ARTICLE

Transcriptional patterns identify resource controls on the diazotroph *Trichodesmium* in the Atlantic and Pacific oceans

Mónica Rouco[1,2] · Kyle R. Frischkorn[1,2] · Sheean T. Haley[1] · Harriet Alexander [3] · Sonya T. Dyhrman[1,2]

Abstract

The N$_2$-fixing cyanobacterium *Trichodesmium* is intensely studied because of the control this organism exerts over the cycling of carbon and nitrogen in the low nutrient ocean gyres. Although iron (Fe) and phosphorus (P) bioavailability are thought to be major drivers of *Trichodesmium* distributions and activities, identifying resource controls on *Trichodesmium* is challenging, as Fe and P are often organically complexed and their bioavailability to a single species in a mixed community is difficult to constrain. Further, Fe and P geochemistries are linked through the activities of metalloenzymes, such as the alkaline phosphatases (APs) PhoX and PhoA, which are used by microbes to access dissolved organic P (DOP). Here we identified significant correlations between *Trichodesmium*-specific transcriptional patterns in the North Atlantic (NASG) and North Pacific Subtropical Gyres (NPSG) and patterns in Fe and P biogeochemistry, with the relative enrichment of Fe stress markers in the NPSG, and P stress markers in the NASG. We also observed the differential enrichment of Fe-requiring PhoX transcripts in the NASG and Fe-insensitive PhoA transcripts in the NPSG, suggesting that metalloenzyme switching may be used to mitigate Fe limitation of DOP metabolism in *Trichodesmium*. This trait may underpin *Trichodesmium* success across disparate ecosystems.

Introduction

Nitrogen (N)-fixing diazotrophs like cyanobacteria in the genus *Trichodesmium* are common in surface waters of the vast tropical and subtropical oligotrophic regions of the ocean such as the North Pacific and North Atlantic sub-tropical gyres (NPSG and NASG) [1, 2]. Atmospheric dinitrogen (N$_2$) fixation by diazotrophs such as *Tricho-desmium* plays an important role in these low nutrient

✉ Sonya T. Dyhrman
sdyhrman@ldeo.columbia.edu

[1] Biology and Paleo Environment Division, Lamont-Doherty Earth Observatory, Columbia University, Palisades, NY 10964, USA

[2] Department of Earth and Environmental Sciences, Columbia University, New York, NY 10027, USA

[3] Department of Population Health and Reproduction, School of Veterinary Medicine, University of California, Davis, CA 95616, USA

subtropical gyres by fueling primary production and carbon (C) export in the ocean [1,3–7]. N$_2$ fixation by the genus *Trichodesmium* alone accounts for approximately 50% of the input of new N in these oligotrophic regions, contributing significantly to the cycling of N and C [8, 9]. In addition, new estimates reveal that *Trichodesmium* contributes to the turnover of reduced phosphorus (P) compounds such as phosphite and phosphonate [10], driving the cycling of P and the production of the greenhouse gas methane in low P oligotrophic oceans [11, 12].

Characterization of the geochemical constraints of oceanic N$_2$ fixation is the subject of intense study. P and iron (Fe) are recognized as major drivers of the distributions and activities of diazotrophs in the oligotrophic ocean gyres over both modern and geological time scales [5, 13]. Fe bioavailability is particularly important due to the high Fe quota of enzymes involved in N$_2$ fixation (e.g., nitrogenase) and photosynthetic electron transport [14]. Models used to predict patterns in global N$_2$ fixation suggest that P constrains diazotroph growth in the NASG, whereas Fe constrains diazotroph growth in the NPSG [13]. However, parameterizing phytoplankton activities and growth in biogeochemical models is challenging, thus leading to uncertainty in their outputs. Model parameters are typically

derived from culture measurements of P and Fe uptake and these studies are predominantly done with model strains that are not representative of field populations. For example, *Trichodesmium erythraeum* IMS101 is a commonly used isolate in laboratory-based studies [2], yet it belongs to a different phylogenetic clade than that which dominates the NASG and NPSG [15, 16]. An additional challenge is that the geochemical conditions in culture are not fully representative of the field, where Fe and P may both be low and where bioavailability varies over the range of chemical species and organic complexes that exist, which are difficult to characterize in marine environments [17, 18]. Recent modeling efforts have focused on some of these challenges by using resource competition theory to evaluate the role of organically complexed P in fueling NASG N_2 fixation [19] and using resource ratio theory to highlight the apparent importance of both Fe and P in driving *Trichodesmium* biogeography and N_2 fixation across oceanic provinces [20, 21]. These studies support the regional importance of Fe and P biogeochemistry in *Trichodesmium* physiological ecology.

Trichodesmium has acquired a number of adaptations to meet P and Fe demands in the oligotrophic regions where it occurs, such as the ability to regulate cellular P [22] and Fe quotas [23], take up different Fe [24] and P species [25], and access dissolved organic P (DOP) with enzymes like alkaline phosphatases (APs) and C-P lyases [23,26–29]. Notably, some of these adaptations could impose tradeoffs by increasing requirements for metal co-factors. For instance, access to the DOP pool might be limited by Fe bioavailability, because PhoX, the most widely distributed type of AP among marine bacteria [30], requires Fe as a metal cofactor [31, 32]. In fact, whole-water field incubations with added Fe in low Fe regions showed an enhancement of AP activity [33]. Such potential interactions and adaptations further exacerbate the complexities surrounding the identification and modeling of the geochemical drivers of N_2 fixation.

In the field, *Trichodesmium* colonies are complex communities with a rich consortium of heterotrophic epibionts and other microorganisms [34–36]. As a consequence, many useful approaches for evaluating Fe and P bioavailability, such as enzymatic activity assays, are not necessarily specific to *Trichodesmium* [27, 35, 37]. Molecular approaches show promise in this regard since they can be tuned to examine *Trichodesmium*-specific signals, especially as the pathways for resource acquisition are increasingly better understood [25, 29,38–40]. *Trichodesmium*-specific transcriptional profiles may yield insights into the tradeoffs in resource acquisition and limitation patterns, as protein profiles have done for populations of the cyanobacteria *Prochlorococcus* [41]. Uncertainties remain in characterizing how resource bioavailability drives *Trichodesmium* physiological ecology in the field, and this, in turn, limits modeling efforts. Here, metatrancriptome profiling was used to identify *Trichodesmium*-specific signals from colonies in the NASG and NPSG in order to ascertain ecosystem-specific geochemical drivers of this globally significant diazotroph.

