AMERICAN SOCIETY FOR MICROBIOLOGY | mSystems

Microbial Ecology | Research Article

Taxonomic distribution of metabolic functions in bacteria associated with *Trichodesmium* consortia

Coco Koedooder,[1,2,3] Futing Zhang,[1,2] Siyuan Wang,[1,2] Subhajit Basu,[1,2,4] Sheean T. Haley,[5] Nikola Tolic,[6] Carrie D. Nicora,[6] Tijana Glavina del Rio,[7] Sonya T. Dyhrman,[5,8] Martha Gledhill,[9] Rene M. Boiteau,[10,11] Maxim Rubin-Blum,[3] Yeala Shaked[1,2]

AUTHOR AFFILIATIONS See affiliation list on p. 19.

ABSTRACT The photosynthetic and diazotrophic cyanobacterium *Trichodesmium* is a key contributor to marine biogeochemical cycles in the subtropical-oligotrophic oceans. *Trichodesmium* form colonies that harbor a distinct microbial community in comparison to the surrounding seawater. The presence of their associated bacteria can expand *Trichodesmium*'s functional potential and is predicted to influence the cycling of carbon, nitrogen, phosphorus, and iron (C, N, P, and Fe). To link the bacteria associated with *Trichodesmium* to key functional traits and elucidate how community structure can influence nutrient cycling, we characterized Red Sea *Trichodesmium* colonies using metagenomics and metaproteomics. Colonies harbored bacteria that typically associate with algae and particles, such as the ubiquitous *Alteromonas macleodii,* but also lineages specific to *Trichodesmium*, such as members from the order Balneolales. The majority of associated bacteria were auxotrophic for different vitamins, indicating their dependency on vitamin production by *Trichodesmium*. The associated bacteria carry functional traits including siderophore biosynthesis, reduced phosphorus metabolism, and denitrification pathways. The analysis supports *Trichodesmium* as an active hotspot for C, N, P, Fe, and vitamin exchange. In turn, *Trichodesmium* may rely on associated bacteria to meet its high Fe demand as several lineages synthesize photolabile siderophores (e.g., vibrioferrin, rhizoferrin, petrobactin) which can enhance the bioavailability of particulate Fe to the entire consortium. Collectively, the results indicate that *Trichodesmium* colonies provide a structure where these interactions can take place. While further studies are required to clarify the exact nature of these interactions, *Trichodesmium*'s reliance on particle and algae-associated bacteria and the observed redundancy of key functional traits likely underpins the resilience of *Trichodesmium* within an ever-changing global environment.

IMPORTANCE Colonies of the cyanobacteria *Trichodesmium* act as a biological hotspot for the usage and recycling of key resources such as C, N, P, and Fe within an otherwise oligotrophic environment. While *Trichodesmium* colonies are known to interact and support a unique community of algae and particle-associated microbes, our understanding of the taxa that populate these colonies and the gene functions they encode is still limited. Characterizing the taxa and adaptive strategies that influence consortium physiology and its concomitant biogeochemistry is critical in a future ocean predicted to have increasingly resource-depleted regions.

KEYWORDS *Trichodesmium*, metagenomes, associated bacteria, interactions, iron, siderophore, siderophore biosynthesis, vitamin B, denitrification, phosphonate, nitrogen fixation

Trichodesmium spp. are a globally relevant group of cyanobacteria that can form surface blooms visible from space (1). Owing to their high abundance and capability to fix both carbon (C) and nitrogen (N), they are often considered key players in the

Editor Jean-Baptiste Raina, University of Technology Sydney, Glebe, Australia

Address correspondence to Coco Koedooder, coco.koedooder@mail.huji.ac.il.

The authors declare no conflict of interest.

See the funding table on p. 19.

Received 26 July 2023
Accepted 21 September 2023
Published 2 November 2023

biogeochemical cycling of C and N within the oligotrophic tropical and subtropical oceans (2–5). A key trait underpinning the functionality of *Trichodesmium* is that its filaments can cluster together to form macroscopic colonies several millimeters in size (1–4 mm) (6, 7). There are several proposed benefits to colony formation including a reduction in grazing, facilitating microbial interactions, and enhancing iron (Fe) and phosphorous (P) uptake (8, 9). One particular advantage of colony formation that has received widespread interest within the scientific community is the ability of *Trichodesmium* colonies to coordinate the capture of dust deposited on the ocean surface (10). *Trichodesmium* has subsequently been reported to actively select Fe and P-rich particles and shuffle them to the core of the colony (11–14). Colony formation hereby adds an intriguing spatial component where *Trichodesmium* spp. can act as a biological hotspot not only for C and N but also for the uptake and recycling of key limiting nutrients such as Fe and P.

Trichodesmium colonies harbor a unique microbial consortium of epibiotic bacteria. Amplicon sequencing has shown these associated bacteria are distinct from populations in the surrounding oligotrophic surface waters (15–17). At the genetic level, the associated microbial community was found to enrich the total functional potential of the collective colony beyond the traits of N_2 and C fixation (18–20). By enhancing the genetic repertoire of metabolic functions, or traits, related to the nutrients C, N, Fe, and P, consortium members were predicted to influence the internal cycling of these nutrients (20–23). To understand the biogeochemical contributions of *Trichodesmium* within the global ocean, it is therefore critical to look beyond the metabolic functions present in *Trichodesmium* alone and toward the processes taking place within the larger microbial community. While metagenomic studies of *Trichodesmium* colonies in the Atlantic (19) and the Pacific Oceans (18, 23, 24) highlighted key functional traits within the consortium, only one study linked pathways to the associated bacteria from eight metagenome-assembled bins (19). The assembly of genomes did not represent the majority of sequencing reads, hindering an exploration of traits to their taxonomic affiliation (19). Technological advances in the field of genomics now enable us to revisit and expand on these previous studies. Characterizing how functional traits are distributed across different taxa is critical to understanding the role associated bacteria play in *Trichodesmium* physiology and its concomitant influence over the cycling of key resources such as C and N.

A key functional trait to explore within the *Trichodesmium* consortium through metagenomics is the biosynthesis of Fe-chelating siderophore molecules. Recent evidence from the Red Sea highlights the unique ability of *Trichodesmium* consortium members to acquire particulate Fe from dust through the production of siderophores (12, 13, 22, 25, 26). Siderophore production has been shown to aid the dissolution of particulate Fe from dust, enhancing its bioavailability (27, 28). While *Trichodesmium* is not known to synthesize siderophores (29), studies have suggested the presence of a putative siderophore uptake system in *T. erythraeum* IMS101 (30, 31). *Trichodesmium* is hereby predicted to take advantage of siderophore-producing bacteria to increase the bioavailability of particulate Fe to the colony as a whole (22, 28, 30, 32). In this manner, *Trichodesmium* interacts with its associated bacteria to meet the high Fe requirements needed for the processes of photosynthesis and N_2 fixation (33). This finding is particularly relevant in light of climate model simulations, which predict an increase in regional dust emissions in areas such as the Red Sea where *Trichodesmium* is known to occur (34, 35). While *Trichodesmium* blooms are a re-occurring phenomenon in this area (36), it represents an under-sampled region where the taxonomic composition of its consortium has not been explored genetically. Furthermore, the presence and taxonomic distribution of siderophore biosynthesis pathways within *Trichodesmium* colonies have not yet been interrogated in depth. Understanding which consortium members positively enhance the bioavailability of Fe within the *Trichodesmium* consortium will further help elucidate key interactions and community dynamics taking place within the colony.

In this study, we conducted an in-depth metagenomic study of *Trichodesmium* colonies collected from the Red Sea. To elucidate how the community structure influences the functional dynamics of the *Trichodesmium* consortium, we identified the different taxa that make up the *Trichodesmium* consortium and examined the presence of functional traits involved in nutrient cycling. We probed our data set for genes and pathways involved in the uptake and metabolism of N, P, Fe, and vitamins. In addition, we contextualize these findings by comparing them to previous metagenomic studies of *Trichodesmium* colonies obtained from the Atlantic and the Pacific oceans. Finally, we analyzed genes of interest to a proteomic data set from *Trichodesmium* colonies collected in parallel.

RESULTS AND DISCUSSION

A set of 52 non-redundant metagenomic-assembled genomes (MAGs) was obtained from *Trichodesmium* colonies, following quality filtering. Of these MAGs, 42 were of high quality (90% complete and <5% redundancy (Table S1) (37). The 52 MAGs accounted for ~80% of the total sample reads indicating that these MAGs sufficiently captured the diversity of the *Trichodesmium* consortium. The relative abundance of the different MAGs was consistent across the three samples (Fig. S1) with 71% ± 6 of the reads mapping to *T. thiebautii* MAG 52. The remaining MAGs collectively represented the associated bacteria of *T. thiebautii* MAG 52, in this analysis. In parallel to the metagenomic analysis, a proteomic data set matched to 1,253 protein sequences from the 52 different MAGs further probed the activity of *Trichodesmium* colonies within the Red Sea with ±90% of the total 14,074 spectral counts matching to *T. thiebautii* MAG 52 (Table S3).

Trichodesmium colonies host a diverse and flexible consortium of bacteria

Trichodesmium colonies consisted of a single *Trichodesmium* species (*T. thiebautii* MAG 52) and a diverse consortium of 51 associated bacteria spanning at least 10 different taxonomic orders (Fig. 1). Unlike a previous study (38), our samples did not include any non-diazotrophic *Trichodesmium* populations and the 51 MAGs hereby represent the associated bacteria of *T. thiebautii* MAG 52. The majority of the associated bacterial MAGs belonged to taxonomic orders known to be present in *Trichodesmium* colonies, including Rhodobacterales, Pseudomonadales, Balneolales, Rhodospirillales, Flavobacterales, Enterobacterales, Rhizobiales, and Sphingomonadales (15, 17–19).

We compared our 52 MAGs with MAGs assembled from previous metagenomic data sets of *Trichodesmium* colonies. To allow intercomparisons across data sets, samples were all processed using the same pipeline. In this manner, we were able to assemble an additional 29 MAGs from the Red Sea (38), and 19 MAGs from colonies collected in the Atlantic (19) and the Pacific Ocean (18, 19), respectively (Table S2). Of these MAGs, 11 were common between both Red Sea data sets. As these samples were obtained from two different seasons, these MAGs putatively represent core re-occurring members of the *Trichodesmium* consortium (Fig. S2). These 11 MAGs included typical particle- and algae-associated lineages (39–41) of *Marinobacter* sp. (MAG23; R-02) and *Alcanivorax* sp. (MAG 47; R-15) from the order Pseudomonadales; *Rhodobacter* sp. (MAG 28; R-29) and *Sulfitobacter* sp. (MAG29; R-10) from the order Rhodobacterales. *Alteromonas macleodii* (MAG 10), a ubiquitous bacterium often found in association with phototrophs (42) and previously isolated from *T. erythraeum* IMS101 cultures (43), was also detected in the previous Red Sea data set but did not reach the genome completeness threshold (>75%). *Roseibium aggregatum* (MAG 17), formerly known as *Labrenzia* sp., had previously been identified as a relevant denitrifying bacteria associated with *Trichodesmium* colonies in the Red Sea (44). When comparing our data set to those obtained from the Atlantic, the North and South Pacific oceans (18, 19, 23, 24, 38), we observed taxonomic similarities at the order level rather than at the species level. While most taxonomic orders re-occurred across all data sets, only one Balneolales genome (MAG 33) was present among all *Trichodesmium* metagenomic data sets (T-09; R-03) (Fig. S2). MAGs assembled from these

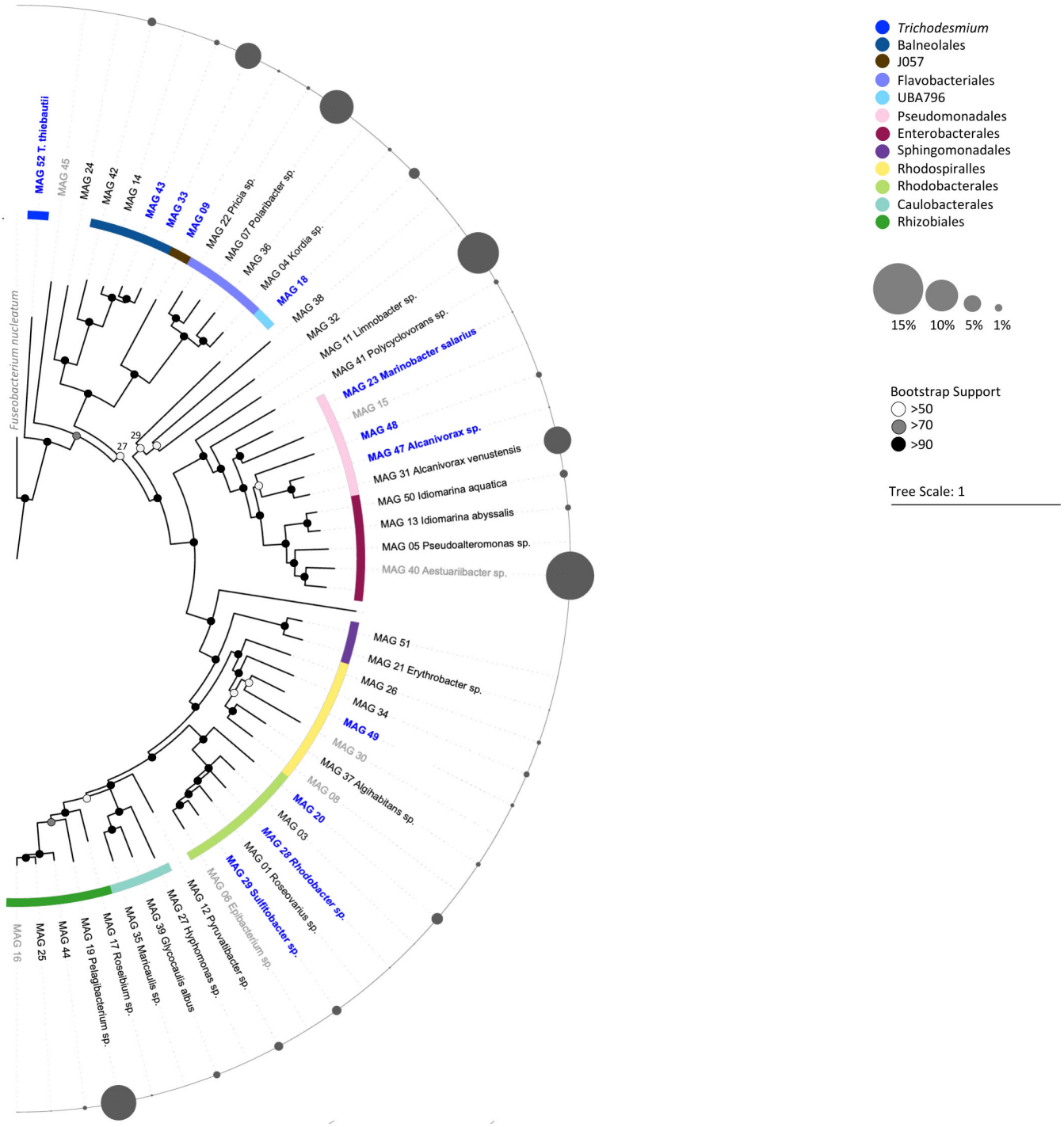

FIG 1 Phylogenetic tree of the 52 MAGs assembled from *Trichodesmium* colonies collected in the Red Sea. The tree is rooted by the outgroup *Fuseobacterium nucleatum*. Known bacterial orders are marked in different colors and several MAGs could taxonomically be identified at the genus level. The relative abundance of read counts mapping to each MAG within the *Trichodesmium* consortium, excluding *Trichodesmium* MAG 52 (±70% of total read counts) is presented as a bubble chart. The MAGs that matched to MAGs assembled from a previous metagenomic data set of Red Sea *Trichodesmium* colonies (38) are highlighted in blue and tentatively represent the core consortium (see Fig. S2). MAGs that did not meet the requirements of a high-quality genome (>90% completeness; <5% redundancy) are indicated in gray. Note that re-occurring MAGs are not necessarily the most abundant MAGs present within the *Trichodesmium* colony (see Tables S1 and S2 for a full taxonomic description of each MAG from each data set).

previous data sets accounted for less than 45% of the total sample reads, which may partially explain the lack of common MAGs across ocean basins. Nonetheless, the

presence of only a few re-occurring MAGs coupled with the re-occurrence of larger taxonomic groups across all ocean basins highlights the presence of a flexible assemblage of associated bacteria typically found in association with particles and algae. These observed characteristics may have important implications for the community structure and functioning of the *Trichodesmium* consortium.

