJ. 2013. *Trichodesmium*–a widespread marine cyanobacterium with unusual nitrogen fixation properties. FEMS Microbiol Rev 37:286–302. https://doi.org/10.1111/j.1574-6976.2012.00352.x.
29. Paerl HW, Kellar PE. 1978. Significance of bacterial-*Anabaena* associations with respect to N_2 fixation in freshwater. J Phycol 14:254–260. https://doi.org/10.1111/j.1529-8817.1978.tb00295.x.
30. Paerl HW, Bebout BM. 1988. Direct measurement of O_2-depleted microzones in marine *Oscillatoria*: relation to N_2 fixation. Science 241:442–445. https://doi.org/10.1126/science.241.4864.442.
31. Fay P. 1992. Oxygen relations of nitrogen fixation in cyanobacteria. Microbiol Rev 56:340–373.
32. Paerl HW, Pinckney JL. 1996. A mini-review of microbial consortia: their roles in aquatic production and biogeochemical cycling. Microb Ecol 31:225–247. https://doi.org/10.1007/BF00171569.
33. Mouget JL, Dakhama A, Lavoie MC, de la Noue J. 1995. Algal growth enhancement by bacteria: is consumption of photosynthetic oxygen involved? FEMS Microbiol Ecol 18:35–43. https://doi.org/10.1111/j.1574-6941.1995.tb00159.x.
34. Fandino LB, Riemann L, Steward GF, Long RA, Azam F. 2001. Variations in bacterial community structure during a dinoflagellate bloom analyzed by DGGE and 16S rDNA sequencing. Aquat Microb Ecol 23:119–130. https://doi.org/10.3354/ame023119.
35. Sheridan CC, Steinberg DK, Kling GW. 2002. The microbial and metazoan community associated with colonies of *Trichodesmium* spp.: a quantitative survey. J Plankton Res 24:913–922. https://doi.org/10.1093/plankt/24.9.913.
36. Hewson I, Poretsky RS, Dyhrman ST, Zielinski B, White AE, Tripp HJ, Montoya JP, Zehr JP. 2009. Microbial community gene expression within colonies of the diazotroph, *Trichodesmium*, from the Southwest Pacific Ocean. ISME J 3:1286–1300. https://doi.org/10.1038/ismej.2009.75.
37. Amin SA, Parker MS, Armbrust EV. 2012. Interactions between diatoms and bacteria. Microbiol Mol Biol Rev 76:667–684. https://doi.org/10.1128/MMBR.00007-12.
38. Bertrand EM, McCrow JP, Moustafa A, Zheng H, McQuaid JB, Delmont TO, Post AF, Sipler RE, Spackeen JL, Xu K, Bronk DA, Hutchins DA, Allen AE. 2015. Phytoplankton-bacterial interactions mediate micronutrient colimitation at the coastal Antarctic sea ice edge. Proc Natl Acad Sci U S A 112:9938–9943. https://doi.org/10.1073/pnas.1501615112.
39. Rouco M, Haley ST, Dyhrman ST. 2016. Microbial diversity within the *Trichodesmium* holobiont. Environ Microbiol 18:5151–5160. https://doi.org/10.1111/1462-2920.13513.
40. Frischkorn KR, Rouco M, Van Mooy BAS, Dyhrman ST. 2017. Epibionts dominate metabolic functional potential of *Trichodesmium* colonies from the oligotrophic ocean. ISME J 11:2090–2101. https://doi.org/10.1038/ismej.2017.74.
41. Lachnit T, Blumel M, Imhoff JF, Wahl M. 2009. Specific epibacterial communities on macroalgae: phylogeny matters more than habitat. Aquat Biol 5:181–186. https://doi.org/10.3354/ab00149.
42. Guannel ML, Horner-Devine MC, Rocap G. 2011. Bacterial community composition differs with species and toxigenicity of the diatom *Pseudo-nitzschia*. Aquat Microb Ecol 64:117–133. https://doi.org/10.3354/ame01513.
43. Sison-Mangus MP, Jiang S, Tran KN, Kudela RM. 2014. Host-specific adaptation governs the interaction of the marine diatom, *Pseudo-nitzschia* and their microbiota. ISME J 8:63–76. https://doi.org/10.1038/ismej.2013.138.