Material and methods

Field sample collection

Sampling took place during a cruise transect in the NASG aboard the R/ V Oceanus (OC471, April and May 2011) and as part of a time-series sampling in the NPSG at Station ALOHA, situated at 22.75°N, 158°W (HOE-DYLAN#7 and #9, August 2012) (Table S1, Fig. S1). A total of 12 samples were sequenced, 9 in the NASG and 3 in the NPSG. All *Trichodesmium* colonies were collected from near the surface (approximately within the upper 25 m) using a handheld 130 μm net, within 2 h of local solar noon. As *nif* genes in *Trichodesmium* display strong diel signals [42], samples were always collected at the same point relative to local noon to remove potential diel variability in overall gene expression. Single colonies were picked and rinsed three times in 0.2 μm filtered local surface water collected at 5 m with a Rosette sampling device. Between 10 and 15 washed *Trichodesmium* colonies per station were filtered onto 47 mm, 10 μm polycarbonate filters, which were then placed in 2 ml cryovials, snap-frozen and stored in liquid N until RNA extraction was performed in the laboratory. Total time from sample collection to preservation was roughly 15 min. In situ profiles of temperature and salinity, and photosynthetically active radiation (PAR) in the NASG were measured by a conductivity, temperature, and depth instrument deployed at each station. In the NPSG, PAR at the sea-surface was measured using a cosine sensor (LI-COR model LI-192) mounted on the top deck of the ship. Vertical profiles of downwelling PAR irradiance were obtained using a Free-Falling Optical Profiler (Satlantic). Both measurements were then used to compute the percent level of PAR with respect to the surface value on 1 m depth bins in the water column. For the NASG samples, nitrite + nitrate ($NO_2^- + NO_3^-$) and soluble reactive silica (Si) were determined from 125 ml of water filtered through a 0.2 μm polycarbonate filter and stored frozen (–20°C) in 10% HCl-cleaned bottles until analysis following the facility's protocols at the Chesapeake Bay Lab at the University of Maryland (http://nasl.cbl.umces.edu/methods/WCC.html). Low-level dissolved inorganic phosphate (PO_4^{3-}) was assayed using a modified MAGIC method [43] with a detection limit of 2.5 nM. All of the NPSG samples were assayed using the protocols of the Hawaii Ocean Time-

series (HOT) program (http://hahana.soest.hawaii.edu/hot/protocols/protocols.html). Dissolved Fe (dFe) concentrations associated with the NPSG samples were collected, analyzed, and reported in Fitzsimmons et al. [17]. Surface dFe concentrations in the NASG were not directly measured in this study, and were computed as the dFe climatological average using data from stations sampled across the same geographical transect as this study (10°N-30°N and 70°W-25°W) throughout different seasons and years. The dFe climatological average was consistent with discrete surface values measured on a similar spring NASG transect reported by Chappell et al. [44]. Notably, all discrete PO_4^{3-} and dFe values were consistent with the regional NASG and NPSG climatology [28, 44–56].

## RNA extraction, mRNA purification, and sequencing

RNA was extracted from each of the filters using the RNeasy Mini Kit (Qiagen), following a modified version of the purification protocol for yeast. Briefly, lysis buffer with 0.01% of β-Mercaptoethanol and RNA-clean zirconia/Si beads (0.5 mm) were added to the filter and samples were vortexed for 5 min, placed on ice for 1 min, and vortexed again for 5 min. Samples were then processed following the remainder of the yeast protocol, as outlined by the manufacturer, and RNA was eluted in water. To eliminate potential DNA contamination, RNA was treated using the RNase-Free DNase Set (Qiagen) and then further purified and concentrated using the RNA Cleanup Protocol from the RNeasy Mini Kit (Qiagen). The RNA was eluted in Tris-EDTA (TE) buffer and potential eukaryotic RNA was removed using the MICROB*Enrich* kit (Thermo Fisher Scientific). Finally, for enrichment of bacterial mRNA and removal of ribosomal RNA, the enriched bacterial RNA was processed through the Ribo-Zero Magnetic Kit for Bacteria (Illumina, cat. no MRZMB126). Successful removal of ribosomal RNA from the samples was confirmed using a Bioanalyzer (Agilent Technologies, Santa Clara, CA, USA). RNA samples were prepared for sequencing using the Illumina TruSeq RNA prep kit and the resulting library was sequenced by the JP Sulzberger Columbia Genome Center (CUGC) on an Illumina HiSeq 2000 resulting in single-end reads of 100 bp, with an overall sequencing depth of ~30 million reads per sample (Table S2). Sequence data are deposited in the Sequence Read Archive, BIOPROJECT PRJNA374879.

Read mapping

Raw sequence data quality was visualized using FastQC and then cleaned and trimmed using Trimmomatic version 0.27 (single-end mode, 6-bp wide sliding window for quality below 20, minimum length of 25 bp). To evaluate the best read mapping approach, trimmed single-end reads from each of the samples were mapped to a combined *Trichodesmium* spp. genome, obtained by merging the *T. erythraeum* IMS101 genome (NCBI reference sequence: NC_008312) with the partial *Trichodesmium thiebautii* H94 genome (GenBank: LAMW00000000), resulting in an average of $6.5 \pm 7.3\%$ reads mapped. Reads were mapped using RNA-Seq by Expectation Maximization (RSEM) [57] with Bowtie2 [58] using default parameters. These low mapping rates were similar to other *Trichodesmium* field studies [59, 60], and likely reflect the presence of reads of heterotrophic bacterial epibionts, which are present at high concentrations in the colonies [35, 36], as well as the potential variability between the genomes of cultured *Trichodesmium* and field populations [59], as field populations encompass a diversity of *Trichodesmium* species [15, 16]. To target a *Trichodesmium* community more representative of the field, reads were mapped to *Trichodesmium*-identified genes from a custom *Trichodesmium* metagenome database assembled by Frischkorn et al. [34]. These field-specific *Trichodesmium* genes were previously generated and analyzed by Frischkorn et al. [34] from four *Trichodesmium*-identified genome bins in *Trichodesmium* colony metagenomes. Briefly, their approach utilized MaxBin 2.0 [61], Prodigal [62], and DIAMOND [63] against the NCBI nr database to generate genome bins, predict coding sequences and proteins, and identify taxonomy, respectively [34]. This custom *Trichodesmium* metagenome mapping approach yielded a 1.6 times higher mapping rate, on average, relative to the culture genome database (Table S2). With this custom *Trichodesmium* metagenome database mapping rates in the NPSG were not consistently lower than those from the NASG (Table S2). Moreover, the average percentage of reads mapping to sequences of each of the four *Trichodesmium* genomic bins defined in Frischkorn et al. [34] was similar in the NASG and NPSG (Fig. S2). Taken together, this suggests the custom *Trichodesmium* metagenome database [34] works well in both regions and, thus, was used for subsequent analyses.