Trichodesmium-associated bacteria synthesize diverse photolabile siderophores which may increase iron availability from dust

We explored the different Fe-uptake mechanisms present within the *Trichodesmium* consortium, as N_2 and C fixation in *Trichodesmium* is often constrained by Fe availability (33, 45). Despite the high Fe requirements of *Trichodesmium* (33), *T. thiebautii* (MAG 52) only contained genes for the uptake of free ferrous (Fe^{+2}) and ferric (Fe^{+3}) iron, but not organic Fe, suggesting that it is limited to a narrow range of forms it can take up directly (Fig. 2). By contrast, MAGs of the associated bacteria contained a large diversity of Fe-uptake genes which include the ability to take up organically bound Fe, such as heme, citrate, and siderophores (Fig. 2). The proteome of the *Trichodesmium* consortium further indicated that, at the time of sampling, *Trichodesmium* MAG 52 was actively taking up Fe through the presence of a Fe^{+3} uptake transporter (K02012) and increasing its intracellular Fe-pool through the presence of Fe-storage proteins (K03594; K04047) (Table S3).

The genomes of 10 *Trichodesmium*-associated bacteria spread across six different taxonomic orders encoded several different siderophore biosynthetic gene clusters (BGCs) (Fig. 2). The presence of a siderophore biosynthesis pathway was strain specific and not shared by an entire clade or lineage. Similar to previous studies, *T. thiebautii* MAG 52 did not contain any known siderophore biosynthesis pathways (18, 31). The siderophore BGCs could further be separated into four NRPS independent synthetase (NIS)-type and six non-ribosomal peptide synthetase (NRPS)-type siderophores (46, 47). Based on the presence of one or more *iucA/iucC* domains, the four NIS-type BGCs matched closely to petrobactin (*A. macleodii* MAG 10), vibrioferrin (*Marinobacteri* sp. MAG 23; *Limnobacter* sp. MAG 11), and rhizoferrin (Rhodospirillales MAG 08) biosynthesis pathways (Fig. 2). From the six different NRPS-type siderophores, only two matched to a known siderophore BGC, namely that of the membrane-bound amphibactin (100%) (*Alcanivorax* sp. MAG 47) and one resembling the BGC of the siderophores agrobactin (93%) or enterobactin (83%) (*Roseibium* sp. MAG 18) (Fig. 2). In addition, our results indicate the potential of several novel NRPS-type siderophore biosynthesis pathways (Fig. S3). NRPS-type siderophore BGCs are not as well characterized as NIS-type siderophores BGCs and can be difficult to distinguish from other secondary metabolites such as toxins and antibiotics (48). Analysis of the four NRPS-type siderophore BGCs using AntiSMASH 7.0 showed a putative siderophore receptor (SMCOG1082) of a TonB-dependent transporter (TBDT) indicating that these NRPS BGCs are involved in siderophore biosynthesis. While these four NRPS BGCs could not be linked to any known siderophore BGCs, several unknown metallophores have previously been identified and associated with *Trichodesmium* colonies and we speculate that some may be linked to these uncharacterized NRPS BGCs (25). Culture-based research will likely be required to link these pathways to a coinciding siderophore molecule.

Our results indicate that the *Trichodesmium* consortium in the Red Sea may act together with several photolabile siderophore-producing bacteria to mine Fe from dust. The re-occurring consortium members *Marinobacter* sp. (MAG 23), *Alcanivorax* sp. (MAG 47), and *A. macleodii* (MAG 10) are all able to synthesize siderophores (Fig. 2) whose citrate-moiety results in their photolability (49). *Trichodesmium* resides in the natural illumination of surface waters, where sunlight can reductively disassociate Fe from these siderophore complexes and provide a flux of available dissolved and inorganic Fe (50, 51). This dissolved inorganic Fe can be taken up by Fe^{+3} and/or Fe^{+2} transport mechanisms shown to be present in most consortium members (Fig. 2). *T. thiebautii* MAG 52 contains a Fe^{+3} and Fe^{+2} uptake mechanism but only contains a partial TonB system with

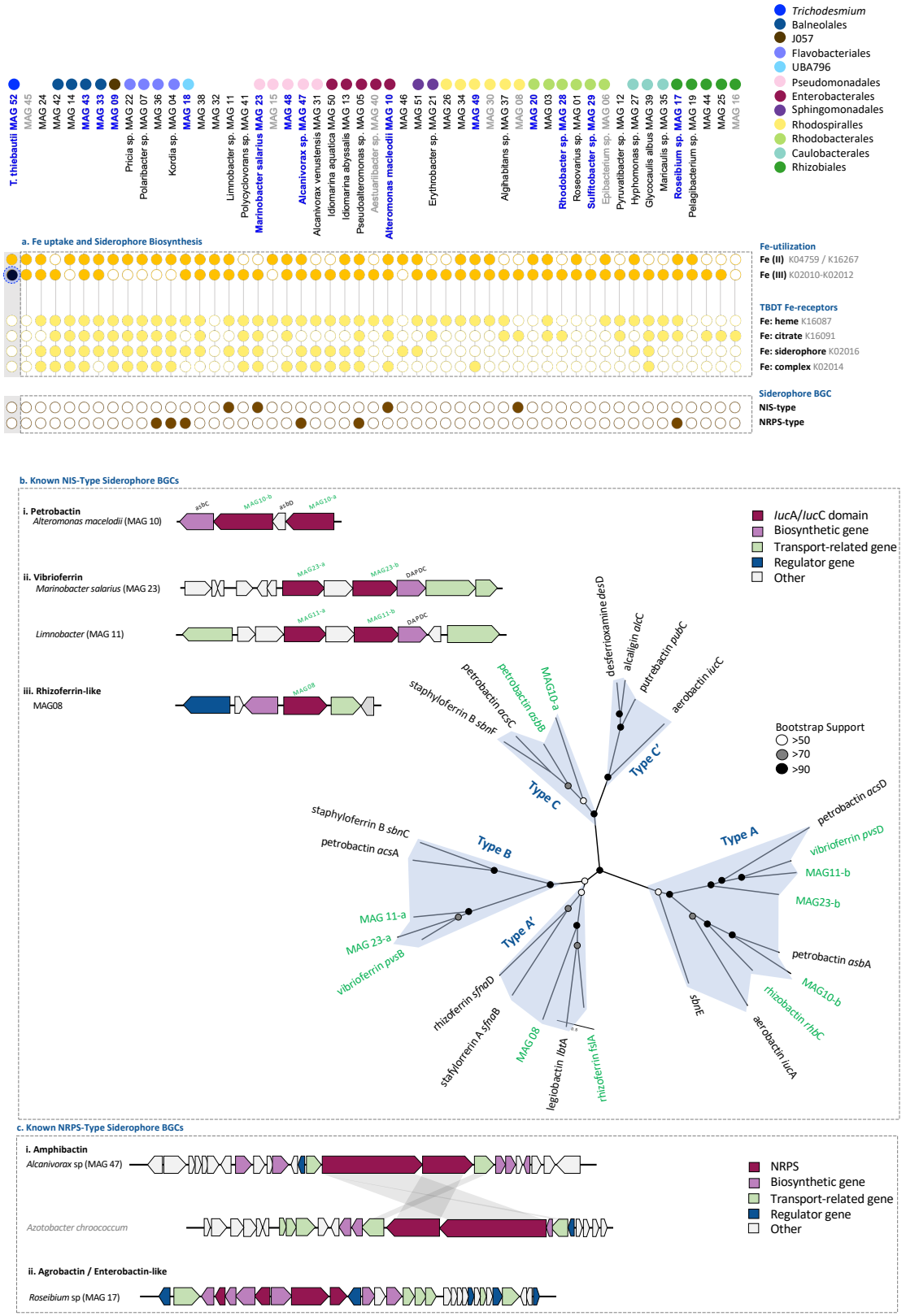

FIG 2 Summary of Fe-uptake and siderophore biosynthesis genes present in MAGs of the *Trichodesmium* consortium (see Table S4 and Materials and Methods for more details). Genes of interest, with a peptide in our proteomic data set are highlighted in black, indicating the protein was present at the time of sampling (see Table S3 for more details). (a) *T. thiebautii* MAG 52 does not contain a wide diversity of Fe-uptake genes, and the majority of associated bacteria MAGs

(Continued on next page)

FIG 2 (Continued)

contain an uptake receptor for organic Fe:complexes (heme, citrate, and siderophores). Several MAGs of associated bacteria contained non-ribosomal peptide synthetase (NRPS) or NRPS independent synthetase (NIS)-type siderophore biosynthetic gene cluster (BGC) pathways. (b) Siderophore BGCs of known NIS-type siderophore BGCs can be clustered according to the presence of one or multiple iucA/iucC domains which enable one to target the closest siderophore BGC match (46). Using this method, we were able to link these pathways to the synthesis of the photolabile (i) petrobactin, (ii) vibrioferrin, and (iii) rhizoferrin. (c) Siderophore BGCs of known NRPS siderophores. Using antiSMASH 7.0, which aligns BGCs to known siderophores within the MIBiG database, we were able to link two of the six putative siderophore NRPS-type pathways to (i) the membrane-bound amphibactin and (ii) agrobactin/enterobactin (see Fig. S3 for more details on the siderophore biosynthesis pathways).

no known TBDT or siderophore receptors (Table S4) (31, 52, 53) suggesting that these photolabile siderophores can be made bioavailable to *Trichodesmium* upon reductive disassociation by sunlight (29). If proven, these photolabile siderophores act as a "common good" within the *Trichodesmium* consortia, where only a few members carrying the trait are needed to benefit the entire consortium (54). This contrasts with the uptake mode of photostable siderophores, such as desferrioxamine-B (DFOB), which require a specialized TBDT (55). Similarly, amphibactin is a membrane-bound siderophore (56), which prevents its diffusion in the environment and subsequently, its ability to act as a public good to the *Trichodesmium* consortium. A recent gene-knockout study in an *A. macleodii* strain verified the ability of petrobactin to enhance the bioavailability of particulate Fe (28). Collectively, these findings support the idea that *Trichodesmium* can interact with several consortium members that produce photolabile siderophores to acquire Fe from dust.

The detected siderophore BGCs were identified in ubiquitous and copiotrophic bacteria typically found in association with particles and algae (17, 28), including *M. salarius* (MAG 23), *Alcanivorax venustensis* (MAG 47), and *A. macleodii* (MAG 10) (Fig. 2). Siderophore production within the *Trichodesmium* consortium therefore appears to be linked to a lifestyle associated with a rapidly changing environment rich in N and C, where regulating Fe bioavailability can provide a competitive advantage. An *A. macleodii* strain, previously isolated from *T. erythraeum* IMS101 cultures, was shown to benefit from the demise of *Trichodesmium* (43, 57), possibly suggesting a more complex interactive relationship between *Trichodesmium* and siderophore-producing bacteria. In line with particle and algae associations, siderophore production has also been associated with other ecological functions including heavy metal detoxification, oxidative stress response, and intra-specific communication (58) (Fig. S4; Table S4). The former two benefits from siderophore production could be relevant to investigate further considering *Trichodesmium*'s ability to center dust, particularly in the presence or absence of sunlight. Although no siderophore BGCs were present in our protein database, the presence of a few, yet reoccurring siderophore-producing MAGs augments previous measurements of siderophore production within the consortium (22, 25, 26) and reveals additional layers of *Trichodesmium* colonies' intriguing ability to sequester Fe from dust.