44. Van Mooy BAS, Hmelo LR, Sofen LE, Campagna SR, May AL, Dyhrman ST, Heithoff A, Webb EA, Momper L, Mincer TJ. 2012. Quorum sensing control of phosphorus acquisition in *Trichodesmium* consortia. ISME J 6:422–429. https://doi.org/10.1038/ismej.2011.115.
45. Wang H, Hill RT, Zheng T, Hu X, Wang B. 2016. Effects of bacterial communities on biofuel-producing microalgae: stimulation, inhibition and harvesting. Crit Rev Biotechnol 36:341–352. https://doi.org/10.3109/07388551.2014.961402.
46. Marumo R, Asaoka O. 1974. *Trichodesmium* in the East China Sea. Distribution of *Trichodesmium theibautii* during 1961-1967. J Mar Sci Technol 30:48–53.
47. Marumo R, Nagasawa S. 1976. Seasonal variation of the standing crop of the pelagic blue-green alga, *Trichodesmium* in the Kuroshio water. Bull Plankton Soc Japan 23:19–25. (In Japanese)
48. Letelier RM, Karl DM. 1996. Role of *Trichodesmium* spp. in the productivity of the subtropical North Pacific Ocean. Mar Ecol Prog Ser 133:263–273. https://doi.org/10.3354/meps133263.
49. Orcutt KM, Lipschultz F, Gundersen K, Arimoto R, Michaels AF, Knap AH, Gallon JR. 2001. A seasonal study of the significance of N_2 fixation by *Trichodesmium* spp. at the Bermuda Atlantic time-series study (BATS) site. Deep Res Part 2 Top Stud Oceanogr 48:1583–1608. https://doi.org/10.1016/S0967-0645(00)00157-0.
50. Kanehisa M, Sato Y, Kawashima M, Furumichi M, Tanabe M. 2016. KEGG as a reference resource for gene and protein annotation. Nucleic Acids Res 44:D457–D462. https://doi.org/10.1093/nar/gkv1070.
51. Markowitz VM, Mavromatis K, Ivanova NN, Chen IM, Chu K, Kyrpides NC. 2009. IMG ER: a system for microbial genome annotation expert review and curation. Bioinformatics 25:2271–2278. https://doi.org/10.1093/bioinformatics/btp393.
52. Momper LM, Reese BK, Carvalho G, Lee P, Webb EA. 2015. A novel cohabitation between two diazotrophic cyanobacteria in the oligotrophic ocean. ISME J 9:882–893. https://doi.org/10.1038/ismej.2014.186.
53. Klawonn I, Bonaglia S, Brüchert V, Ploug H. 2015. Aerobic and anaerobic nitrogen transformation processes in N2-fixing cyanobacterial aggregates. ISME J 9:1456–1466. https://doi.org/10.1038/ismej.2014.232.
54. Wyman M, Hodgson S, Bird C. 2013. Denitrifying *Alphaproteobacteria* from the Arabian Sea that express *nosZ*, the gene encoding nitrous oxide reductase, in oxic and suboxic waters. Appl Environ Microbiol 79:2670–2681. https://doi.org/10.1128/AEM.03705-12.
55. Ganesh S, Parris DJ, DeLong EF, Stewart FJ. 2014. Metagenomic analysis of size-fractionated picoplankton in a marine oxygen minimum zone. ISME J 8:187–211. https://doi.org/10.1038/ismej.2013.144.
56. Barcelos e Ramos J, Biswas H, Schulz KG, LaRoche J, Riebesell U. 2007. Effect of rising atmospheric carbon dioxide on the marine nitrogen fixer *Trichodesmium*. Global Biogeochem Cycles 21:1–6. https://doi.org/10.1029/2006GB002898.
57. Wei Y. 2001. High-density microarray-mediated gene expression profiling of *Escherichia coli*. J Bacteriol 183:545–556. https://doi.org/10.1128/JB.183.2.545-556.2001.
58. Gifford SM, Sharma S, Booth M, Moran MA. 2013. Expression patterns reveal niche diversification in a marine microbial assemblage. ISME J 7:281–298. https://doi.org/10.1038/ismej.2012.96.