Trichodesmium-only proteins predicted from the custom *Trichodesmium* metagenome database were clustered into 6710 orthologous groups (OGs), as described by Frischkorn et al. [34]. This yielded ~2.3 times more OGs than that derived from the *T. erythraeum* IMS101 genome (2982) as per Frischkorn et al. [34]. This disparity could stem from the presence of multiple species with different gene contents in *Trichodesmium* communities in the field and the fact that *T. erythraeum* IMS101 is not the dominant species in field populations [15, 16]. OGs were assigned putative annotations using the UniRef90 database [63] and the Kyoto Encyclopedia of Genes and Genomes (KEGG). Single functional annotations for entire OGs were determined by selecting the most abundant UniRef annotation for all
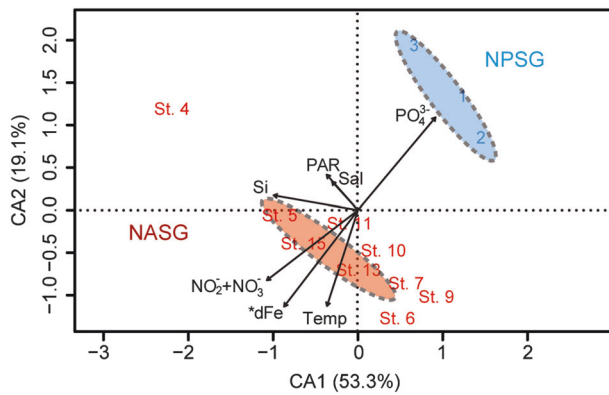
Fig. 1 Global transcriptional patterns of *Trichodesmium* communities in the North Atlantic subtropical gyre (NASG) and North Pacific subtropical gyre (NPSG). Correspondence analysis (CA) of the global metatranscriptome. CA ordinations of communities of the NASG and NPSG are shown, with significant ($p < 0.05$) environmental vectors fitted using the vegan function envfit [65]. Arrows indicate the direction of the (increasing) environmental gradient, and their lengths are proportional to their correlations with the ordination. 95% Confidence ellipses are indicated for each of the sample types by ocean basin. Discrete values for nitrite and nitrate ($NO_2^- + NO_3^-$), phosphate (PO_4^{3-}), dissolved iron (dFe) in the NPSG, the climatological average for dFe in the NASG, and temperature (Temp) used in this analysis are available in Table S1. *In the NASG, dFe was computed as the climatological average (Table S1)

proteins clustered into that group. KEGG annotations were used in some cases to refine the most abundant annotation. To most stringently categorize OGs identified as APs, annotations were refined with DIAMOND searches against representative proteins from COGs 3211 (PhoX), 1785 (PhoA), as well as the protein sequences of three previously identified putative APs in the IMS101 genome (PhoA: YP723031, PhoX: YP723360, and PhoX2: YP723924) [29], and as described in Frischkorn et al. [34]. For comparisons against the putative PhoA gene from *T. erythraeum* IMS101, OGs within the metagenomes were considered homologous to this protein if the blast e-value was $<1 \times 10^{-5}$ with a bit score >50 and contained UniRef blast homologs to the YP723031 gene from *T. erythraeum* IMS101 and other putative AP genes identified through KEGG or UniRef annotation [34]. For the AP OGs, we detected a single PhoX OG and two PhoA OGs termed PhoA1 and PhoA2 herein (Table S4). Additional nucleotide alignments were performed for all *Trichodesmium*-identified sequences in OGs classified as *pstS/sphX*, *phoA*, and *phoX* [34] along with representative sequences [29] of these genes from cultured isolates of Clade I, the dominant population in both regions (*T. thiebautii*, *T. spiralis*, and *T. tenue*). The MUSCLE application in Geneious (v.11) was used with default settings to run the alignment [64]. The culture *pstS/sphX*, *phoA*, and *phoX* sequences for *T. thiebautii*, *T. spiralis*, and *T. tenue* were obtained from GenBank under accession numbers FJ602760-602771. To conservatively approach the expression data, all reads mapping to gene variants within a given OG were pooled for subsequent analyses. OG counts were normalized to total mapped reads across each sample and only those OGs that had 1 read per million in at least three samples were included and considered detectable.
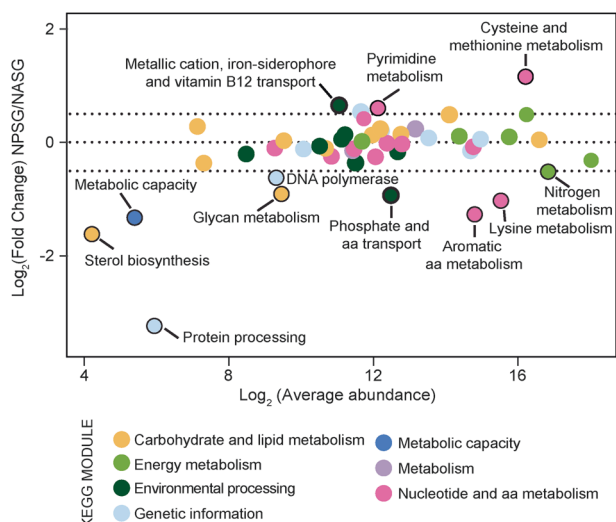
Differential expression analyses

Correspondence analyses (CAs) were conducted with the vegan package in R [65] using OG counts obtained from RSEM to examine differential OG expression patterns across samples. Vector fitting was done using the 'envfit' function in vegan with 9999 permutations to estimate the significance of the correlations between biogeochemical variables (Table S3) and the ordination by CA ($p < 0.05$), using the Benjamini–Hochberg (BH) algorithm [66] to control the false discovery rate (FDR). Statistical differences in global transcriptional patterns between NPSG and NASG communities were tested with a permutational multivariate analysis of variance (PERMANOVA) in R, using a Bray–Curtis dissimilarity matrix from the normalized OG table as the input for the 'adonis' function within the vegan package, with 9999 permutations [67, 68]. This analysis was repeated on the subset of OGs that were co-expressed in both the NPSG and the NASG. OG counts obtained from RSEM were also used to calculate differential expression with the package edgeR in R [69] for individual OGs in the NPSG compared with the NASG *Trichodesmium* sp. populations, treating each sample in the two regions as biological replicates. Default parameters were used to calculate dispersion of normalized counts for the replicates within each ocean basin so that these could be combined for further analyses [70]. Pairwise comparisons of combined counts (i.e., OG relative abundance) between ocean basins (NPSG versus NASG) were made with the exactTest function. FDR was controlled with the BH algorithm. FDR values <0.05 reflect statistically significant differences in OG relative abundance between NPSG and NASG communities (Table S4).

Results and discussion

Transcriptional patterns vary significantly between NASG and NPSG populations

Transcriptional patterns in *Trichodesmium* populations collected in the NASG and NPSG (Fig. S1) were examined to ascertain the geochemical drivers of *Trichodesmium* physiological ecology in these two ocean gyres. Of the 6710 *Trichodesmium* OGs in the custom metagenome database

Fig. 2 KEGG pathway enrichment in the North Atlantic subtropical gyre (NASG) and North Pacific subtropical gyre (NPSG). Differential expression of *Trichodesmium* communities from the NASG versus NPSG across KEGG pathways. Highlighted pathways are those with a Log_2 fold change > 0.5 (above or below the dashed line). Bold outline indicates KEGG modules associated with iron and phosphorus transport. Here, aa denotes amino acid

[34], a total of 4709 OGs were expressed by *Trichodesmium* communities in this study, of which 89% were common between the populations of the two oceans (Fig. S3). CA identified a significant difference (PERMANOVA: $F = 4.51$, $p = 0.013$) in global transcriptional pattern between NASG and NPSG *Trichodesmium* communities, with samples from each ocean basin segregating on the CA ordination plane (Fig. 1). Vector fitting showed that phosphate (PO_4^{3-}), and the climatological average for dFe, nitrate + nitrite ($NO_2^- + NO_3^-$), and temperature (Temp) significantly correlated ($p < 0.05$) with the global transcriptional pattern of these communities, whereas salinity and PAR did not (Table S3). Si, which is not used by *Trichodesmium* as a resource, did not significantly correlate with the differences in transcription. Samples from the NPSG were placed in the region of the ordination plane where relatively higher PO_4^{3-} concentrations were measured and samples from the NASG were placed in the region of the ordination plane where relatively higher dFe concentrations were recorded (Fig. 1). Similar patterns in the CA ordination, vector fitting, and PERMANOVA analysis ($F = 4.31$, $p = 0.021$) were observed when these analyses were run with the subset of OGs co-expressed in both the NASG and the NPSG (Fig. S4). Transcripts in *Trichodesmium* can turn over quickly in response to changes in geochemistry [29, 38] and are averaged here over both time and space, which would tend to minimize potential differences across ocean basins. Yet, the global transcriptional patterns associated with each basin were strikingly different

and linked with the basin-scale climatology and geochemistry, particularly for Fe and P.