The *Trichodesmium* consortium is heterogeneous in genes related to the uptake and use of inorganic and organic phosphorous

We predicted that the majority of bacteria associated with *Trichodesmium* has the ability to metabolize phosphite or phosphonate. This can offer a key competitive advantage among consortium members in a P-limited environment (24, 59) including the Gulf of Aqaba in the Red Sea (60). Within a low-P environment, dissolved inorganic P is scarce, and bacteria must scavenge phosphate from dissolved organic P to meet their P-requirements (61). *Trichodesmium* is highly efficient at scavenging and utilizing organic-P through the production of alkaline phosphatase (AP), and the uptake of phosphonate and phosphite (62). Multiple studies have shown that *Trichodesmium* colonies in the Atlantic and Pacific Oceans are hotspots for the reduced P-cycling of phosphite or phosphonate (24, 59) suggesting that the majority of consortium members can process reduced-P. Confirming past studies (24, 63), *T. thiebautii* MAG 52 encodes a

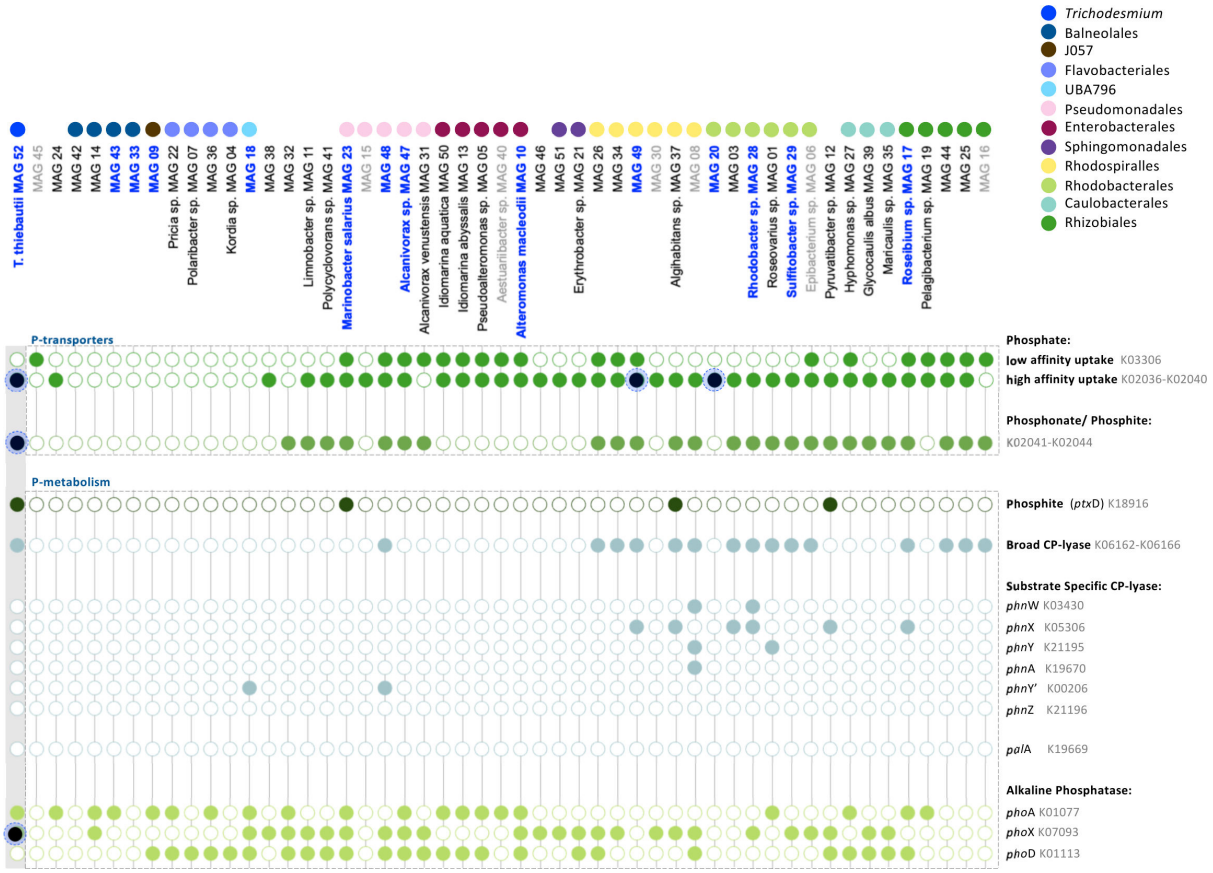

FIG 3 Summary of P-metabolism genes present in MAGs of the *Trichodesmium* consortium (see Table S4 and Materials and Methods for more details). Genes of interest, with a peptide in our proteomic data set are highlighted in black, indicating the protein was present at the time of sampling (see Table S3 for more details). *Trichodesmium* is known to be highly competitive for P, as exemplified by the multitude of (a) P-uptake systems for phosphate (*pst*S and *pst*B), phosphonate (*phn*CDE), and phosphonate/phosphite (*phn*CDE/*ptx*ABC). (b) Within the *Trichodesmium* consortium, the ability to metabolize a range of phosphonate compounds using a broad specificity C-P lyase (*phn*GHIJKLM) is more prominent than the substrate-specific hydrolysis of 2-aminoethylphosphonate (2-AEP; *phn*WXYZA) or phosphonopyruvate (*pal*A) and phosphite metabolism (*ptx*D).

set of functions to transport and metabolize diverse forms of P (Fig. 3). This includes the presence of a high-affinity phosphate transporter (*pst*SCAB) and a phosphonate/phosphite transporter (*phn*CDE/*pxt*ABC). In addition to these transport systems, genes related to the utilization and metabolism of organic-P included alkaline phosphatase (AP) (*pho*A, *pho*X), and a carbon-phosphorus (C-P) lyase (*phn*GHIJKLM) which hydrolyses a broad range of phosphonate compounds that can then be taken up. *T. thiebautii* MAG 52 also encodes phosphite dehydrogenase (*ptx*D) which catalyzes the oxidation of phosphite to phosphate. The presence of three *phn*D homologs (K02044), *pho*X (K07093), and *pst*B (K02036), within the proteome, suggests *Trichodesmium* was taking up phosphate, phosphonate, and hydrolyzing phosphoester compounds to meet its P-demand (Table S3). Several of these proteins are regulated by P (63), and their detection in the proteome is consistent with intense competition for P in the low P Red Sea.

As KEGG annotation cannot accurately distinguish between a phosphonate (*phn*CDE) and a phosphite (*pxt*ABC) transporter (64, 65), the presence of adjacent P-metabolism genes can further indicate the P-source linked to it (63). Specifically, the presence of phosphite dehydrogenase (*ptxD*) points to phosphite metabolism while a broad specificity C-P lyase (*phnGHIJKLM*) or the substrate-specific hydrolysis of phosphonopyruvate (*pal*A) or 2-aminoethylphosphonate (2-AEP) to acetaldehyde (*phn*WX), acetate (*phn*WYA), or glycine (*phn*Y'Z) can point to the metabolism of phosphonate (66, 67). Our results showed that the presence of a *phn*CDE/*ptx*ABC uptake system was specific

to certain taxonomic orders including Pseudomonadales, Rhodospirillales, Rhodobacterales, Caulobacterales, and Rhizobiales (Fig. 3). These taxonomic orders may therefore have a competitive advantage over members that lack them, particularly when faced with low P conditions. Of these 25 MAGs, 16 could be linked to phosphonate metabolism through the presence of a broad specificity C-P lyase. By contrast, only four MAGs spread across different taxonomic orders appeared to be able to process phosphite due to the presence of phosphite dehydrogenase (Fig. 3). Seven MAGs contained genes associated with the substrate-specific hydrolysis of 2-AEP but pathways were not complete, making it difficult to conclude (Fig. 3). While C-P lyases were not found within our proteomic data set, the presence of broad specificity C-P lyase (*phn*GHIJKLM) appears to be the most prevalent strategy for phosphonate utilization among consortium members, in comparison to 2-AEP hydrolysis (Table S3). This finding is in contrast to a meta-analysis of marine metagenomes where the most common phosphonate degradation strategy within the water column was the substrate-specific hydrolysis of 2-AEP (68). While 2-AEP is the most abundant phosphonate source in the marine environment (69), the prevalence of broad specificity C-P lyases may reflect the relevance of different phosphonate compounds to bacteria residing in *Trichodesmium* colonies.

The *Trichodesmium* consortium contains genes related to multiple N-metabolic pathways including denitrification

We explored the presence of multiple N-metabolic pathways predicted to influence the biogeochemical contribution of *Trichodesmium* regarding N_2 fixation. In *Trichodesmium* colonies sampled from the Red Sea, N_2 fixation appears to be uniquely attributed to *T. thiebautii* MAG 52 (Fig. 4). This contrasts with a previous study that observed *nif*H genes in both *Trichodesmium* and their associated bacteria (18). In our study, only *T. thiebautii* (MAG 52) contained the full N_2 fixation pathway (*nif*DKHEB: M00175) and therefore the ability to fix N_2 within the colony. Similar to other studies in *Trichodesmium* (18, 23, 70), the consortium contained several genes linked to the N-transformation pathways of assimilatory nitrate reduction to ammonium (ANRA), dissimilatory nitrate reduction to ammonium (DNRA), and denitrification (Fig. 4). In the absence of ammonium, the ANRA pathway enables marine bacteria to utilize nitrate as an alternative nitrogen source (71) while denitrification and DNRA are anaerobic respiratory pathways (72). The latter two pathways are N-loss and N-recycling processes, respectively, and are predicted to enable the residing community to conserve and fully utilize N_2 fixed by *Trichodesmium* (18). Collectively, the N-metabolic pathways present in the consortium can alter the net contribution of *Trichodesmium* colonies to local N-cycling dynamics (21, 72, 73). Nitrate and nitrite reductases are present in all three pathways and the same genes can be associated with multiple pathways in different bacteria (71, 74, 75). Therefore, in addition to KEGG annotations, we linked genes to N-metabolic pathways based on the presence of additional marker genes. Our classification of genes hereby contrasts with previous studies in *Trichodesmium* which relied solely on KEGG annotation (18, 23, 70). In this manner, 19 MAGs containing the nitrite reductase genes *nir*BD could instead be coupled to the nitrate reductase *nas*A, a marker gene for ANRA (71, 74, 75). Contrastingly, previous studies linked the presence of *nir*BD to DNRA within the *Trichodesmium* consortium (18, 23, 70). In our data set, only one MAG could be associated with DNRA using the marker gene *nrf*AH (76). Our functional assignment of genes to N-metabolic pathways hereby remains putative and to confirm whether these reductases are correctly associated with DNRA or ANRA pathways will require a thorough experimental investigation.

Denitrification appeared to be modular within the *Trichodesmium* consortium as different genomes collectively completed the entire pathway (Fig. 4 and 6). Three MAGs contained *nos*Z, the terminal step in denitrification, including *Roseibium aggregatum* (MAG 17), formerly known as *Labrenzia* sp., which was previously identified as an important denitrifying bacteria associated with *Trichodesmium* colonies in the Red Sea (44). While the presence of modularity in denitrification pathways itself is not uncommon

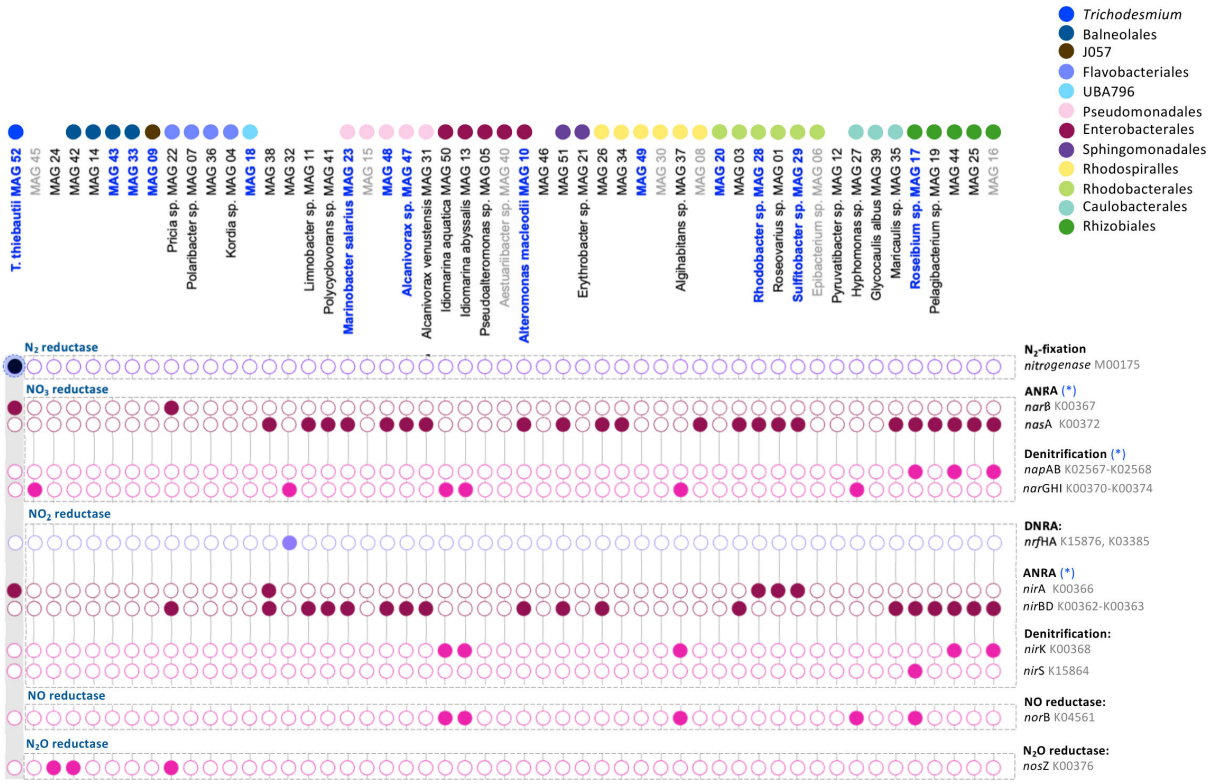

FIG 4 Summary of N-metabolism genes present in MAGs of the *Trichodesmium* consortium (see Table S4 and Materials and Methods for more details). Genes of interest, with a peptide in our proteomic data set are highlighted in black, indicating the protein was present at the time of sampling (see Table S3 for more details). *T. thiebautii* MAG 52 is the only MAG with an N_2 fixation pathway. When a gene is not considered a marker gene for a specific pathway (and the assignment is hereby putative), the pathway is marked with a blue asterisk. The majority of MAGs contained a putative ANRA pathway to take up nitrate from the environment, while others contained several steps within the pathways of denitrification. Denitrification appears to be modular, where several MAGs completed the denitrification pathway but none contained genes for its entirety.

(77–79), finding its presence within the *Trichodesmium* consortium indicates the spatial role *Trichodesmium* plays in providing a platform where each step can be performed by a distinct subset of the microbial consortium. The presence of either denitrification or DNRA pathways is rare in the oxic surface ocean, as these processes require anoxic and suboxic conditions (80). In the surface waters, suboxic and anoxic micro-environments were detected in marine aggregates, such as marine snow and fecal pellets, where enhanced microbial respiration results in the consumption of O_2 and the formation of anoxic microzones (81, 82). *Trichodesmium* colonies may represent an additional environment where such processes can take place within the surface waters of the ocean (44) (83). Our proteomic data set indicated *Trichodesmium* MAG 52 was actively fixing nitrogen (K02588, K02591, and K02586), taking up ammonium (K03320), and metabolizing nitrogen *via* a glutamine synthetase (K01915). Peptides also matched to a glutamine synthetase in consortium members (K01915 and K00982). Peptides, however, did not match any sequences related to ANRA, DNRA, or denitrification pathways (Table S3). While evidence suggests DNRA and denitrification processes are taking place within natural *Trichodesmium* colonies (44, 83), measured O_2 concentrations in *Trichodesmium* colonies are often well above zero even during darkness (21) and it is still not entirely clear when and where these N-cycling pathways are actively taking place (21, 73). It is suggested that the rapid O_2 consumption by residing bacteria and the presence of dust (84) could support the formation of suboxic microenvironments in *Trichodesmium* colonies but will require more studies to resolve (73, 84). Altogether, more work is required to fully elucidate the functional potential of consortium members and the exact

tradeoffs between interactive members of the *Trichodesmium* consortium regarding N-metabolism.