59. Cottrell MT, Kirchman DL. 2000. Natural assemblages of marine proteobacteria and members of the *Cytophaga-Flavobacter* cluster consuming low- and high-molecular-weight dissolved organic matter. Appl Environ Microbiol 66:1692–1697. https://doi.org/10.1128/AEM.66.4.1692-1697.2000.
60. Pedrotti ML, Beauvais S, Kerros ME, Iversen K, Peters F. 2009. Bacterial colonization of transparent exopolymeric particles in mesocosms under different turbulence intensities and nutrient conditions. Aquat Microb Ecol 55:301–312. https://doi.org/10.3354/ame01308.
61. DeLong EF, Franks DG, Alldredge AL. 1993. Phylogenetic diversity of aggregate-attached vs. free-living marine bacterial assemblages. Limnol Oceanogr 38:924–934. https://doi.org/10.4319/lo.1993.38.5.0924.
62. Thomas F, Hehemann JH, Rebuffet E, Czjzek M, Michel G. 2011. Environmental and gut *Bacteroidetes*: the food connection. Front Microbiol 2:93. https://doi.org/10.3389/fmicb.2011.00093.
63. Fernandez-Gomez B, Richter M, Schuler M, Pinhassi J, Acinas SG, Gonzalez JM, Pedros-Alio C. 2013. Ecology of marine *Bacteroidetes*: a comparative genomics approach. ISME J 7:1026–1037. https://doi.org/10.1038/ismej.2012.169.
64. Chen Z, Lei X, Li Y, Zhang J, Zhang H, Yang L, Zheng W, Xu H, Zheng T. 2014. Whole-genome sequence of marine bacterium *Phaeodactylibacter xiamenensis* strain KD52, isolated from the phycosphere of microalga *Phacodactylum tricornutum*. Genome Announc 2:e01289-14. https://doi.org/10.1128/genomeA.01289-14.
65. Chen Z, Lei X, Lai Q, Li Y, Zhang B, Zhang J, Zhang H, Yang L, Zheng W, Tian Y, Yu Z, Xu H, Zheng T. 2014. *Phaeodactylibacter xiamenensis* gen. nov., sp. nov., a member of the family *Saprospiraceae* isolated from the marine alga *Phaeodactylum tricornutum*. Int J Syst Evol Microbiol 64:3496–3502. https://doi.org/10.1099/ijs.0.063909-0.
66. Hall-Stoodley L, Costerton JW, Stoodley P. 2004. Bacterial biofilms: from

Downloaded from http://aem.asm.org/ on December 15, 2017 by USC Norris Medical Library

the natural environment to infectious diseases. Nat Rev Microbiol 2:95–108. https://doi.org/10.1038/nrmicro821.
67. Bauer M, Kube M, Teeling H, Richter M, Lombardot T, Allers E, Würdemann CA, Quast C, Kuhl H, Knaust F, Woebken D, Bischof K, Mussmann M, Choudhuri JV, Meyer F, Reinhardt R, Amann RI, Glöckner FO. 2006. Whole-genome analysis of the marine Bacteroidetes '*Gramella forsetii*' reveals adaptations to degradation of polymeric organic matter. Environ Microbiol 8:2201–2213. https://doi.org/10.1111/j.1462-2920.2006.01152.x.
68. Noinaj N, Guillier M, Barnard TJ, Buchanan SK. 2010. TonB-dependent transporters: regulation, structure, and function. Annu Rev Microbiol 64:43–60. https://doi.org/10.1146/annurev.micro.112408.134247.
69. Zubkov MV, Fuchs BM, Archer SD, Kiene RP, Amann R, Burkill PH. 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. https://doi.org/10.1046/j.1462-2920.2001.00196.x.
70. Nikrad MP, Cottrell MT, Kirchman DL. 2014. Uptake of dissolved organic carbon by gammaproteobacterial subgroups in coastal waters of the west Antarctic peninsula. Appl Environ Microbiol 80:3362–3368. https://doi.org/10.1128/AEM.00121-14.