Despite clear differences in the transcriptional patterns (Fig. 1), *Trichodesmium* populations were similar in the two regions. Samples in this study, with the exception of NPac_3, were concurrently collected with those of Rouco et al. [15, 16], confirming that Clade I dominated (94% and 92% in the NASG and NPSG, respectively) *Trichodesmium* field communities in both oceans. Additionally, the average percentage of reads mapping to sequences of each of the four *Trichodesmium* genomic bins defined in Frischkorn et al. [34] was consistent across regions, with bin 1 the most abundant across all samples, followed by bins 3, 2, and 9 (Fig. S2). Taken together, this evidence suggests that *Trichodesmium* populations were comparable between regions.

As a first approach to evaluate which pathways might drive the apparent differences in regional transcriptional patterns, OGs were grouped at the module level based on KEGG orthology (Fig. 2). Of the KEGG pathways with a Log_2 fold change > 0.5, the metallic cation and Fe-siderophore module, including Fe metabolic pathways, was enriched with a higher proportion of reads in the NPSG relative to the NASG (Fig. 2). In contrast, the module including phosphate transport pathways was enriched in the NASG relative to the NPSG (Fig. 2). Taken together, the observed transcriptional patterns empirically corroborate models that suggest the relative importance of Fe and P in controlling phytoplankton productivity in these two ocean gyres [13, 71] and reflect ecosystem-specific geochemical drivers of *Trichodesmium* physiology, suggesting that low Fe is a driver of transcriptional patterns in the NPSG, whereas low P shapes observed transcriptional patterns in the NASG.

Marker genes reflect ecosystem-specific traits that underpin *Trichodesmium* resource acquisition

The relative expression patterns of individual OGs were compared between the NPSG and NASG to examine specific functions outside of a KEGG framework (Table S4). NPSG and NASG *Trichodesmium* communities had significant differences in their transcriptional patterns at the OG level, with 30% of the OGs (1401 of 4709 OG) having significant (FDR < 0.05) differences (Table S4, Fig. S5). All of the individual OGs in this analysis were detectable in both the NASG and NPSG, indicating that significant differences in their expression across oceans were not derived from a complete lack of read mapping in a specific region. The relative abundance of the RotA OG, which has been used as a constitutively expressed marker gene [29], did not significantly differ between regions (Fig. S5). Interestingly, the relative abundances for OGs coding for the multi-subunit proteins of the metalloprotein, nitrogenase, involved
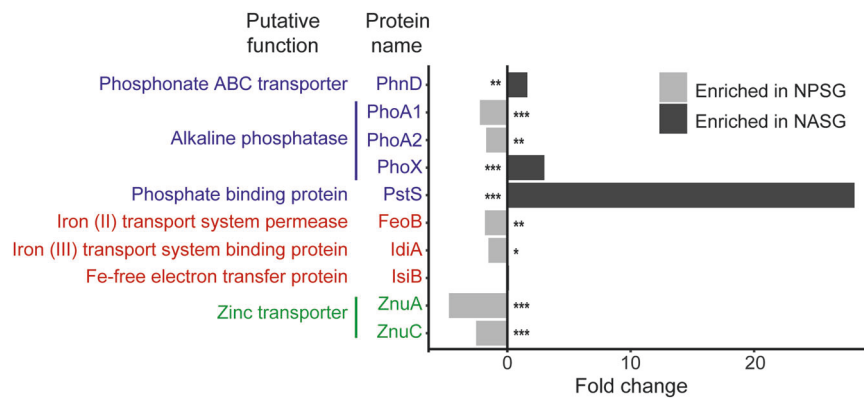
Fig. 3 Marker gene expression enrichment in *Trichodesmium* field communities. Fold change in expression of orthologous groups (OG) used as markers of P and Fe limitation and Zn uptake (highlighted in blue, red, and green, respectively) in North Atlantic subtropical gyre (NASG) and Nor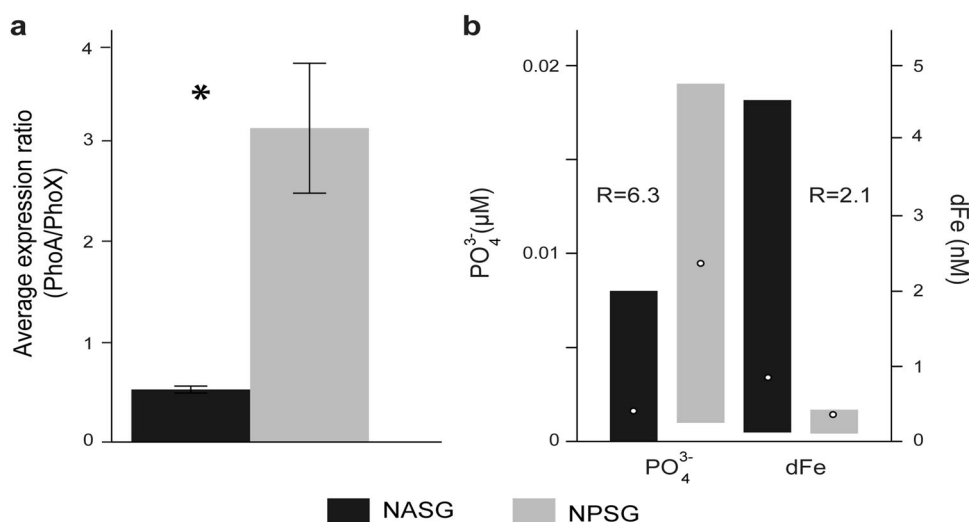th Pacific subtropical gyre (NPSG) communities. Genes listed for Fe and P were all experimentally validated as Fe or P-regulated in *Trichodesmium* [26, 29, 38–40, 44, 74]. Asterisks next to the horizontal bars indicate significance (***$p < 0.001$, **$p < 0.05$, *$p < 0.1$)

in N_2 fixation, including the Fe nitrogenase reductase protein (NifH) and the Fe-molybdenum dinitrogen reductases (NifD and NifK), were not statistically differentially abundant between NPSG and NASG populations (Fig. S5) where the range in N_2 fixation is overlapping but highly variable [5, 55]. Although *nif* expression can track with *Trichodesmium* N_2 fixation in culture, maximum *nif* expression generally occurs 2 h before the maximum nitrogenase activity [42, 72] and the nitrogenase enzyme can be regulated both at the transcriptional and post-transcriptional level [73]. As a result, the *nif* gene expression data derived here cannot be used to predict patterns in *Trichodesmium* N_2 fixation rates with confidence.