Trichodesmium consortium members are auxotrophic for vitamins B_1, B_7, and B_{12}

The presence of vitamin auxotrophy within the consortium indicates a dependency of members on *Trichodesmium* for vitamin provision. Our results show that while *Trichodesmium* MAG 52 encodes the biosynthesis pathways for cobalamin (B_{12}), biotin (B_7), and thiamin (B_1), the majority of associated bacteria (49 of the 52 MAGs) lack a biosynthesis pathway for at least one of these vitamins (Fig. 5). Bacteroidetes, Sphingomonadales, and Enterobacterales species, including *A.s macleodii* (MAG 10), were all auxotrophic for vitamin B_{12}. MAGs from the order Rhodospirillales lacked the ability to synthesize vitamin B_7, whereas Rhizobiales and Rhodobacterales appeared to be auxotrophic for both vitamin B_1 and B_7. Only *Pseudoalteromonas* encoded all three biosynthesis pathways, including the re-occurring taxa *Alcanivorax* sp. (MAG 47) and *M. salarius* (MAG 23). These consortium members may also contribute to the synthesis and provision of vitamins B_1, B_7, and B_{12} to their auxotrophic counterparts (Fig. 5) but, based on overall relative abundance, *Trichodesmium* represents the most likely source of B vitamins. Previous studies have shown that *Trichodesmium* can synthesize and secrete the vitamin cobalamin (B_{12}) to the environment (23). Vitamin uptake genes were not found in our proteomic data set, and only one peptide matched the biosynthesis pathway of vitamin B_{12} (K19221) in *Trichodesmium* (Table S3). Nonetheless, the production and secretion of different vitamins by *Trichodesmium* may be a key determinant of community composition, stimulating vitamin-based interactions among auxotrophic consortium members.

Nutrient cycling in *Trichodesmium* consortia is likely underpinned by interactions with copiotrophic bacteria

A literature study of the bacterial lineages present within the *Trichodesmium* consortium describes these taxonomic orders as typical primary colonizers of marine surfaces such as algae or sinking particles (39–41). Strains of *Alteromonas*, *Marinobacter,* and *Pseudomonas* often co-occur with phototrophs such as cyanobacteria as well as oil spills, likely as a result of their ability to degrade hydrocarbons and aromatic compounds

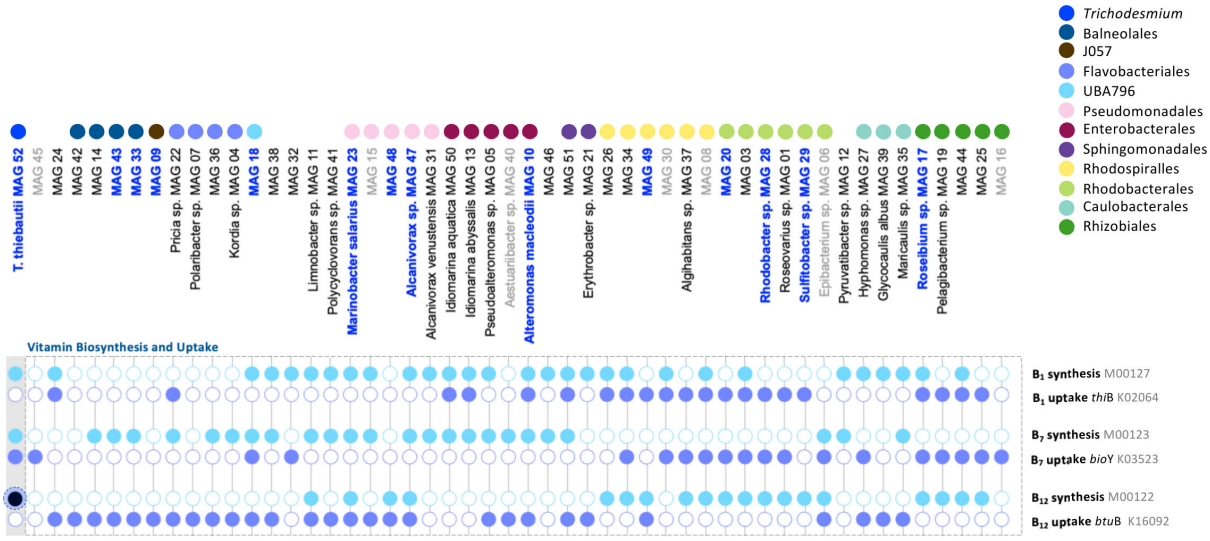

FIG 5 Summary of vitamin B_1, B_7, and B_{12} biosynthesis and uptake pathways (see Table S4 and Materials and Methods for more details). Genes of interest, with a peptide in our proteomic data set are highlighted in black, indicating the protein was present at the time of sampling (see Table S3 for more details). *Trichodesmium* contains the biosynthesis pathways for all three vitamins, but the vast majority of associated bacteria were auxotrophic for one or more vitamins. These results suggest that the associated bacteria rely on *Trichodesmium* to obtain their vitamin requirements.

(42, 85). While the prevalence of these copiotrophic lifestyles within the *Trichodesmium* consortium has been observed in previous studies (17, 19), it is still unclear whether their association with *Trichodesmium* is specific or reflective of a more general interaction between algae or particle associations. Copiotrophy is a widely encompassing term, but it generally concerns opportunistic bacteria with a versatile metabolic capacity to use a wide range of substrates, acting as important remineralizers by rapidly responding toward areas enriched in organic matter (86, 87). It is, therefore, possible to consider the majority of the consortium as active remineralizers of the C, N, P, Fe, and vitamins *Trichodesmium* provides. Viewing *Trichodesmium* as a hotspot for nutrients is supported by our proteomic samples where the most abundant peptides attributed to *T. thiebautii* MAG 52 were related to photosynthetic (K05376, K05377, and K02284) and diazotrophic machinery (K03839, K02588, and K02591) indicating that the processes of N and C fixation were actively taking place at the time of collection (Table S3). Meanwhile, *T. thiebautii* MAG 52 appeared to be limited in Fe and P and actively taking up their various forms through different transporter systems. Due to low peptide biomass, it was more difficult to make conclusions about the protein hits associated with consortium members but included the uptake of nutrients such as amino acids (K09969 and K01999), phosphate (K02040), Fe (K02014), sugars (K02055, K02027, and K10552), and peptides (K02035 and K02032) (Table S3).

In general, copiotrophic bacteria are characterized by the presence of several bacterial-interactive traits (88), indicating the potential of the *Trichodesmium* consortium to form complex interactive networks. The prevalence of copiotrophic bacteria can be supported by the presence of several bacterial-interactive and particle-associated traits (88, 89) (Fig. S4; Table S4), and a previous study has shown gene expression among *Trichodesmium* and associated bacteria appears to be coordinated (23). Although speculative, this could be supported by the presence of several proteins involved in interactive traits such as pilus formation (K02658, K02662, and K02669) and secretion systems (K11003; K03072; K03070, K03110, K03073; K03217) within *T. thiebautii* MAG 52 (Table S3). In this study, 15 associated bacteria MAGs, including re-occurring members from the order Rhodobacterales, contained an acyl homoserine lactone (AHL) biosynthesis pathway (Fig. S4a). Previous studies have shown that the presence of AHLs can modulate N_2 fixation rates (20) and increase alkaline phosphatase (AP) activity within *Trichodesmium* colonies (90). In particular, consortium members belonging to the phylum Proteobacteria (e.g., Rhodobacterales) are predicted to coordinate their behavior through AHLs (15, 90–92). While our proteomic data set did not indicate their expression at the time of sampling, our results do highlight several potential candidates to further investigate. An AHL biosynthesis pathway was also found in *T. thiebautii* MAG 52. The presence of AHL-homologs had previously been reported for *T. erythraeum* IMS101; however, it lacked an AHL-binding residue suggesting that it is likely not functional (90). Whether *T. thiebautii* MAG 52 contains a functional AHL and can coordinate its gene expression through AHLs is beyond the scope of this analysis. Lastly, the functional association of poorly studied bacterial clades, in particular Balneolales (Table S1), remained more elusive and indicated that we still lack a clear understanding of these bacteria and their interactive role within the *Trichodesmium* consortium. Further investigations will be required to untangle the dynamics of both general and specific interactions taking place between members of the *Trichodesmium* consortium.

Conclusions and future perspectives

Our results characterize *Trichodesmium* colonies as a hotspot for C, N, P, Fe, and vitamins in the residing community (Fig. 6). Colonies harbored a flexible assemblage of bacteria, and several core-community members, mostly characterized by a copiotrophic lifestyle adept at processing the nutrients *Trichodesmium* can provide. Vitamins are likely a major component influencing community structure as the majority of bacteria were auxotrophic for one or more vitamins and likely relied on *Trichodesmium* for their synthesis. A key beneficial trait *Trichodesmium* can acquire from its residing community is related to

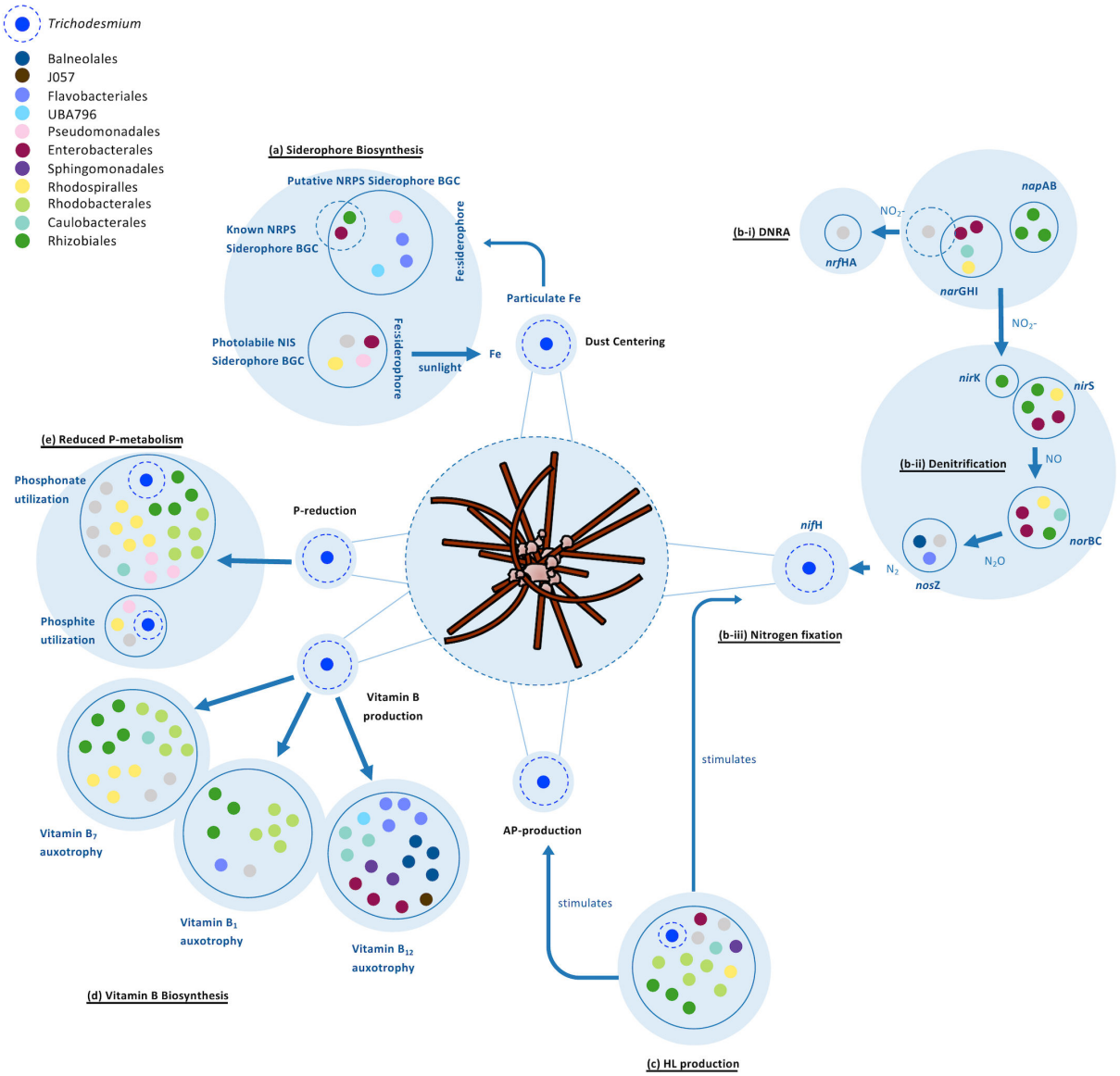

FIG 6 Schematic summary of the potential interactions taking place within the *Trichodesmium* consortia as discussed in this study. Genes for each trait are noted in Table S4. Each blue bubble indicates a different function, and each blue circle pools together the different MAGs containing a particular trait for that function. Each MAG is represented by a colored dot reflective of the taxonomic order it is associated with (gray dots represent MAGs that were not attributed to a specific taxonomic order). *Trichodesmium* is highlighted in cobalt blue, and the important functions *Trichodesmium* displays are highlighted with a dotted circle (see Table S4 for the full list of functional traits, presence/absence, and number of homologs). (a) Siderophore biosynthesis. Several siderophore biosynthesis genes are present in a diverse set of MAGs. Siderophores are predicted to mine Fe from dust that has been collected and centered within *Trichodesmium* colonies. Six of the four were identified as known NIS-type siderophores of which petrobactin, vibrioferrin, and rhizoferrin are photolabile. In the presence of sunlight, the Fe complexed to these siderophores disassociates and acts as a public good to the larger consortium. The genes and pathways for (b-i) dissimilatory nitrate reduction to ammonium (DNRA) and (b-ii) denitrification are present in several MAGs of the consortium. Denitrification appears to be modular within the *Trichodesmium* consortium as a different subset of bacteria contains genes for the terminal step (*nosZ*). (b-iii) *Trichodesmium* MAG 52 was the only MAG containing N_2 fixation genes in this data set. (c) HL production. Several MAGs contain an acyl-homoserine lactone synthesis pathway (AHL), including *Trichodesmium*. AHLs act as an intercommunicating quorum sensing molecule and have been shown to influence the production of alkaline phosphatase (AP) and N_2 fixation in *Trichodesmium* (associated bacteria can also produce AP but is not depicted here—see Fig. 3). (d) *Trichodesmium* can produce the vitamins cobalamin (B_{12}), thiamin (B_1), and biotin (B_7). Different subsets of MAGs were auxotrophic for different vitamins, defined for each MAG lacking a biosynthesis pathway but containing an uptake pathway for one or more of these vitamins. Very few MAGs are capable of synthesizing all three vitamins and consortium members likely exhibit an interdependent relationship for these vitamins (associated bacteria can also produce vitamins but is not depicted here—see Fig. 5). (e) Reduced P-metabolism. *Trichodesmium* colonies have been shown to take up and utilize phosphonate and phosphite in a low-P environment and MAGs containing these pathways may have a competitive advantage over members that lack them.

the acquisition of iron from particles through the production of siderophores. *Trichodesmium* harbored several bacteria that can secrete photolabile siderophores, supporting the idea that *Trichodesmium* can meet their high Fe requirements through a collaborative effort to obtain Fe from dust particles collected within the colony. Whether the production of siderophores can further benefit the colony through processes such as heavy metal detoxification and the reduction of oxidative stress may be intriguing to further investigate. To determine the conditions in which such processes may benefit the colony could be explored by moving away from bulk measurements and toward the application of single-colony approaches.

