Point by point responses follow, with relevant changes noted in red.
A marked up manuscript appears at the end of the document.
Changes have been made to the text, Figures 1, 2, 3, and 4, and to the supplementary text

Reviewer #2 - point by point response

General overview

The authors present a metaproteomic study of field-collected Trichodesmium colonies, focused on phosphate and iron stress markers, and complement that study with a membrane crowding model, which I think is a nice approach to try and understand the observed co-limitation patterns for iron and phosphate. The study is comprehensive in that samples from multiple cruises and years are used; with all but one station (HOT) located in the Atlantic Ocean. Increasing the knowledge of nutrient limitation in natural Trichodesmium populations is certainly of interest, given that it seems to be connected to aggregation of Trichodesmium in some way, either directly or through a general C1 stress response. While the study as such is valuable and should be published, I have a few major remarks that I think should be addressed before it is ready.

We thank the reviewer for their thoughtful comments on our manuscript and discuss changes below in red. Updates to the text are also provided and changes are highlighted in green.

Major Remarks

1. The whole conclusion of co-occurring phosphate and iron stress relies on the assumption that protein abundances of IdiA and SphX are good proxies for iron or phosphate limitation, respectively. The authors do cite the relevant literature that showed upregulation of the respective markers under the corresponding nutrient stress. What I am missing is information on which fold-changes in protein abundance were measured in the cited studies under the respective nutrient limiting conditions. For example, what are the base levels of IdiA and SphX protein in the cell? If there is three times more SphX than IdiA, such as in Fig 3 for some of the Tricolim samples, does that really indicate co-limitation, or does that just reflect the base level of IdiA? For example, Snow et al, 2015 (Fig. 4) only report a two-fold change for IdiA from ~100 fmol/ug to ~200 fmol/ug under iron stress. I suggest presenting the evidence for IdiA and SphX being markers for the respective stresses clearly in a table, including the type of experiment (culture or field), absolute quantifications if stated, and fold-changes measured. I also suggest then being a little more careful in wording throughout the paper, and differentiating better how the results from different stations could be interpreted.

The reviewer asks us to address how our results compare with previous reports of IdiA and SphX abundance during nutrient limitation. Per the reviewer, we provide new supplementary tables (Table S8 and S9) containing fold-change values from the literature for IdiA and SphX as during nutrient

limitation. The biomarkers increased in abundance during nutrient depletion, however the magnitude of the response varied. This is likely due to experimental differences such as the analytical methodology (Western blots versus LC-MS), nutrient concentrations or growth rate of the culture being examined. For instance, IdiA responded less strongly to iron stress in the Walworth et al., 2016 experiments compared to the Webb et al., 2001 and Snow et al., 2015 experiments, and we hypothesize this is because the Fe-depleted condition had 10 nM added iron as opposed to 0 nM iron in Webb et al. 2001 and Snow et al. 2015. We agree with the reviewer that applying a quantitative framework to this data would be valuable once the necessary data becomes available and now note this in the text.

Based on the consistent responsiveness of IdiA and SphX to nutrient limitation in the laboratory, we concluded that the high relative abundance of these biomarkers was indicative of nutrient stress. The reasoning is that the cells were clearly devoting a large fraction of their proteome to Fe and P uptake, likely at the expense of other nutrient uptake systems such as for organic nitrogen as is discussed later in the text. The reviewer's point that there may be basal expression of IdiA and SphX in replete cells is well taken. Because we report relative abundance data, we cannot directly compare our results to those of other researchers who took different quantitative approaches. However, the possibility of basal expression should be clarified in the text. Carefully calibrated datasets relating IdiA and SphX protein abundance to nutrient limited growth rates, while outside the scope of this paper, would be valuable in facilitating interpretation of this data. Based on the reviewer's comments we have updated Section 3.2 below to clarify the assumptions and caveats in our interpretation, and to highlight the need for calibrated biomarker studies.

3.2 *Trichodesmium* is simultaneously iron and phosphate stressed throughout its habitat

A surprising emergent observation from the *Trichodesmium* metaproteomes was the co-occurrence of the iron (IdiA) and phosphate (SphX) stress biomarkers across the samples. The ubiquitous and highly abundant presence of these proteins relative to total protein implied that co-stress may be the norm rather than the exception for *Trichodesmium* colonies in the field, particularly in the North Atlantic. Even though low-level basal expression of IdiA and SphX has been observed, it was clear that the colonies were devoting a large fraction of their cellular resources to Fe and P uptake, respectively (see Tables S8 and S9) (Webb et al., 2001, Webb et al., 2007, Chappell et al., 2010, Orchard et al., 2010, Snow et al., 2015, Walworth et al., 2016, Frischkorn et al., 2019). This, combined with the responsiveness of IdiA and SphX to nutrient availability in *Trichodesmium* filaments in the laboratory, indicated that co-stress was occurring.