71. Brissette JL, Weiner L, Ripmaster TL, Model P. 1991. Characterization and sequence of the *Escherichia coli* stress-induced psp operon. J Mol Biol 220:35–48. https://doi.org/10.1016/0022-2836(91)90379-K.
72. Battesti A, Majdalani N, Gottesman S. 2015. Stress sigma factor RpoS degradation and translation are sensitive to the state of central metabolism. Proc Natl Acad Sci U S A 112:5159–5164. https://doi.org/10.1073/pnas.1504639112.
73. McCarren J, Becker JW, Repeta DJ, Shi Y, Young CR, Malmstrom RR, Chisholm SW, DeLong EF. 2010. Microbial community transcriptomes reveal microbes and metabolic pathways associated with dissolved organic matter turnover in the sea. Proc Natl Acad Sci U S A 107:16420–16427. https://doi.org/10.1073/pnas.1010732107.
74. Siddiqui PJA, Bergman B, Carpenter EJ. 1992. Filamentous cyanobacterial associates of the marine planktonic cyanobacterium *Trichodesmium*. Phycologia 31:326–337. https://doi.org/10.2216/i0031-8884-31-3-4-326.1.
75. Amin SA, Hmelo LR, van Tol HM, Durham BP, Carlson LT, Heal KR, Morales RL, Berthiaume CT, Parker MS, Djunaedi B, Ingalls AE, Parsek MR, Moran MA, Armbrust EV. 2015. Interaction and signalling between a cosmopolitan phytoplankton and associated bacteria. Nature 522:98–101. https://doi.org/10.1038/nature14488.
76. Walworth N, Pfreundt U, Nelson WC, Mincer T, Heidelberg JF, Fu F, Waterbury JB, Glavina del Rio T, Goodwin L, Kyrpides NC, Land ML, Woyke T, Hutchins DA, Hess WR, Webb EA. 2015. *Trichodesmium* genome maintains abundant, widespread noncoding DNA in situ, despite oligotrophic lifestyle. Proc Natl Acad Sci U S A 112:4251–4256. https://doi.org/10.1073/pnas.1422332112.
77. Haas BJ, Papanicolaou A, Yassour M, Grabherr M, Blood PD, Bowden J, Couger MB, Eccles D, Li B, Lieber M, Macmanes MD, Ott M, Orvis J, Pochet N, Strozzi F, Weeks N, Westerman R, William T, Dewey CN, Henschel R, Leduc RD, Friedman N, Regev A. 2013. *De novo* transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. Nat Protoc 8:1494–1512. https://doi.org/10.1038/nprot.2013.084.
78. Hyatt D, Locascio PF, Hauser LJ, Uberbacher EC. 2012. Gene and translation initiation site prediction in metagenomic sequences. Bioinformatics 28:2223–2230. https://doi.org/10.1093/bioinformatics/bts429.
79. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. 1990. Basic local alignment search tool. J Mol Biol 215:403–410. https://doi.org/10.1016/S0022-2836(05)80360-2.
80. O'Leary. 2016. Reference sequence (RefSeq) database at NCBI: current status, taxonomic expansion, and functional annotation. Nucleic Acids Res 44:D733–D745. https://doi.org/10.1093/nar/gkv1189.
81. Huson DH, Beier S, Flade I, Górska A, El-Hadidi M, Mitra S, Ruscheweyh HJ, Tappu R. 2016. MEGAN community edition - interactive exploration and analysis of large-scale microbiome sequencing data. PLoS Comput Biol 12:e1004957. https://doi.org/10.1371/journal.pcbi.1004957.
82. Racine J. 2012. RSTUDIO: a platform-independent IDE for R and Sweave. J Appl Econ 27:167–172. https://doi.org/10.1002/jae.1278.
83. Barrett T, Wilhite SE, Ledoux P, Evangelista C, Kim IF, Tomashevsky M, Marshall KA, Phillippy KH, Sherman PM, Holko M, Yefanov A, Lee H, Zhang N, Robertson CL, Serova N, Davis S, Soboleva A. 2013. NCBI GEO: archive for functional genomics data sets–update. Nucleic Acids Res 41:D991–D995. https://doi.org/10.1093/nar/gks1193.

Downloaded from http://aem.asm.org/ on December 15, 2017 by USC Norris Medical Library