The complex *Trichodesmium* consortium represents an intriguing system to study the intricate interactions taking place within a particle-rich system (Fig. 6), asking how such processes can impact the C, N, P, and Fe cycles in a rapidly changing ocean where shifts in the microbial community composition and their abundance are expected to occur (35, 93). Understanding the factors that drive its community structure across temporal and spatial scales remains an important avenue to continue to explore through a combination of transcriptomic and proteomic techniques. The observed redundancy of multiple traits involved in nutrient cycling among *Trichodesmium* consortium members, including the production of siderophores, indicates the flexibility of a colony while maintaining such traits within the community. Altogether, our results highlight the presence of a dynamic consortium where functional traits are well conserved within the *Trichodesmium* consortium, underpinning the resilience of the colony within an ever-changing ocean.

MATERIALS AND METHODS

Trichodesmium metagenomic sampling and extraction

Trichodesmium puff colonies, 1–2 mm in size, were hand-picked during the spring bloom (6th May 2019) and separated into three samples for metagenomics (~100–200 colonies each) using a 100 µm phytoplankton net at 20 m depth in the Gulf of Aqaba (Eilat, Israel) (29.56°N, 34.95°E). Colonies were washed three times, by gently picking colonies with a pipette and placing them in new petri dishes with fresh 0.2 µm filtered and sterilized seawater, before being filtered on a 0.2-µm, 45-mm polycarbonate (PC) filter and vacuum filtered using a peristaltic vacuum pump. Free-living bacteria are hereby removed while bacteria associated with the *Trichodesmium* colony, both loosely and strongly, are captured and sequenced. This method is also similar to that of other studies (18, 19, 23). Filters were flash-frozen in liquid N_2 and kept at −80°C. DNA was extracted from the three *Trichodesmium* samples using the DNeasy Plant Pro Kit (Qiagen), with a minor modification to the lysis procedure. The kit-provided tissue disruption tubes were not used. Rather, ~250 µL zirconia/silica beads (0.5 mm) were added to each sample tube before the addition of Solution CD1, and samples were vortexed for 5 min. The resulting lysate was processed as per the remainder of the manufacturer's instructions. DNA was quantified fluorometrically on a Qubit 4.0 Fluorometer (Thermo Fisher Scientific, USA). From the three samples, metagenomic libraries were prepared and sequenced with 0.7 billion reads 2 × 151 bp reads on Illumina NovaSeq S4 at the DOE Joint Genome Institute (California, USA). This study can be found under JGI Gold Study ID number Gs0149370 (https://gold.jgi.doe.gov/study?id=Gs0149370).

Metagenomic raw read assembly and binning

Sequences were analyzed from this study and compared to previously published metagenomic studies of *Trichodesmium* colonies isolated from the Red Sea (PRJNA804487) (38) from the Pacific (PRJNA435427) (23) (PRJNA358796) (18) and the Atlantic (PRJNA330990) (19) Oceans. All raw sequences, from all data sets, were analyzed using the same protocol described below. The analysis can also be found in the following github depository: https://github.com/cocokoedooder/Trichodesmium_consortium

Metagenomes were assembled, binned, and quantified using the ATLAS (v2) pipeline (94). Briefly, raw sequences underwent quality control through the BBTools suite (95, 96)

and were assembled using metaSPAdes (97) (k-mer lengths: 21, 33, 55, 99, and 121 bp). For each sample, MAGs were binned automatically using MetaBAT 2 (98), MaxBin 2.0 (99), and VAMB (100). Using DAS Tool (101) and dRep (102) for quality filtering, a final set of 52 non-redundant MAGs was obtained based on a 97.5% average nucleotide identity (ANI) and 75% completeness cutoff. The 97.5% ANI cutoff was chosen to account for the possibility of multiple *Trichodesmium* sub-species that may be present in our samples, as had been shown in a previous study (38). The completeness and redundancy of each bin were subsequently assessed using CheckM2 v0.1.2 (103). MAGs were considered of high quality if they reached 90% completeness and 10% redundancy and were highlighted throughout our results and figures (37). Genes were predicted using Prodigal (104). In this study, the final set of 52 MAGs did not undergo manual inspection for further MAG refinement. The coverage of each MAG was quantified across each sample by mapping reads back to each MAG. We refer to the MAGs from this study as MAG XX, bacterial MAGs from a previous Red Sea study as R-XX (38), and bacterial MAGs from other metagenomic studies from the Pacific and Atlantic Oceans (18, 19, 23) as T-XX.

Phylogenetic diversity

MAGs were taxonomically characterized using the genome taxonomy database tool kit GTDB-tk v2.1.1 (105). The taxonomic names of MAGs were classified according to the GTDB taxonomy database of bacterial genomes (version r207.0) where phylogeny is inferred using a concatenated alignment of 120 ubiquitous single-copy proteins ("bac120") (Tables S1 and S2). The GTDB taxonomic match for each MAG was subsequently linked to its corresponding NCBI taxonomy (https://gtdb.ecogenomic.org/). GTDB taxonomy was preferably used throughout the analysis, but to simplify comparisons within the literature, we used the larger NCBI taxonomic order Rhodospirillales which encompasses the smaller GTDB taxonomic groups of Kiloniellales and Thalassobaculaceae.

Phylogeny was inferred using GTo-Tree (v.1.16.12; default settings) (106) from a concatenation of 74 conserved single-copy HMM markers for bacteria using the best-fit model Q.pfam +R7 model in IQ-Tree (v2.1.3), using the Bayesian information criterion (BIC) (107, 108). Shimodaira-Hasegawa approximate likelihood-ratio test (SH-aLRT) and ultrafast bootstrap approximation (UFBoot) branch support values were estimated from 1,000 bootstraps. The tree was visualized using FigTree (v1.4.4) and rooted using the genome of *Fuseobacterium nucleatum* (PRJNA1419) as an outgroup according to reference (109). MAGs were considered to re-occur in another metagenomic data set if their genome sequence identity was >97.5% similar (Fig. S2).

Functional diversity

The amino-acid sequence for each MAG was annotated using HMMscan and GhostKoala under default parameters (50% identity cutoff), utilizing the PFAM and KEGG databases, respectively (110, 111). Secondary metabolites (e.g., biosynthesis of siderophores and AHLs) were screened for each MAG using AntiSMASH 7.0. In certain cases, the Rapid Annotation using Subsystem Technology (RAST) server was used to annotate amino-acid sequences using the collection of protein families, FIGfams (112, 113). In addition, each MAG was assessed using METABOLIC (v4.0) which lists the presence of interactive KEGG modules for each MAG (114). Based on these different annotation formats, we inspected each MAG for the presence of functional traits. For each functional pathway or gene, the annotation method, the coinciding KEGG and PFAM ID number, and the number of hits in each MAG can be found in Table S4. Pathways consisting of multiple genes were considered present using a general cutoff of 50% completeness followed by a manual inspection of each pathway for each MAG. For transporter genes, marker genes of substrate-specific receptors were a determining factor in marking their presence or absence in a genome. The presence and absence of different traits for each MAG were displayed using the Interactive Tree of Life (iTOL) visualization tool (115).

Iron metabolism

To explore the relationship between *Trichodesmium* and its associated bacteria for the bioavailability of Fe from dust, we searched MAGs for the KEGG IDs of multiple Fe-uptake pathways. Fe-uptake receptors were separated according to abiotic Fe^{+2} (*feo*B: K04759), Fe^{+3} (*afu*A: K02012), and organic Fe uptake which is further separated into Fe-heme (*hem*R: K16087), Fe-citrate (*fec*A: K16091), and Fe-siderophores (*fev*S: K02016). The presence of a Fe:siderophore receptor K02016 should be taken cautiously as its KEGG annotation does not clearly distinguish between siderophores or vitamin B_{12} as a substrate.

Siderophore biosynthesis pathways within MAGs were identified using a conservative approach that required a combination of both KEGG and HMM annotation for each MAG and the use of AntiSMASH 7.0 and FeGenie tools (116) (Fig. S3). Pathways were separately analyzed for both NRPS-type and NIS-type biosynthesis pathways. NIS-type siderophore biosynthesis pathways are more easily identifiable than the two, and we searched for the presence of the Fe-uptake chelate *iuc*A/*iuc*C domain (PF04183). NIS genes were clustered into different types using known NIS biosynthesis genes as defined and described by Carroll and Moore (46). Briefly, genes were aligned using the online Multiple Alignment Fast Fourier Transformation (MAFFT v7.490; L-INS-i) software (117). A phylogenetic tree was constructed from this alignment using IQtree2 (v2.1.3). The resulting best-fit substitution model Q.pfam + F + R4 was selected using Bayesian information criteria (BIC). Branches were assigned using SH-aLRT and UFBoot branch support values were estimated from 1,000 bootstraps. The resulting consensus tree was visualized using FigTree (v1.4.4).

The identification of NRPS-type siderophore biosynthesis pathways required FeGenie, which utilizes biosynthetic pHMMs, and was used to highlight putative NRPS-type siderophores that were further inspected using AntiSMASH 7.0. In AntiSMASH 7.0, MAGs containing an NRPS pathway coupled with a siderophore receptor (SMCOG1082) of a TBDT within the gene cluster were marked as putative NRPS siderophore producers. These pathways were further compared to known NRPS siderophore pathways using the MiBIG database (118) and the KEGG pathway map for NRPS-type siderophores BGC map01053.

Vitamin B_1, B_7, and B_{12} biosynthesis and uptake

We assume that a bacterial strain is auxotrophic for thiamin (B_1), biotin (B_7), or cobalamin (B_{12}) if their coinciding genome lacks a vitamin B_{12} biosynthesis pathway but contains a transporter for its uptake. For each vitamin, we searched MAGs for the KEGG IDs and determined the presence of a vitamin biosynthesis pathway using KEGG modules. The presence of a B_1 biosynthesis pathway (M00127) was determined by the presence of *thi*CDEGL (K00946; K00788; K03149; K00941; K03147). A vitamin B_1 uptake system was considered present if the marker gene *thi*B (K02064), encoding a vitamin B_1 binding subunit of an ABC transporter, was present (119) alongside a coinciding transporter system *thi*PQ (K02063, K02062). The presence of a vitamin B_7 biosynthesis pathway (M00123) was determined by the presence of *bio*ABDF (K01012; K01935; K00833; K00652). A vitamin B_7 uptake system was present if the marker gene *bio*Y (K03523), encoding a vitamin B_7 binding subunit of an ABC transporter (120), was present alongside a coinciding transporter system *bio*NM (K16783 and K16784) or *efc*TA_1A_2 (K16785, K16786, and K16787). Finally, the presence of a vitamin B_{12} biosynthesis pathway was determined by the KEGG ID module M00122. A putative vitamin B_{12} uptake system was considered present when the marker gene *btu*B (K16092) was observed (121, 122), which encodes a vitamin B_{12} receptor for a TBDT. In addition, we queried the presence of two *btu*F genes (K06858 and K25034) that encode a vitamin B_{12} binding subunit of the ABC transporter *btu*CDF (K06074 and K06073).

Nitrogen metabolism

N-metabolism pathways of MAGs were compiled from a selected set of KEGG ID numbers. MAGs were queried for the presence of genes related to N_2 fixation, dissimilatory nitrate reduction to ammonium (DNRA), assimilatory nitrate reduction to ammonium (ANRA), and denitrification. The ability to fix N_2 was determined through the presence of *nif*HKD (M00175: K02588, K02591, K02586, and K00531). Denitrification, ANRA, and DNRA can be separated into several steps. All three pathways involve a nitrate reductase which reduces nitrate to nitrite (*nap*AB: K02567, K02568*; nar*GHI: K00370–K00374; *nar*B: K00367). In ANRA and DNRA, the nitrite is further reduced to ammonia by a nitrite reductase (*nrf*AH: K15876 and K03385*, nir*A: K00366; *nir*BD: K00362 and K00363). In denitrification, the nitrite is further reduced to nitric oxide (*nir*S: K15864*; nir*K: K00368), nitric oxide to nitrous oxide (*nor*BC: K04561 and K02305), and nitrous oxide to N_2 (*nos*Z: K00376). Classification of genes in KEGG (map00910) is largely based on *Escherichia coli*, which appears to form the exception rather than the rule when linking nitrate reductases to different N-metabolic pathways (71, 123–125). We therefore performed an extensive literature search of marker genes to putatively screen and assign genes in each genome to their respective pathways. In this manner, the genes *nor*B and *nir*A present in *Trichodesmium* MAG 52 could be associated with nitrate assimilation (126). Similarly, the genes *nas*A and *nir*BD in 21 MAGs were associated with the ANRA pathway (71) rather than DNRA as had been done in previous studies (18, 23, 70). The nitrite reductases *nir*K or *nir*S acted as marker genes for the presence of denitrification pathways (124) and enabled us to putatively link the nitrate reductases *nar*GHI and *nap*AB to this pathway instead of DNRA in several genomes (10 MAGs) (125). Similarly, n*rf*A acted as a diagnostic for the presence of DNRA within a genome (76), with only one MAG associated with this pathway.

Phosphorus metabolism

The presence of multiple P-metabolic pathways within the *Trichodesmium* consortium was explored. We explored several transporters for high-affinity phosphate (*pst*SCAB: K02036–K02038, K02040), low-affinity phosphate (*pit*A: K03306), and a phosphite/phosphonate transporter (*phn*CDE: K02041; K02042; and K02044). We searched MAGs for KEGG ID numbers for the presence of phosphite dehydrogenase (*ptx*D: K18916), which oxidizes phosphite to phosphate, and the broad specificity C-P lyase (*phn*GHIJKLM: K06162-6 and K05780-1), which hydrolyses phosphonate bonds in organic P. In addition, MAGs were queried for the presence of the substrate-specific phosphonate metabolism enzymes including *phn*A (K19670), *phn*X (K05306), and *phn*W (K03430), and collectively these genes were used to indicate the functional potential for reduced P-metabolism. Lastly, MAGs were examined for the presence of alkaline phosphatases (AP) *pho*A (K01077), *pho*X (K07093), and *pho*D (K01113). AP hydrolyses phosphoesters enabling the metabolism of organic-P.