To further constrain the relationship between Fe, P, and *Trichodesmium* physiology, we tracked a number of OGs corresponding to proteins previously characterized as markers of P or Fe stress. There were three main OGs that contained proteins that have been used in expression studies as markers of Fe stress in *Trichodesmium* communities: IdiA, a Fe (III) transport system binding protein; FeoB, an Fe (III) transport system Fe permease; and IsiB or flavodoxin, an Fe-free electron transfer protein that can replace ferrodoxin during Fe stress [38, 40, 44, 74, 75]. The relative abundance of two of the three OGs, IdiA and FeoB, were significantly enriched in *Trichodesmium* populations in the NPSG relative to the NASG (Fig. 3, Fig. S5). With a read mapping approach, we cannot fully discount the potential presence of sequence variants for these gene targets in the NPSG that could bias the OG enrichment pattern. However, if sequence variants for IdiA and FeoB were abundant in the NPSG, they would reduce the relative NPSG enrichment signal, not increase it. As such, our observations for these OGs likely represent only minimum levels of NPSG enrichment. IsiB relative abundance was not significantly different between the NPSG and NASG (Fig. 3, Fig. S5).

Previous work identified IsiB expression consistent with Fe limitation in the South Pacific, but not the NASG using a quantitative reverse transcriptase-PCR (qRT-PCR) approach [44]. It is possible that IsiB expression is sensitive to the relative Fe and P geochemistry that differs between the NPSG and South Pacific, which could explain the lack of significant IsiB enrichment in the NPSG observed here. IsiB sequence variants have been detected in the South Pacific [44]. Such variants in primer sites are more of a concern for qRT-PCR than metatranscriptome profiling as applied here. However, we cannot exclude the possibility that NPSG sequence variants for IsiB may be influencing the OG enrichment pattern. Regardless, the enrichment of markers of Fe stress in the NPSG versus NASG populations (Fig. 3) appears to underpin the global transcriptional patterns (Fig. 1) and these data highlight the importance of Fe as a driver of *Trichodesmium* physiological ecology in the NPSG.

Similar to Fe, there are a number of P-scavenging OGs that have been identified as markers of P stress in *Trichodesmium* [26, 29, 39]. These include OGs involved in the acquisition of PO_4^{3-}, the preferred and most abundant source of P for marine bacteria [76, 77], as well as those involved in the acquisition of other P compounds, such as phosphonates and phosphoesters, which dominate the dissolved P pool in oligotrophic regions [76, 78]. The relative abundance of representative OGs for a number of P stress markers were statistically higher in *Trichodesmium* populations in the NASG compared with those of NPSG (Fig. 3, Fig. S5). These P stress markers included the high-affinity PO_4^{3-} binding protein, PstS [29], the high-affinity phosphonate binding protein, PhnD [26], and the AP PhoX, which hydrolyzes P from ester-bond P compounds [29]. It is unlikely that the enrichment patterns observed for the expression of these P-related OGs are an artifact of read

Fig. 4 PhoA-/PhoX-type alkaline phosphatase average expression ratios in *Trichodesmium* populations of the North Atlantic subtropical gyre (NASG) and North Pacific subtropical gyre (NPSG) with relevant geochemical climatology. **a** PhoA-/PhoX-type alkaline phosphatase average expression ratios. PhoA here represents the total PhoA signal where normalized values for PhoA1 and PhoA2 OGs were summed. Error bars indicate SEM. Asterisk (*) indicates significance ($p = 0.009$) using the Wilcoxon test in R. **b** Climatological averages of phosphate (PO_4^{3-}) and dissolved iron (dFe) concentrations in the North Atlantic subtropical gyre (NASG) and North Pacific subtropical gyre (NPSG). The analysis includes surface (~5 m) climatological data from the same region (10°N-30°N and 70°W-25°W in NASG and 20°N-30°N and 150°W-160°W in NPSG) as this study collected throughout different years and seasons. White points indicate the average, and the bars correspond to the maximum and minimum number recorded in the region. R represents the ratio of the higher to the lower nutrient concentration. Notably, PO_4^{3-} concentrations are ~6 times higher in the NPSG than the NASG. In contrast, dFe concentrations are ~2 times higher in the NASG than the NPSG. PO_4^{3-} data from [28, 44, 46–48, 52–55], and this study. dFe data from [44, 45, 49–51, 54, 56], and this study

mapping. As highlighted above, *Trichodesmium* diversity is likely similar across regions, and there is a high degree of nucleotide identity for the *phoX*, *phoA*, and *pstS* genes between species and even different clades of *Trichodesmium* [29]. The maximum % nucleotide identity between sequences within the PstS, PhoA, and PhoX OGs and cultured isolates was >97% (Table S5). This indicates that, although the OG as a whole might contain some slightly divergent gene sequences, in each OG there is at least one sequence that is nearly identical to a gene in *T. thiebautii*, *T. spiralis*, and *T. tenue*. Thus, it is unlikely that expression differences observed for these OGs are consistently an artifact of read mapping. Again, these expression patterns corroborate the global transcriptional patterns (Fig. 1), as well as model results [13] and previous non-species-specific field experiments focused on AP activity and phosphate uptake [51,79–81]. Taken together, these data highlight the importance of P as a driver of *Trichodesmium* physiological ecology in the NASG.

Trichodesmium switches the relative transcript abundance of Fe- and Zn-requiring metalloenzymes for DOP hydrolysis

The PhoX-type AP is more common than the PhoA-type in marine bacteria [30]. Both metalloenzymes are present in the *T. erythraeum*, *T. tenue*, *T. spiralis*, and *T. theibautii* genomes [29, 30], and detected in *Trichodesmium* meta-genomes [34]. In the metatranscriptome data here, the relative abundance of the PhoX-type AP was significantly higher in the NASG compared with the NPSG and the relative abundance of the PhoA-type AP was significantly higher in the NPSG (Fig. 3, Fig. S5). Both APs are known to be upregulated under conditions of P stress in *Trichodesmium* laboratory cultures, although patterns in PhoX transcription suggested that it was the dominant AP under Fe replete culture conditions [29]. The two APs have different substrate specificities and metal co-factors. PhoX can hydrolyze both phosphomono- and di-esters and contains calcium (Ca) and Fe as co-factors [31, 32]. By contrast, PhoA is typically specific to phosphomonoesters and contains zinc (Zn) and magnesium (Mg) co-factors [82, 83]. The patterns of expression of these two OGs across the two basins correlated with patterns of Fe and P biogeochemistry, with a higher PhoA/PhoX expression ratio in the NPSG where the average Fe is lower than in the NASG (Fig. 4). These data suggest that *Trichodesmium* may alter its PhoA/PhoX ratio to minimize its Fe requirement in low Fe regions like the NPSG. In addition, the expression signals of OGs involved in Zn homeostasis and uptake, such as ZnuA and ZnuC [84], were significantly enriched in *Trichodesmium* communities in the NPSG (Fig. 3), suggesting that there