Bacterial and particle interaction traits

Known traits that are indicative of putative bacterial interactions were previously summarized and identified by Zoccarato et al. for marine bacterial genomes (88). In addition to siderophore biosynthesis and vitamin B_1, B_7, and B_{12} biosynthesis (see previously), MAGs were probed for the presence of AHL biosynthesis genes using AntiSMASH 7.0. All AHLs, including the AHL present in *Trichodesmium,* were homologs of *lux*R and could further be identified by the presence of the pfam domain PF00765. The presence of auxin efflux genes was determined according to RAST annotation. MAGs were screened for KEGG ID numbers for the presence of the quorum sensing regulation gene *lux*R (K07782) and *bja*R1 (K18098), secretion systems (KEGG map03070), and motility and adhesion genes (KEGG map02020) (Table S4). Traits involved in particle interaction were previously described and summarized by Debeljak et al. (89). In addition to Fe-uptake and siderophore biosynthesis pathways, MAGs were screened for KEGG ID

numbers involved in metal-response, metal-efflux, and metal-storage pathways (Table S4).

Trichodesmium proteomic sampling and extraction

In parallel to the metagenomic sampling, 20 puff colonies were picked, in triplicate, for proteomic analysis at the Environmental Molecular Sciences Laboratory (EMSL) at Pacific Northwest National Laboratory according to standardized protocols. Colonies were washed three times with filtered seawater, as described above. Samples were centrifuged, and the resulting pellet was diluted in 200 µL 8 M urea and transferred to 2 mL pre-filled micro-organism lysing mix glass bead tubes and bead beat in a Bead Ruptor Elite Bead Mill Homogenizer (OMNI International, Kennesaw, GA) at a speed of 5.5 for 45 s. After bead beating, the lysate was transferred to a new 4 mL tube and immediately placed in an ice block and spun at 1,000× *g* for 10 min at 4°C. An amount of 200 µL of the lysed samples was transferred into 2 mL centrifuge tubes. A bicinchoninic acid (BCA) assay (Thermo Scientific, MA, USA) was performed to determine protein concentration. Following the assay, 10 mM dithiothreitol was added to the samples and incubated at 60°C for 30 min with constant shaking at 800 rpm. Samples were diluted eightfold for preparation for digestion with 100 mM NH_4HCO_3, 1 mM $CaCl_2$, and sequencing grade trypsin (Promega, WI) was added to all protein samples at a 1:50 (wt/wt) trypsin-to-protein ratio for 3 h at 37 ˚C, 450 rpm. Digested samples were desalted using a 4-probe-positive pressure Gilson GX-274 ASPEC system (Gilson Inc., WI) with Discovery C18 50 mg/1 mL solid phase extraction tubes (Supelco, MO), using the following protocol: 3 mL of methanol was added for conditioning followed by 3 mL of 0.1% trifluoroacetic acid (TFA) in H_2O. The samples were then loaded onto each column followed by 4 mL of 95:5: H_2O:ACN, 0.1% TFA. Samples were eluted with 1 mL 80:20 ACN:H_2O, 0.1% TFA. The samples were concentrated using a speed vac and a final BCA was performed to determine the peptide concentration and samples were fractionated into 12 fractions for LC-MS/MS analysis.

MS analysis was performed using a Q-Exactive Plus mass spectrometer (Thermo Scientific) outfitted with a homemade nano‐electrospray ionization interface. Electrospray emitters used 150 µm o.d. × 20 µm i.d. chemically etched fused silica (124). The ion transfer tube temperature and spray voltage were 300°C and 2.2 kV, respectively. The data were collected for 120 min following a 10-min delay after completion of sample trapping and start of gradient. FT-MS spectra were acquired from 300 to 1,800 m/z at a resolution of 70 k (AGC target 3e6) and the top 12 FT-HCD-MS/MS spectra were acquired in data‐dependent mode with an isolation window of 1.5 m/z at a resolution of 17.5 k (AGC target 1e5) using a normalized collision energy of 30, dynamic exclusion time of 30 s, and detected charge state of an ion 2 or higher. The resulting proteomic samples from this study were deposited in MassIVE (PXD040628).

Proteomic analysis

Peptide matching was performed using MSGF+ software (127) against the amino-acid sequences obtained from the matching metagenome data set through PROKKA (128) using the ATLAS pipeline (v2) (94) resulting in 5,292 unique clean peptide sequences that could subsequently be mapped to 1,253 different protein sequences. MSGF+ was used in target/decoy mode with 20 ppm parent ion tolerance, partial tryptic rule, and M-oxidation (+15.9949) as a dynamic modification. Best matches from the MSGF+ searches were filtered at a 1% false discovery rate based on the target/decoy model and this final set of peptides was used in the consequent quantitative analysis. We obtained a total of 14,074 spectral counts without consideration of peptide specificity for proteins and MAGs, to encompass the widest range of putative protein matches. 91.6% of peptides mapped to a unique protein sequence and 92.7% of peptides were found in a single MAG, indicating that for most peptides there was little ambiguity regarding its protein and taxonomic match. Peptide sequences could be linked to 81 proteins that were associated with 65 genes of interest. Their sequences were manually checked for MAG and protein

specificity to assess the reliability of these counts. Spectral count normalization was performed using the normalized spectral abundance factor (129). Briefly, counts were divided by the matching peptide length and their relative abundance was calculated from the total number of spectral hits per sample (Table S3).

ACKNOWLEDGMENTS

We would like to thank Julia Cerruti for her help in the analysis of our functional annotation using R Studio.

This research was performed under the Facilities Integrating Collaborations for User Science (FICUS) program (proposal: 10.46936/fics.proj.2018.50403/60000053) and used resources at the DOE Joint Genome Institute and the Environmental Molecular Sciences Laboratory, which are DOE Office of Science User Facilities. Both facilities are sponsored by the Biological and Environmental Research program and operated under Contract Nos. DE-AC02-05CH11231 (JGI) and DE-AC05-76RL01830 (EMSL). This research was supported by Grant no. 2020041 awarded to Y.S. and R.M.B. from the United States-Israel Binational Science Foundation (BSF) and the Israel Ministry of Science and Technology (MOST) Grant No. 001126 to M.R.-B. Additional support was provided by a grant from the Simons Foundation award ID 721225 to S.T.D.

Y.S., R.M.B., M.G., and S.B. conceived of this study. S.B., F.Z., and S.W. collected the DNA and protein samples. S.B., F.Z., and S.W. performed the experiments together with Y.S., and S.T.D., S.T.H., and T.G.D.R. extracted and sequenced the DNA. C.D.N. extracted and processed protein samples that were analyzed by N.T. and R.M.B. M.R.-B. and C.K. performed the metagenomic analysis. C.K. wrote the manuscript under the guidance of Y.S., M.R.-B., R.M.B., M.G., and S.T.D. with further input from S.W., F.Z., and N.T.

AUTHOR AFFILIATIONS

[1]The Fredy and Nadine Herrmann Institute of Earth Sciences, Hebrew University of Jerusalem, Jerusalem, Israel
[2]The Interuniversity Institute for Marine Sciences in Eilat, Eilat, Israel
[3]Israel Oceanographic and Limnological Research, Haifa, Israel
[4]Microsensor Research Group, Max Planck Institute for Marine Microbiology, Bremen, Germany
[5]Lamont-Doherty Earth Observatory, Columbia University, New York, USA
[6]Earth and Biological Sciences, Pacific Northwest National Laboratory, Richland, Washington, USA
[7]Joint Genome Institute, Lawrence Berkeley National Laboratory, Berkeley, California, USA
[8]Department of Earth and Environmental Sciences, Columbia University, New York, USA
[9]GEOMAR, Helmholtz Center for Ocean Research, Kiel, Germany
[10]Environmental Molecular Sciences Laboratory, Pacific Northwest National Laboratory, Richland, Washington, USA
[11]College of Earth, Ocean, and Atmospheric Sciences, Oregon State University, Corvallis, Oregon, USA

AUTHOR ORCIDs

Coco Koedooder http://orcid.org/0000-0001-5454-8969
Rene M. Boiteau http://orcid.org/0000-0002-4127-4417
Maxim Rubin-Blum http://orcid.org/0000-0003-2953-3161

FUNDING

Funder	Grant(s)	Author(s)
Joint Genome Institute (JGI)	EMSL 50403	Yeala Shaked
Joint Genome Institute (JGI)	EMSL 50403	Rene M. Boiteau

Funder	Grant(s)	Author(s)
United States-Israel Binational Science Foundation (BSF)	2020041	Yeala Shaked
United States-Israel Binational Science Foundation (BSF)	2020041	Rene M. Boiteau
Ministry of Science, Technology and Space (MOST)	001126	Maxim Rubin-Blum
Simons Foundation (SF)	721225	Sonya T. Dyhrman

AUTHOR CONTRIBUTIONS

Coco Koedooder, Formal analysis, Investigation, Methodology, Software, Visualization, Writing – original draft | Futing Zhang, Investigation, Writing – review and editing | Siyuan Wang, Investigation, Writing – review and editing | Subhajit Basu, Investigation | Sheean T. Haley, Methodology, Resources, Writing – review and editing | Nikola Tolic, Resources, Software, Writing – review and editing | Carrie D. Nicora, Resources, Software | Tijana Glavina del Rio, Resources, Software | Sonya T. Dyhrman, Validation, Writing – review and editing | Martha Gledhill, Conceptualization, Funding acquisition, Investigation, Writing – review and editing | Rene M. Boiteau, Conceptualization, Funding acquisition, Investigation, Validation, Writing – review and editing | Maxim Rubin-Blum, Funding acquisition, Investigation, Supervision, Writing – review and editing | Yeala Shaked, Conceptualization, Funding acquisition, Supervision, Writing – review and editing

DATA AVAILABILITY

Raw sequences are available under the Gold Study ID number Gs0149370 and correspond to Gp0503298 (PRJNA1022442; SAMN37607796), Gp0503299 (PRJNA1022443, SAMN37607797), and Gp0503300 (PRJNA1022444, SAMN37607795). The 52 metagenomic assembled genomes analyzed in this study are available at NCBI under the BioProject number PRJNA944101. Proteomic samples are available at MassIVE under the project accession number MSV000091416. Scripts of the data analysis and figures can be found in the github depository: https://github.com/cocokoedooder/Trichodesmium_consortium

ADDITIONAL FILES

The following material is available online.

Supplemental Material

Figure S1 (mSystems00742-23-s0001.pdf). Relative abundances of bacterial MAGs from sampled *Trichodesmium* colonies indicate that the associated bacterial distribution is consistent.
Figure S2 (mSystems00742-23-s0002.pdf). Phylogenetic tree of the 52 MAGs together with MAGs assembled from 3 other *Trichodesmium* metagenomic data sets from colonies collected in the Red Sea, the Pacific Ocean, and the Atlantic Ocean.
Figure S3 (mSystems00742-23-s0003.pdf). NIS and NRPS-type siderophore biosynthesis pathways identified in MAGs of associated bacteria.
Figure S4 (mSystems00742-23-s0004.pdf). MAGs were probed for bacterial-interactive and particle-interactive traits.
Supplemental Tables (mSystems00742-23-s0005.xlsx). Tables S1-S4.

REFERENCES

1. 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
2. Zehr JP. 2011. Nitrogen fixation by marine cyanobacteria. Trends Microbiol 19:162–173. https://doi.org/10.1016/j.tim.2010.12.004
3. Capone DG, Burns JA, Montoya JP, Subramaniam A, Mahaffey C, Gunderson T, Michaels AF, Carpenter EJ, Capone C, Burns JA, Montoya