may be an increased Zn demand associated with the expression of the PhoA AP, which requires Zn. Tradeoffs associated with this strategy would be a higher Zn quota for populations in the NPSG and reduced bioavailability of some DOP substrates. Although there is limited information on Zn concentrations for these regions, average surface values are higher (~0.24 nM) in the NASG than in the NPSG (~0.07 nM) [47, 85–87]. It is not clear why *Trichodesmium* would favor a Zn-requiring enzyme, PhoA, in a low Zn environment like the NPSG, but the coincident enrichment of Zn transport functions suggests an increase in Zn demand, and that the Zn may be from the environment and not just recycled internally. Regardless, these data suggest a switch between Fe- and Zn-rich metalloenzymes for DOP hydrolysis consistent with climatological patterns in Fe geochemistry of the NASG and NPSG. Fe and Zn limitation of community AP activity had been previously shown using data from bulk AP activity assays, where AP increased after Fe or Zn additions in different ocean regions [33, 80]. This is most likely driven by shifts in the PhoA and PhoX expression ratios, as observed here for *Trichodesmium*, but not all marine bacteria carry both PhoX and PhoA like *Trichodesmium*. The apparent metalloenzyme switching observed here might underpin *Trichodesmium* fitness across variable oligotrophic ocean ecosystems by allowing cells to modulate trace metal requirements and mitigate Fe limitation of DOP metabolism while maintaining other high Fe-requiring processes such as N_2 fixation and photosynthesis in low Fe regions [14].

Conclusion

Taken together, the transcriptional patterns reflect ecosystem-specific geochemical drivers of *Trichodesmium* physiological ecology and empirically validate geochemical models that predict the importance of P in the control of *Trichodesmium* growth and N_2 fixation in the NASG and Fe as a driver in the NPSG [13]. The expression of phosphate and phosphonate transport genes and AP genes highlights the importance of DOP metabolism in both ecosystems and the tradeoffs that this organism uses to maintain N_2 fixation under different geochemical conditions. The findings also suggest that trace elements, such as Zn, should also be included in models given their influence on the activity of APs. *Trichodesmium* is predicted to increase growth and N_2 fixation with elevated CO_2 in the future ocean [88, 89]. Modeling the traits and tradeoffs observed here in the context of future ocean conditions will help predict concomitant impacts on C and N cycling and their control on marine primary production.

Compliance with ethical standards

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

References

1. Marconi D, Sigman DM, Casciotti KL, Campbell EC, Alexandra Weigand M, Fawcett SE, et al. Tropical dominance of N_2 fixation in the North Atlantic Ocean. Glob Biogeochem Cycles. 2017;31:1608–23.
2. Bergman B, Sandh G, Lin S, Larsson J, Carpenter EJ. *Trichodesmium* - a widespread marine cyanobacterium with unusual nitrogen fixation properties. FEMS Microbiol Rev. 2013;37:286–302.
3. Arrigo KR. Marine microorganisms and global nutrient cycles. Nature. 2005;437:343–8.
4. Mulholland MR. The fate of nitrogen fixed by diazotrophs in the ocean. Biogeosciences. 2007;4:37–51.
5. Sohm JA, Webb EA, Capone DG. Emerging patterns of marine nitrogen fixation. Nat Rev Microbiol. 2011;9:499–508.
6. Tyrrell T. The relative influences of nitrogen and phosohorus on oceanic primary production. Nature. 1999;400:525–31.
7. Böttjer D, Dore JE, Karl DM, Letelier RM, Mahaffey C, Wilson ST, et al. Temporal variability of nitrogen fixation and particulate nitrogen export at Station ALOHA. Limnol Oceanogr. 2017;62:200–16.
8. Capone DG, Burns JA, Montoya JP, Subramaniam A, Mahaffey C, Gunderson T, et al. Nitrogen fixation by *Trichodesmium* spp.: an important source of new nitrogen to the tropical and subtropical North Atlantic Ocean. Glob Biogeochem Cycles. 2005;19:GB2024.
9. Karl D, Letelier R, Tupas L, Dore J, Christian J, Hebel D. The role of nitrogen fixation in biogeochemical cycling in the subtropical North Pacific Ocean. Nature. 1997;388:533–8.
10. Van Mooy BAS, Krupke A, Dyhrman ST, Fredricks HF, Frischkorn KR, Ossolinski JE, et al. Major role of planktonic phosphate reduction in the marine phosphorus redox cycle. Science. 2015;348:783–5.
11. Karl DM, Beversdorf L, Björkman KM, Church MJ, Martinez A, Delong EF. Aerobic production of methane in the sea. Nat Geosci. 2008;1:473–8.
12. Repeta DJ, Ferrón S, Sosa OA, Johnson CG, Repeta LD, Acker M, et al. Marine methane paradox explained by bacterial degradation of dissolved organic matter. Nat Geosci. 2016;9:884–7.
13. Moore JK, Doney SC, Lindsay K. Upper ocean ecosystem dynamics and iron cycling in a global three-dimensional model. Glob Biogeochem Cycles. 2004;18:1–21.

14. Berman-Frank I, Cullen J, Shaked Y, Sherrell R, Falkowski PG. Iron availability, cellular iron quotas, and nitrogen fixation in *Trichodesmium*. Limnol Oceanogr. 2001;46:1249–60.

15. Rouco M, Haley ST, Alexander H, Wilson ST, Karl DM, Dyhrman ST. Variable depth distribution of *Trichodesmium* clades in the North Pacific Ocean. Environ Microbiol Rep. 2016a;8:1058–66.

16. Rouco M, Joy-Warren H, McGillicuddy DJ, Waterbury JB, Dyhrman ST. *Trichodesmium* sp. clade distributions in the western North Atlantic Ocean. Limnol Oceanogr. 2014;59:1899–909.

17. Fitzsimmons JN, Hayes CT, Al-Subiai SN, Zhang R, Morton P, Weisend R, et al. Daily to decadal variability of size-fractionated iron and iron-binding ligands at the Hawaii Ocean Time-series Station ALOHA. Geochim Cosmochim Acta. 2015;171:303–24.

18. Saito MA, Goepfert TJ, Ritt JT. Some thoughts on the concept of colimitation: three definitions and the importance of bioavailability. Limnol Oceanogr. 2008;53:276–90.

19. Landolfi A, Koeve W, Dietze H, Kähler P, Oschlies A. A new perspective on environmental controls of marine nitrogen fixation. Geophys Res Lett. 2015;42:4482–9.

20. Snow JT, Schlosser C, Woodward EMS, Mills MM, Achterberg EP, Mahaffey C, et al. Environmental controls on the biogeography of diazotrophy and *Trichodesmium* in the Atlantic Ocean. Glob Biogeochem Cycles. 2015;29:865–84.

21. Ward BA, Dutkiewicz S, Mark Moore C, Follows MJ. Iron, phosphorus, and nitrogen supply ratios define the biogeography of nitrogen fixation. Limnol Oceanogr. 2013;58:2059–75.