JP, Subramaniam A, Mahaffey C, Gunderson T, Michaels AF. 2005. Nitrogen fixation by *Trichodesmium* spp.: an important source of new nitrogen to the tropical and subtropical North Atlantic ocean. Global Biogeochem Cycles 19:n. https://doi.org/10.1029/2004GB002331
4. Hutchins DA, Fu FX, 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
5. Thompson AW, Zehr JP. 2013. Cellular interactions: lessons from the nitrogen-fixing cyanobacteria. J Phycol 49:1024–1035. https://doi.org/10.1111/jpy.12117
6. 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
7. Pierella Karlusich JJ, Pelletier E, Lombard F, Carsique M, Dvorak E, Colin S, Picheral M, Cornejo-Castillo FM, Acinas SG, Pepperkok R, Karsenti E, de Vargas C, Wincker P, Bowler C, Foster RA. 2021. Global distribution patterns of marine nitrogen-fixers by imaging and molecular methods. Nat Commun 12:4160. https://doi.org/10.1038/s41467-021-24299-y
8. Eichner M, Inomura K, Pierella Karlusich JJ, Shaked Y. 2023. Better together? Lessons on sociality from *Trichodesmium*. Trends Microbiol 31:1072–1084. https://doi.org/10.1016/j.tim.2023.05.001
9. Agarwal V, Inomura K, Mouw CB. 2022. Quantitative analysis of the trade-offs of colony formation for *Trichodesmium*. Microbiol Spectr 10:e0202522. https://doi.org/10.1128/spectrum.02025-22
10. Rubin M, Berman-Frank I, Shaked Y. 2011. Dust-and mineral-iron utilization by the marine dinitrogen-fixer *Trichodesmium*. Nature Geosci 4:529–534. https://doi.org/10.1038/ngeo1181
11. Wang S, Koedooder C, Zhang F, Kessler N, Eichner M, Shi D, Shaked Y. 2022. Colonies of the marine cyanobacterium *Trichodesmium* optimize dust utilization by selective collection and retention of nutrient-rich particles. iScience 25:103587. https://doi.org/10.1016/j.isci.2021.103587
12. Kessler N, Armoza-Zvuloni R, Wang S, Basu S, Weber PK, Stuart RK, Shaked Y. 2020. Selective collection of iron-rich dust particles by natural *Trichodesmium* colonies. ISME J 14:91–103. https://doi.org/10.1038/s41396-019-0505-x
13. Rubin M, Berman-Frank I, Shaked Y. 2011. Dust- and mineral-iron utilization by the marine dinitrogen-fixer *Trichodesmium*. Nature Geosci 4:529–534. https://doi.org/10.1038/ngeo1181
14. Pfreundt U, Słomka J, Schneider G, Sengupta A, Carrara F, Fernandez V, Ackermann M, Stocker R. 2023. Controlled motility in the cyanobacterium *Trichodesmium* regulates aggregate architecture. Science 380:830–835. https://doi.org/10.1126/science.adf2753
15. Hmelo L, Van Mooy B, Mincer T. 2012. Characterization of bacterial epibionts on the cyanobacterium *Trichodesmium*. Aquat Microb Ecol 67:1–14. https://doi.org/10.3354/ame01571
16. 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
17. Lee MD, Walworth NG, McParland EL, Fu FX, 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. Gradoville MR, Crump BC, Letelier RM, Church MJ, White AE. 2017. Microbiome of *Trichodesmium* colonies from the North Pacific subtropical gyre. Front Microbiol 8:1122. https://doi.org/10.3389/fmicb.2017.01122
19. Frischkorn Kyle R, 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
20. Frischkorn KR, Rouco M, Van Mooy BAS, Dyhrman ST. 2018. The*Trichodesmium*microbiome can modulate host N_2fixation. Limnol Oceanogr Lett 3:401–408. https://doi.org/10.1002/lol2.10092
21. Klawonn I, Eichner MJ, Wilson ST, Moradi N, Thamdrup B, Kümmel S, Gehre M, Khalili A, Grossart HP, Karl DM, Ploug H. 2020. Distinct nitrogen cycling and steep chemical gradients in *Trichodesmium* colonies. ISME J 14:399–412. https://doi.org/10.1038/s41396-019-0514-9
22. Basu S, Gledhill M, de Beer D, Prabhu Matondkar SG, Shaked Y. 2019. Colonies of marine cyanobacteria *Trichodesmium* interact with associated bacteria to acquire iron from dust. Commun Biol 2:284. https://doi.org/10.1038/s42003-019-0534-z
23. Frischkorn K.R, Haley ST, Dyhrman ST. 2018. Coordinated gene expression between *Trichodesmium* and its microbiome over day-night cycles in the North Pacific subtropical gyre. ISME J 12:997–1007. https://doi.org/10.1038/s41396-017-0041-5
24. Frischkorn Kyle R., Krupke A, Guieu C, Louis J, Rouco M, Salazar Estrada AE, Van Mooy BAS, Dyhrman ST. 2018. *Trichodesmium* physiological ecology and phosphate reduction in the western tropical South Pacific. Biogeosciences 15:5761–5778. https://doi.org/10.5194/bg-15-5761-2018
25. Gledhill M, Basu S, Shaked Y. 2019. Metallophores associated with *Trichodesmium* erythraeum colonies from the Gulf of Aqaba. Metallomics 11:1547–1557. https://doi.org/10.1039/c9mt00121b
26. Basu S, Shaked Y. 2018. Mineral iron utilization by natural and cultured *Trichodesmium* and associated bacteria. Limnol Oceanogr 63:2307–2320. https://doi.org/10.1002/lno.10939
27. Kraemer SM, Butler A, Borer P, Cervini-Silva J. 2005. Siderophores and the dissolution of iron-bearing minerals in marine systems. Rev Mineral Geochem 59:53–84. https://doi.org/10.2138/rmg.2005.59.4
28. Manck LE, Park J, Tully BJ, Poire AM, Bundy RM, Dupont CL, Barbeau KA. 2022. Petrobactin, a siderophore produced by alteromonas, mediates community iron acquisition in the global ocean. ISME J 16:358–369. https://doi.org/10.1038/s41396-021-01065-y
29. Roe KL, Barbeau K, Mann EL, Haygood MG. 2012. Acquisition of iron by *Trichodesmium* and associated bacteria in culture. Environ Microbiol 14:1681–1695. https://doi.org/10.1111/j.1462-2920.2011.02653.x
30. Polyviou D, Baylay AJ, Hitchcock A, Robidart J, Moore CM, Bibby TS. 2017. Desert dust as a source of iron to the globally important diazotroph *Trichodesmium*. Front Microbiol 8:2683. https://doi.org/10.3389/fmicb.2017.02683
31. Chappell PD, Webb EA. 2010. A molecular assessment of the iron stress response in the two phylogenetic clades of *Trichodesmium*. Environ Microbiol 12:13–27. https://doi.org/10.1111/j.1462-2920.2009.02026.x
32. Snow JT, Polyviou D, Skipp P, Chrismas NAM, Hitchcock A, Geider R, Moore CM, Bibby TS. 2015. Quantifying integrated proteomic responses to iron stress in the globally important marine diazotroph *Trichodesmium*. PLoS One 10:e0142626. https://doi.org/10.1371/journal.pone.0142626
33. Berman-Frank I, Quigg A, Finkel ZV, Irwin AJ, Haramaty L. 2007. Nitrogen-fixation strategies and Fe requirements in cyanobacteria. Limnol Oceanogr 52:2260–2269. https://doi.org/10.4319/lo.2007.52.5.2260
34. Jickells TD, An ZS, Andersen KK, Baker AR, Bergametti G, Brooks N, Cao JJ, Boyd PW, Duce RA, Hunter KA, Kawahata H, Kubilay N, laRoche J, Liss PS, Mahowald N, Prospero JM, Ridgwell AJ, Tegen I, Torres R. 2005. Global iron connections between desert dust, ocean biogeochemistry, and climate. Science 308:67–71. https://doi.org/10.1126/science.1105959
35. Kok JF, Ward DS, Mahowald NM, Evan AT. 2018. Global and regional importance of the direct dust-climate feedback. Nat Commun 9:241. https://doi.org/10.1038/s41467-017-02620-y
36. Post AF, Dedej Z, Gottlieb R, Li H, Thomas DN, El-Absawi M, El-Naggar A, El-Gharabawi M, Sommer U. 2002. Spatial and temporal distribution of *Trichodesmium* spp. in the stratified Gulf of Aqaba, Red Sea. Mar Ecol Prog Ser 239:241–250. https://doi.org/10.3354/meps239241
37. Bowers RM, Kyrpides NC, Stepanauskas R, Harmon-Smith M, Doud D, Reddy TBK, Schulz F, Jarett J, Rivers AR, Eloe-Fadrosh EA, et al. 2017. Minimum information about a single amplified genome (MISAG) and a metagenome-assembled genome (MIMAG) of bacteria and archaea. Nat Biotechnol 35:725–731. https://doi.org/10.1038/nbt.3893
38. Koedooder C, Landou E, Zhang F, Wang S, Basu S, Berman-Frank I, Shaked Y, Rubin-Blum M. 2022. Metagenomes of Red Sea subpopulations challenge the use of marker genes and morphology to assess *Trichodesmium* diversity. Front Microbiol 13:879970. https://doi.org/10.3389/fmicb.2022.879970
39. Dang H, Lovell CR. 2000. Bacterial primary colonization and early succession on surfaces in marine waters as determined by amplified rRNA gene restriction analysis and sequence analysis of 16S rRNA

genes. Appl Environ Microbiol 66:467–475. https://doi.org/10.1128/AEM.66.2.467-475.2000
40. Dang H, Li T, Chen M, Huang G. 2008. Cross-ocean distribution of Rhodobacterales bacteria as primary surface colonizers in temperate coastal marine waters. Appl Environ Microbiol 74:52–60. https://doi.org/10.1128/AEM.01400-07
41. Duret MT, Lampitt RS, Lam P. 2019. Prokaryotic niche partitioning between suspended and sinking marine particles. Environ Microbiol Rep 11:386–400. https://doi.org/10.1111/1758-2229.12692
42. Koch H, Germscheid N, Freese HM, Noriega-Ortega B, Lücking D, Berger M, Qiu G, Marzinelli EM, Campbell AH, Steinberg PD, Overmann J, Dittmar T, Simon M, Wietz M. 2020. Genomic, metabolic and phenotypic variability shapes ecological differentiation and intraspecies interactions of *Alteromonas macleodii*. Sci Rep 10:1–14. https://doi.org/10.1038/s41598-020-57526-5
43. Hou S, López-Pérez M, Pfreundt U, Belkin N, Stüber K, Huettel B, Reinhardt R, Berman-Frank I, Rodriguez-Valera F, Hess WR. 2018. Benefit from decline: the primary transcriptome of *Alteromonas macleodii* str. Te101 during *Trichodesmium* demise. ISME J 12:981–996. https://doi.org/10.1038/s41396-017-0034-4
44. 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
45. Held NA, Webb EA, McIlvin MM, Hutchins DA, Cohen NR, Moran DM, Kunde K, Lohan MC, Mahaffey C, Woodward EMS, Saito MA. 2020. Co-occurrence of Fe and P stress in natural populations of the marine diazotroph *Trichodesmium*. Biogeosciences 17:2537–2551. https://doi.org/10.5194/bg-17-2537-2020
46. Carroll CS, Moore MM. 2018. Ironing out siderophore biosynthesis: a review of non-ribosomal peptide synthetase (NRPS)-independent siderophore synthetases. Crit Rev Biochem Mol Biol 53:356–381. https://doi.org/10.1080/10409238.2018.1476449
47. Gulick AM. 2017. Nonribosomal peptide synthetase biosynthetic clusters of ESKAPE pathogens. Nat Prod Rep 34:981–1009. https://doi.org/10.1039/c7np00029d
48. Reitz ZL, Medema MH. 2022. Genome mining strategies for metallophore discovery. Curr Opin Biotechnol 77:102757. https://doi.org/10.1016/j.copbio.2022.102757
49. Butler A, Harder T, Ostrowski AD, Carrano CJ. 2021. Photoactive siderophores: structure, function and biology. J Inorg Biochem 221:111457. https://doi.org/10.1016/j.jinorgbio.2021.111457
50. Amin Shady A, Green DH, Küpper FC, Carrano CJ. 2009. Vibrioferrin, an unusual marine siderophore: iron binding, photochemistry, and biological implications. Inorg Chem 48:11451–11458. https://doi.org/10.1021/ic9016883
51. Amin SA, Green DH, Hart MC, Küpper FC, Sunda WG, Carrano CJ. 2009. Photolysis of iron-siderophore chelates promotes bacterial-algal mutualism. Proc Natl Acad Sci U S A 106:17071–17076. https://doi.org/10.1073/pnas.0905512106
52. Chappell PD, Moffett JW, Hynes AM, Webb EA. 2012. Molecular evidence of iron limitation and availability in the global diazotroph *Trichodesmium*. ISME J 6:1728–1739. https://doi.org/10.1038/ismej.2012.13
53. Kranzler C, Lis H, Shaked Y, Keren N. 2011. The role of reduction in iron uptake processes in a unicellular, planktonic cyanobacterium. Environ Microbiol 13:2990–2999. https://doi.org/10.1111/j.1462-2920.2011.02572.x
54. Cordero OX, Ventouras LA, DeLong EF, Polz MF. 2012. Public good dynamics drive evolution of iron acquisition strategies in natural bacterioplankton populations. Proc Natl Acad Sci U S A 109:20059–20064. https://doi.org/10.1073/pnas.1213344109
55. Barbeau K, Rue EL, Trick CG, Bruland KW, Butler A. 2003. Photochemical reactivity of siderophores produced by marine heterotrophic bacteria and cyanobacteria based on characteristic Fe(III) binding groups. Limnol Oceanogr 48:1069–1078. https://doi.org/10.4319/lo.2003.48.3.1069
56. Martinez JS, Carter-Franklin JN, Mann EL, Martin JD, Haygood MG, Butler A. 2003. Structure and membrane affinity of a suite of amphiphilic siderophores produced by a marine bacterium. Proc Natl Acad Sci U S A 100:3754–3759. https://doi.org/10.1073/pnas.0637444100
57. Lee MD, Webb EA, Walworth NG, Fu FX, Held NA, Saito MA, Hutchins DA. 2018. Transcriptional activities of the microbial consortium living with the marine nitrogenfixing cyanobacterium *Trichodesmium* reveal potential roles in community-level nitrogen cycling. Appl Environ Microbiol 84:2026–2043. https://doi.org/10.1128/AEM.02026-17
58. Kramer J, Özkaya Ö, Kümmerli R. 2020. Bacterial siderophores in community and host interactions. Nat Rev Microbiol 18:152–163. https://doi.org/10.1038/s41579-019-0284-4
59. Van Mooy BAS, Krupke A, Dyhrman ST, Fredricks HF, Frischkorn KR, Ossolinski JE, Repeta DJ, Rouco M, Seewald JD, Sylva SP. 2015. Phosphorus cycling. major role of planktonic phosphate reduction in the marine phosphorus redox cycle. Science 348:783–785. https://doi.org/10.1126/science.aaa8181
60. Mackey KRM, Labiosa RG, Calhoun M, Street JH, Post AF, Paytan A. 2007. Phosphorus availability, phytoplankton community dynamics, and taxon-specific phosphorus status in the Gulf of Aqaba, Red Sea. Limnol Oceanogr 52:873–885. https://doi.org/10.4319/lo.2007.52.2.0873
61. Karl DM, Björkman KM. 2015. Dynamics of dissolved organic phosphorus, p 233–334. In Biogeochemistry of marine dissolved organic matter
62. Orchard ED, Webb EA, Dyhrman ST. 2009. Molecular analysis of the phosphorus starvation response in *Trichodesmium* spp. Environ Microbiol 11:2400–2411. https://doi.org/10.1111/j.1462-2920.2009.01968.x
63. Frischkorn KR, Haley ST, Dyhrman ST. 2019. Transcriptional and proteomic choreography under phosphorus deficiency and re-supply in the N2 fixing cyanobacterium *Trichodesmium* erythraeum. Front Microbiol 10:330. https://doi.org/10.3389/fmicb.2019.00330
64. Metcalf WW, Wanner BL. 1991. Involvement of the *Escherichia coli* phn (*psiD*) gene cluster in assimilation of phosphorus in the form of phosphonates, phosphite, Pi esters, and Pi. J Bacteriol 173:587–600. https://doi.org/10.1128/jb.173.2.587-600.1991
65. White AK, Metcalf WW. 2007. Microbial metabolism of reduced phosphorus compounds. Annu Rev Microbiol 61:379–400. https://doi.org/10.1146/annurev.micro.61.080706.093357
66. Stosiek N, Talma M, Klimek-Ochab M. 2020. Carbon-phosphorus lyase-the state of the art. Appl Biochem Biotechnol 190:1525–1552. https://doi.org/10.1007/s12010-019-03161-4
67. Kim AD, Baker AS, Dunaway-Mariano D, Metcalf WW, Wanner BL, Martin BM. 2002. The 2-aminoethylphosphonate-specific transaminase of the 2-aminoethylphosphonate degradation pathway. J Bacteriol 184:4134–4140. https://doi.org/10.1128/JB.184.15.4134-4140.2002
68. Lockwood S, Greening C, Baltar F, Morales SE. 2022. Global and seasonal variation of marine phosphonate metabolism. ISME J 16:2198–2212. https://doi.org/10.1038/s41396-022-01266-z
69. Hilderbrand RL. 2018. The effects of synthetic phosphonates on living systems, p 139–169. In The role of phosphonates in living systems
70. Coates CJ, Wyman M. 2017. A denitrifying community associated with a major, marine nitrogen fixer. Environ Microbiol 19:4978–4992. https://doi.org/10.1111/1462-2920.14007
71. Allen AE, Booth MG, Frischer ME, Verity PG, Zehr JP, Zani S. 2001. Diversity and detection of nitrate assimilation genes in marine bacteria. Appl Environ Microbiol 67:5343–5348. https://doi.org/10.1128/AEM.67.11.5343-5348.2001
72. Kuypers MMM, Marchant HK, Kartal B. 2018. The microbial nitrogen-cycling network. Nat Rev Microbiol 16:263–276. https://doi.org/10.1038/nrmicro.2018.9
73. Eichner M, Basu S, Wang S, de Beer D, Shaked Y. 2020. Mineral iron dissolution in*Trichodesmium*colonies: the role of $O_2$and pH microenvironments. Limnol Oceanogr 65:1149–1160. https://doi.org/10.1002/lno.11377
74. Dalsing BL, Allen C. 2014. Nitrate assimilation contributes to Ralstonia solanacearum root attachment, stem colonization, and virulence. J Bacteriol 196:949–960. https://doi.org/10.1128/JB.01378-13
75. Durand S, Guillier M. 2021. Transcriptional and post-transcriptional control of the nitrate respiration in bacteria. Front Mol Biosci 8:667758. https://doi.org/10.3389/fmolb.2021.667758
76. Cannon J, Sanford RA, Connor L, Yang WH, Chee-Sanford J. 2019. Optimization of PCR primers to detect phylogenetically diverse *nrfA* genes associated with nitrite ammonification. J Microbiol Methods 160:49–59. https://doi.org/10.1016/j.mimet.2019.03.020