22. Van Mooy BAS, Fredricks HF, Pedler BE, Dyhrman ST, Karl DM, Koblízek M, et al. Phytoplankton in the ocean use non-phosphorus lipids in response to phosphorus scarcity. Nature. 2009;458:69–72.

23. Morrissey J, Bowler C. Iron utilization in marine cyanobacteria and eukaryotic algae. Front Microbiol. 2012;3:1–13.

24. Roe KL, Barbeau K. Uptake mechanisms for inorganic iron and ferric citrate in *Trichodesmium erythraeum* IMS101. Metallomics. 2014;6:2042–51.

25. Polyviou D, Hitchcock A, Baylay AJ, Moore CM, Bibby TS. Phosphite utilization by the globally important marine diazotroph *Trichodesmium*. Environ Microbiol Rep. 2015;7:824–30.

26. Dyhrman ST, Chappell PD, Haley ST, Moffett JW, Orchard ED, Waterbury JB, et al. Phosphonate utilization by the globally important marine diazotroph *Trichodesmium*. Nature. 2006;439:68–71.

27. Dyhrman ST, Webb EA, Anderson DM, Moffett JW, Waterbury JB. Cell-specific detection of phosphorus stress in *Trichodesmium* from the Western North Atlantic. Limnol Oceanogr. 2002;47:1832–6.

28. Orchard ED, Ammerman JW, Lomas MW, Dyhrman ST. Dissolved inorganic and organic phosphorus uptake in *Trichodesmium* and the microbial community: the importance of phosphorus ester in the Sargasso Sea. Limnol Oceanogr. 2010;55:1390–9.

29. Orchard ED, Webb EA, Dyhrman ST. Molecular analysis of the phosphorus starvation response in *Trichodesmium* spp. Environ Microbiol. 2009;11:2400–11.

30. Sebastian M, Ammerman JW. The alkaline phosphatase PhoX is more widely distributed in marine bacteria than the classical PhoA. ISME J. 2009;3:563–72.

31. Wu J-R, Shien J-H, Shieh HK, Hu C-C, Gong S-R, Chen L-Y, et al. Cloning of the gene and characterization of the enzymatic properties of the monomeric alkaline phosphatase (PhoX) from *Pasteurella multocida* strain X-73. FEMS Microbiol Lett. 2007;267:113–20.

32. Yong SC, Roversi P, Lillington J, Rodriguez F, Krehenbrink M, Zeldin OB, et al. A complex iron-calcium cofactor catalyzing phosphotransfer chemistry. Science. 2014;345:1170–3.

33. Browning TJ, Achterberg EP, Yong JC, Rapp I, Utermann C, Engel A, et al. Iron limitation of microbial phosphorus acquisition in the tropical North Atlantic. Nat Commun. 2017;8:15465.

34. Frischkorn KR, Rouco M, Van Mooy BAS, Dyhrman ST. Epibionts dominate metabolic functional potential of *Trichodesmium* colonies from the oligotrophic ocean. ISME J. 2017;11:2090–101.

35. Hmelo LR, Van Mooy BAS, Mincer T. Characterization of bacterial epibionts on the cyanobacterium *Trichodesmium*. Aquat Microb Ecol. 2012;67:1–14.

36. Rouco M, Haley ST, Dyhrman ST. Microbial diversity within the *Trichodesmium* holobiont. Environ Microbiol. 2016b;18:5151–60.

37. Van Mooy BA, Hmelo LR, Sofen LE, Campagna SR, May AL, Dyhrman ST, et al. Quorum sensing control of phosphorus acquisition in *Trichodesmium* consortia. ISME J. 2012;6:422–9.

38. Chappell PD, Webb EA. A molecular assessment of the iron stress response in the two phylogenetic clades of *Trichodesmium*. Environ Microbiol. 2010;12:13–27.

39. Martínez A, Osburne MS, Sharma AK, DeLong EF, Chisholm SW. Phosphite utilization by the marine picocyanobacterium *Prochlorococcus* MIT9301. Environ Microbiol. 2012;14:1363–77.

40. Webb E, Moffett J, Waterbury J. Iron stress in open-ocean cyanobacteria (*Synechococcus*, *Trichodesmium*, and *Crocosphaera* spp.): identification of the IdiA protein. Appl Environ Microbiol. 2001;67:5444–52.

41. Saito MA, McIlvin MR, Moran DM, Goepfert TJ, DiTullio GR, Post AF, et al. Multiple nutrient stresses at intersecting Pacific Ocean biomes detected by protein biomarkers. Science. 2014;345:1173–7.

42. El-Shehawy R, Lugomela C, Ernst A, Bergman B. Diurnal expression of hetR and diazocyte development in the filamentous non-heterocystous cyanobacterium *Trichodesmium erythraeum*. Microbiology. 2003;149:1139–46.

43. Rimmelin P, Moutin T. Re-examination of the MAGIC method to determine low orthophosphate concentration in seawater. Anal Chim Acta. 2005;548:174–82.

44. Chappell PD, Moffett JW, Hynes AM, Webb EA. Molecular evidence of iron limitation and availability in the global diazotroph *Trichodesmium*. ISME J. 2012;6:1728–39.

45. Bruland KW, Orians KJ, Cowen JP. Reactive trace metals in the stratified central North Pacific. Geochim Cosmochim Acta. 1994;58:3171–82.

46. Diaz JM, Björkman KM, Haley ST, Ingall ED, Karl DM, Longo AF, et al. Polyphosphate dynamics at Station ALOHA, North Pacific subtropical gyre. Limnol Oceanogr. 2016;61:227–39.

47. Jakuba RW, Moffett JW, Dyhrman ST. Evidence for the linked biogeochemical cycling of zinc, cobalt, and phosphorus in the western North Atltic Ocean. Glob Biogeochem Cycles. 2008;22:GB4012.

48. Reinthaler T, Sintes E, Herndl GJ. Dissolved organic matter and bacterial production and respiration in the sea-surface microlayer of the open Atlantic and the western Mediterranean Sea. Limnol Oceanogr. 2008;53:122–36.

49. Bergquist BA, Boyle EA. Dissolved iron in the tropical and subtropical Atlantic Ocean. Glob Biogeochem Cycles. 2006;20:GB1015.

50. Rue EL, Bruland KW. Complexation of iron(III) by natural organic ligands in the Central North Pacific as determined by a new competitive ligand equilibration/adsorptive cathodic stripping voltammetric method. Mar Chem. 1995;50:117–38.

51. Sañudo-Wilhelmy SA, Kustka AB, Gobler CJ, Hutchins DA, Yang M, Lwiza K, et al. Phosphorus limitation of nitrogen fixation by *Trichodesmium* in the central Atlantic Ocean. Nature. 2001;411:66–9.

52. Sohm J, Capone D. Phosphorus dynamics of the tropical and subtropical north Atlantic: *Trichodesmium* spp. versus bulk plankton. Mar Ecol Prog Ser. 2006;317:21–8.

53. Wilson ST, Barone B, Ascani F, Bidigare RR, Church MJ, Del Valle DA, et al. Short-term variability in euphotic zone biogeochemistry and primary productivity at Station ALOHA: a case study of summer 2012. Glob Biogeochem Cycles. 2015;29:1145–64.