77. Pajares S, Ramos R. 2019. Processes and microorganisms involved in the marine nitrogen cycle: knowledge and gaps. Front Mar Sci 6:739. https://doi.org/10.3389/fmars.2019.00739
78. Graf DRH, Jones CM, Hallin S, de Crécy-Lagard V. 2014. Intergenomic comparisons highlight modularity of the denitrification pathway and underpin the importance of community structure for N2O emissions. PLoS One 9:e114118. https://doi.org/10.1371/journal.pone.0114118
79. Kreimer A, Borenstein E, Gophna U, Ruppin E. 2008. The evolution of modularity in bacterial metabolic networks. Proc Natl Acad Sci U S A 105:6976–6981. https://doi.org/10.1073/pnas.0712149105
80. Hutchins DA, Capone DG. 2022. The marine nitrogen cycle: new developments and global change. Nat Rev Microbiol 20:401–414. https://doi.org/10.1038/s41579-022-00752-7
81. Smriga S, Ciccarese D, Babbin AR. 2021. Denitrifying bacteria respond to and shape microscale gradients within particulate matrices. Commun Biol 4:570. https://doi.org/10.1038/s42003-021-02102-4
82. Stief P, Kamp A, Thamdrup B, Glud RN. 2016. Anaerobic nitrogen turnover by sinking diatom aggregates at varying ambient oxygen levels. Front Microbiol 7:98. https://doi.org/10.3389/fmicb.2016.00098
83. Paerl HW, Bebout BM. 1988. Direct measurement of O2-depleted microzones in marine *Oscillatoria*: relation to N2 fixation. Science 241:442–445. https://doi.org/10.1126/science.241.4864.442
84. Paerly HW, Bebout BM, Prufert LE. 1989. Bacterial associations with marine *Oscillatoria* sp. (*Trichodesmium* sp.) populations: ecophysiological implications1. J Phycol 25:773–784. https://doi.org/10.1111/j.0022-3646.1989.00773.x
85. McGenity TJ. 2019. Taxonomy, genomics and ecophysiology of hydrocarbon-degrading microbes. Springer International Publishing.
86. Ho A, Di Lonardo DP, Bodelier PLE. 2017. Revisiting life strategy concepts in environmental microbial ecology. FEMS Microbiol Ecol 93:6. https://doi.org/10.1093/femsec/fix006
87. Lauro FM, McDougald D, Thomas T, Williams TJ, Egan S, Rice S, DeMaere MZ, Ting L, Ertan H, Johnson J, Ferriera S, Lapidus A, Anderson I, Kyrpides N, Munk AC, Detter C, Han CS, Brown MV, Robb FT, Kjelleberg S, Cavicchioli R. 2009. The genomic basis of trophic strategy in marine bacteria. Proc Natl Acad Sci U S A 106:15527–15533. https://doi.org/10.1073/pnas.0903507106
88. Zoccarato L, Sher D, Miki T, Segrè D, Grossart HP. 2022. A comparative whole-genome approach identifies bacterial traits for marine microbial interactions. Commun Biol 5:276. https://doi.org/10.1038/s42003-022-03184-4
89. Debeljak P, Blain S, Bowie A, van der Merwe P, Bayer B, Obernosterer I. 2021. Homeostasis drives intense microbial trace metal processing on marine particles. Limnol Oceanogr 66:3842–3855. https://doi.org/10.1002/lno.11923
90. 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
91. 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
92. Hmelo LR. 2017. Quorum sensing in marine microbial environments. Ann Rev Mar Sci 9:257–281. https://doi.org/10.1146/annurev-marine-010816-060656
93. Azam F, Fenchel T, Field J, Gray J, Meyer-Reil L, Thingstad F. 1983. The ecological role of water-column microbes in the sea*. Mar Ecol Prog Ser 10:257–263. https://doi.org/10.3354/meps010257
94. Kieser S, Brown J, Zdobnov EM, Trajkovski M, McCue LA. 2020. ATLAS: a snakemake workflow for assembly, annotation, and genomic binning of metagenome sequence data. BMC Bioinformatics 21:257. https://doi.org/10.1186/s12859-020-03585-4
95. Bushnell B. 2021. Bbtools software package. Sourceforge.Net/projects/Bbmap/
96. Bushnell B, Rood J, Singer E. 2017. BBMerge - accurate paired shotgun read merging via overlap. PLoS One 12:e0185056. https://doi.org/10.1371/journal.pone.0185056
97. Nurk S, Meleshko D, Korobeynikov A, Pevzner PA. 2017. metaSPAdes: a new versatile metagenomic assembler. Genome Res 27:824–834. https://doi.org/10.1101/gr.213959.116
98. Kang DD, Li F, Kirton E, Thomas A, Egan R, An H, Wang Z. 2019. MetaBAT 2: an adaptive binning algorithm for robust and efficient genome reconstruction from metagenome assemblies. PeerJ 7:e7359. https://doi.org/10.7717/peerj.7359
99. Wu YW, Simmons BA, Singer SW. 2016. MaxBin 2.0: an automated binning algorithm to recover genomes from multiple metagenomic datasets. Bioinformatics 32:605–607. https://doi.org/10.1093/bioinformatics/btv638
100. Nissen JN, Johansen J, Allesøe RL, Sønderby CK, Armenteros JJA, Grønbech CH, Jensen LJ, Nielsen HB, Petersen TN, Winther O, Rasmussen S. 2021. Improved metagenome binning and assembly using deep variational autoencoders. Nat Biotechnol 39:555–560. https://doi.org/10.1038/s41587-020-00777-4
101. Sieber CMK, Probst AJ, Sharrar A, Thomas BC, Hess M, Tringe SG, Banfield JF. 2018. Recovery of genomes from metagenomes via a dereplication, aggregation and scoring strategy. Nat Microbiol 3:836–843. https://doi.org/10.1038/s41564-018-0171-1
102. Olm MR, Brown CT, Brooks B, Banfield JF. 2017. dRep: a tool for fast and accurate genomic comparisons that enables improved genome recovery from metagenomes through de-replication. ISME J 11:2864–2868. https://doi.org/10.1038/ismej.2017.126
103. Chklovski A, Parks DH, Woodcroft BJ, Tyson GW. 2022. CheckM2: a rapid, scalable and accurate tool for assessing microbial genome quality using machine learning. bioRxiv. https://doi.org/10.1101/2022.07.11.499243
104. Hyatt D, Chen G-L, Locascio PF, Land ML, Larimer FW, Hauser LJ. 2010. Prodigal: prokaryotic gene recognition and translation initiation site identification. BMC Bioinformatics 11:1–11. https://doi.org/10.1186/1471-2105-11-119
105. Chaumeil P-A, Mussig AJ, Hugenholtz P, Parks DH, Hancock J. 2020. GTDB-Tk: a toolkit to classify genomes with the genome taxonomy database. Bioinformatics 36:1925–1927. https://doi.org/10.1093/bioinformatics/btz848
106. Lee MD. 2019. GToTree: a user-friendly workflow for phylogenomics. Bioinformatics 35:4162–4164. https://doi.org/10.1093/bioinformatics/btz188
107. Kalyaanamoorthy S, Minh BQ, Wong TKF, von Haeseler A, Jermiin LS. 2017. ModelFinder: fast model selection for accurate phylogenetic estimates. Nat Methods 14:587–589. https://doi.org/10.1038/nmeth.4285
108. Nguyen L-T, Schmidt HA, von Haeseler A, Minh BQ. 2015. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. Mol Biol Evol 32:268–274. https://doi.org/10.1093/molbev/msu300
109. Coleman GA, Davín AA, Mahendrarajah TA, Szánthó LL, Spang A, Hugenholtz P, Szöllősi GJ, Williams TA. 2021. A rooted phylogeny resolves early bacterial evolution. Science 372:eabe0511. https://doi.org/10.1126/science.abe0511
110. Kanehisa M, Sato Y, Morishima K. 2016. BlastKOALA and GhostKOALA: KEGG tools for functional characterization of genome and metagenome sequences. J Mol Biol 428:726–731. https://doi.org/10.1016/j.jmb.2015.11.006
111. Bateman A, Coin L, Durbin R, Finn RD, Hollich V, Griffiths-Jones S, Khanna A, Marshall M, Moxon S, Sonnhammer ELL, Studholme DJ, Yeats C, Eddy SR. 2004. The Pfam protein families database. Nucleic Acids Res 32:D138–41. https://doi.org/10.1093/nar/gkh121
112. Aziz RK, Bartels D, Best AA, DeJongh M, Disz T, Edwards RA, Formsma K, Gerdes S, Glass EM, Kubal M, Meyer F, Olsen GJ, Olson R, Osterman AL, Overbeek RA, McNeil LK, Paarmann D, Paczian T, Parrello B, Pusch GD, Reich C, Stevens R, Vassieva O, Vonstein V, Wilke A, Zagnitko O. 2008. The RAST server: rapid annotations using subsystems technology. BMC Genomics 9:1–15. https://doi.org/10.1186/1471-2164-9-75
113. Meyer F, Overbeek R, Rodriguez A. 2009. FIGfams: yet another set of protein families. Nucleic Acids Res 37:6643–6654. https://doi.org/10.1093/nar/gkp698
114. Zhou Z, Tran PQ, Breister AM, Liu Y, Kieft K, Cowley ES, Karaoz U, Anantharaman K. 2022. METABOLIC: high-throughput profiling of microbial genomes for functional traits, metabolism, biogeochemistry, and community-scale functional networks. Microbiome 10:33. https://doi.org/10.1186/s40168-021-01213-8

115. Letunic I, Bork P. 2021. Interactive tree of life (iTOL) v5: an online tool for phylogenetic tree display and annotation. Nucleic Acids Res 49:W293–W296. https://doi.org/10.1093/nar/gkab301
116. Garber AI, Nealson KH, Okamoto A, McAllister SM, Chan CS, Barco RA, Merino N. 2020. FeGenie: a comprehensive tool for the identification of iron genes and iron gene neighborhoods in genome and metagenome assemblies. Front Microbiol 11:37. https://doi.org/10.3389/fmicb.2020.00037
117. Katoh K, Kuma K, Toh H, Miyata T. 2005. MAFFT version 5: improvement in accuracy of multiple sequence alignment. Nucleic Acids Res 33:511–518. https://doi.org/10.1093/nar/gki198
118. Kautsar SA, Blin K, Shaw S, Navarro-Muñoz JC, Terlouw BR, van der Hooft JJJ, van Santen JA, Tracanna V, Suarez Duran HG, Pascal Andreu V, Selem-Mojica N, Alanjary M, Robinson SL, Lund G, Epstein SC, Sisto AC, Charkoudian LK, Collemare J, Linington RG, Weber T, Medema MH. 2020. MIBiG 2.0: a repository for biosynthetic gene clusters of known function. Nucleic Acids Res 48:D454–D458. https://doi.org/10.1093/nar/gkz882
119. Webb E, Claas K, Downs D. 1998. thiBPQ encodes an ABC transporter required for transport of thiamine and thiamine pyrophosphate in *Salmonella typhimurium*. J Biol Chem 273:8946–8950. https://doi.org/10.1074/jbc.273.15.8946
120. Hebbeln P, Rodionov DA, Alfandega A, Eitinger T. 2007. Biotin uptake in prokaryotes by solute transporters with an optional ATP-binding cassette-containing module. Proc Natl Acad Sci U S A 104:2909–2914. https://doi.org/10.1073/pnas.0609905104
121. Chimento DP, Mohanty AK, Kadner RJ, Wiener MC. 2003. Substrate-induced transmembrane signaling in the cobalamin transporter BtuB. Nat Struct Biol 10:394–401. https://doi.org/10.1038/nsb914
122. Heller K, Mann BJ, Kadner RJ. 1985. Cloning and expression of the gene for the vitamin B12 receptor protein in the outer membrane of *Escherichia coli*. J Bacteriol 161:896–903. https://doi.org/10.1128/jb.161.3.896-903.1985
123. Sparacino-Watkins C, Stolz JF, Basu P. 2014. Nitrate and periplasmic nitrate reductases. Chem Soc Rev 43:676–706. https://doi.org/10.1039/c3cs60249d
124. Wei W, Isobe K, Nishizawa T, Zhu L, Shiratori Y, Ohte N, Koba K, Otsuka S, Senoo K. 2015. Higher diversity and abundance of denitrifying microorganisms in environments than considered previously. ISME J 9:1954–1965. https://doi.org/10.1038/ismej.2015.9
125. Richardson DJ, Berks BC, Russell DA, Spiro S, Taylor CJ. 2001. Functional, biochemical and genetic diversity of prokaryotic nitrate reductases. Cell Mol Life Sci 58:165–178. https://doi.org/10.1007/PL00000845
126. Wang Q, Li H, Post AF. 2000. Nitrate assimilation genes of the marine diazotrophic, filamentous cyanobacterium *Trichodesmium* sp. strain WH9601. J Bacteriol 182:1764–1767. https://doi.org/10.1128/JB.182.6.1764-1767.2000
127. Kim S, Gupta N, Pevzner PA. 2008. Spectral probabilities and generating functions of tandem mass spectra: a strike against decoy databases. J Proteome Res 7:3354–3363. https://doi.org/10.1021/pr8001244
128. Seemann T. 2014. Prokka: rapid prokaryotic genome annotation. Bioinformatics 30:2068–2069. https://doi.org/10.1093/bioinformatics/btu153
129. Florens L, Carozza MJ, Swanson SK, Fournier M, Coleman MK, Workman JL, Washburn MP. 2006. Analyzing chromatin remodeling complexes using shotgun proteomics and normalized spectral abundance factors. Methods 40:303–311. https://doi.org/10.1016/j.ymeth.2006.07.028