54. Wu J, Boyle E. Iron in the Sargasso Sea: implications for the processes controlling dissolved Fe distribution in the ocean. Glob Biogeochem Cycles. 2002;16:1–8.

55. Wu J, Sunda W, Boyle EA, Karl DM. Phosphate depletion in the western North Atlantic. Ocean Sci. 2000;289:759–62.

56. Brown MT, Landing WM, Measures CI. Dissolved and particulate Fe in the western and central North Pacific: results from the 2002 IOC cruise. Geochem, Geophys Geosystems. 2005;6:Q10001.

57. Li B, Dewey CN. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. BMC Bioinforma. 2011;12:323.

58. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. Nat Meth. 2012;9:357–9.

59. Hilton JA, Satinsky BM, Doherty M, Zielinski B, Zehr JP. Metatranscriptomics of N$_2$-fixing cyanobacteria in the Amazon River plume. ISME J. 2015;9:1557–69.

60. Spungin D, Pfreundt U, Berthelot H, Bonnet S, AlRoumi D, Natale F, et al. Mechanisms of *Trichodesmium* bloom demise within the New Caledonia Lagoon during the VAHINE mesocosm experiment. Biogeosci Discuss. 2016;13:1–44.

61. Wu YW, Simmons BA, Singer SW. MaxBin 2.0: an automated binning algorithm to recover genomes from multiple metagenomic datasets. Bioinformatics. 2015;32:605–7.

62. Hyatt D, Chen G-L, Locascio PF, Land ML, Larimer FW, Hauser LJ. Prodigal: prokaryotic gene recognition and translation initiation site identification. BMC Bioinforma. 2010;11:119.

63. Buchfink B, Xie C, Huson DH. Fast and sensitive protein alignment using DIAMOND. Nat Methods. 2015;12:59–60.

64. Edgar RC. MUSCLE: a multiple sequence alignment method with reduced time and space complexity. BMC Bioinforma. 2004;5:1–19.

65. Oksanen J. Multivariate analysis of ecological communities in R: vegan tutorial. 2015. URL: http://www.mooreecology.com/uploads/2/4/2/1/24213970/vegantutor.pdf.

66. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. J R Stat Soc Ser B. 1995;57:289–300.

67. Anderson MJ. A new method for non-parametric multivariate analysis of variance. Austral Ecol. 2001;26:32–46.

68. Zapala MA, Schork NJ. Multivariate regression analysis of distance matrices for testing associations between gene expression patterns and related variables. Proc Natl Acad Sci USA. 2006;103:19430–5.

69. Robinson M, Oshlack A. A scaling normalization method for differential expression analysis of RNA-seq data. Genome Biol. 2010;11:R25.

70. Robinson MD, Smyth GK. Moderated statistical tests for assessing differences in tag abundance. Bioinformatics. 2007;23:2881–7.

71. Krishnamurthy A, Moore JK, Mahowald N, Luo C, Zender CS. Impacts of atmospheric nutrient inputs on marine biogeochemistry. J Geophys Res. 2010;115:1–13.

72. Chen YB, Chen YB, Dominic B, Dominic B, Mellon MT, Mellon MT, et al. Circadian rhythm of nitrogenase gene expression in the diazotrophic filamentous nonheterocystous cyanobacterium *Trichodesmium* sp strain IMS101. J Bacteriol. 1998;180:3598–605.

73. Wyman M, Zehr JP, Capone DG. Temporal variability in nitrogenase gene expression in natural populations of the marine cyanobacterium *Trichodesmium thiebautii*. Appl Environ Microbiol. 1996;62:1073–5.

74. Shi T, Sun Y, Falkowski PG. Effects of iron limitation on the expression of metabolic genes in the marine cyanobacterium *Trichodesmium erythraeum* IMS101. Environ Microbiol. 2007;9:2945–56.

75. Webb E, Jakuba R, Moffett J, Dyhrman S. Molecular assessment of phosphorus and iron physiology in *Trichodesmium* populations from the western Central and western South Atlantic. Limnol Oceanogr. 2007;52:2221–32.

76. Karl DM. Microbially mediated transformations of phosphorus in the sea: new views of an old cycle. Ann Rev Mar Sci. 2014;6:279–337.

77. Vershinina O, Znamenskaya L. The Pho regulons of bacteria. Microbiology. 2002;71:581–95.

78. Young CL, Ingall ED. Marine dissolved organic phosphorus composition: insights from samples recovered using combined electrodialysis/reverse osmosis. Aquat Geochem. 2010;16:563–74.

79. Hynes AM, Chappell PD, Dyhrman ST, Doney SC, Webb EA. Cross-basin comparison of phosphorus stress and nitrogen fixation in *Trichodesmium*. Limnol Oceanogr. 2009;54:1438–48.

80. Mahaffey C, Reynolds S, Davis CE, Lohan MC, Lomas MW. Alkaline phosphatase activity in the subtropical ocean: insights from nutrient, dust and trace metal addition experiments. Front Mar Sci. 2014;1:1–13.

81. Sohm JA, Mahaffey C, Capone DG. Assessment of relative phosphorus limitation of *Trichodesmium* spp. in the North Pacific, North Atlantic, and the north coast of Australia. Limnol Oceanogr. 2008;53:2495–502.

82. Kim EE, Wyckoff HW. Reaction mechanism of alkaline phosphatase based on crystal structures. J Mol Biol. 1991;218:449–64.

83. Sowadski JM, Handschumacher MD, Krishna Murthy HM, Foster BA, Wyckoff HW. Refined structure of alkaline phosphatase from *Escherichia coli* at 2.8 A resolution. J Mol Biol. 1985;186:417–33.

84. Barnett JP, Millard A, Ksibe AZ, Scanlan DJ, Schmid R, Blindauer CA. Mining genomes of marine cyanobacteria for elements of zinc homeostasis. Front Microbiol. 2012;3:1–21.

85. Bruland KW. Oceanographic distributions of cadmium, zinc, nickel, and copper in the North Pacific. Earth Planet Sci Lett. 1980;47:176–198.

86. Kim T, Obata H, Kondo Y, Ogawa H, Gamo T. Distribution and speciation of dissolved zinc in the western North Pacific and its adjacent seas. Mar Chem. 2015;173:330–41.

87. Lohan MC, Statham PJ, Crawford DW. Total dissolved zinc in the upper water column of the subarctic North East Pacific. Deep Res Part II Top Stud Oceanogr. 2002;49:5793–808.

88. Hutchins DA, Fu F-X, Zhang Y, Warner ME, Feng Y, Portune K, et al. CO$_2$ control of *Trichodesmium* N$_2$ fixation, photosynthesis, growth rates, and elemental ratios: implications for past, present, and future ocean biogeochemistry. Limnol Oceanogr. 2007;52:1293–304.

89. Hutchins DA, Walworth NG, Webb EA, Saito MA, Moran D, McIlvin MR, et al. Irreversibly increased nitrogen fixation in *Trichodesmium* experimentally adapted to elevated carbon dioxide. Nat Commun. 2015;6:8155.