Interestingly, biomarker abundance was not necessarily associated with nutrient concentrations in the surface ocean, suggesting that the colonies were experiencing stress despite variation in nutrient availability (Figure 3 C-D). SphX abundance varied up to 7.5 fold and were negatively associated with dissolved phosphate concentrations, though analytical differences across the field expeditions may have forced this relationship (Figure S2). Oceanographically, SphX was most abundant in the P-deplete, summer-stratified North Atlantic gyre (JC150 expedition) compared with winter waters near the Amazon river plume (Tricolim expedition) or at station ALOHA, where phosphate concentrations were greater (Hynes et al., 2009; Sañudo-Wilhelmy et al., 2001; Wu et al., 2000). IdiA varied up to 8 fold but there was no observable relationship with

dFe concentrations at the surface. Instead, IdiA may be responsive to other factors such as the varying iron requirements of the populations/species examined. It should be highlighted that in this study only *Trichodesmium* colonies were considered, so factors such as colony size may affect iron availability and biomarker expression. Additionally, because the surface ocean iron inventory is low, transient inputs such as from the Sahara desert can dramatically impact iron availability on short time scales, and the time scale of these inputs relative to changes in biomarker abundance is not well understood (Kunde et al., 2019). Carefully calibrated datasets relating IdiA and SphX abundance to nutrient-limited growth rates of *Trichodesmium* in both the filamentous and colonial forms would facilitate further interpretation of this data.

Table S8. Literature values for IdiA as a biomarker of Fe stress

Study	**Fe addition replete**	**Fe addition deplete**	**IdiA fold change**
Webb 2001	50nM	0nM	1.54
Walworth 2016	250nM	10nM	1.07 at 380 pCO2, 1.31 at 780 pCO2
Snow 2015	120nM	0nM	2.38

*Note that while the cells in Walworth, et al., 2016 were clearly Fe-limited according to growth rate measurements, they had access to 10nM total iron as opposed to 0nM total iron for the other experiments. This may explain the discrepancy in the IdiA fold change values.

Table S9. Literature values for SphX as a biomarker of P stress

Study	**P addition replete**	**P addition deplete**	**SphX fold change**
Walworth 2016	0.25uM	10uM	2.64 at 380 pCO2, 3.10 at 780 pCO2
Frischkorn 2019	0uM	50uM	2.62

2. Figure 3, and the corresponding Figure S2 are nice and the basis for some important claims being made in part 3.2. of this manuscript. However, these claims should be supported with the necessary

statistics, and it would help if Fig S2 was not in the Supplement, but presented together. For example, in line 210 ff, the authors claim that a) "Biomarkers for iron (IdiA) and phosphate (SphX) stress were highly abundant and positively associated with surface Fe or P concentrations" and b) "IdiA varied up to 8 fold, and increased moving West to East across the JC150 transect, consistent with an observed decrease in dFe concentrations" . For a) I think the authors mean "negatively", not "positively", correlated. And while I believe this correlation for SphX, it is not obvious for IdiA. For b) I cannot see increasing protein abundance from west Please prove this statistically before claiming it.

First, we thank the reviewer for correcting our mistake in line 210 – we did indeed mean to write "negatively." For clarity, we have moved the panels in Figure S2 alongside Figure 3 in the main text. Figure S2 now provides scatter plots of IdiA and SphX versus dFe and phosphate concentrations in the surface ocean (see updated figures section at the end of this document). There was a statistically observable relationship between SphX and dissolved phosphate, however the relationship may be forced by the different analytical approaches used on the JC150 versus Tricolim expeditions as is noted in the text below. By contrast there was no statistically observable relationship between IdiA and dFe, even though laboratory experiments clearly indicated that IdiA was a good biomarker of Fe stress in *Trichodesmium*. There are many factors that could influence this association, particularly iron speciation, changes in iron quotas, and factors such as colony size, which are not controlled in the field. We thank the reviewer for calling this point to our attention as it provides an opportunity to discuss these points in the updated Section 3.2 (see Major Remarks #1).

3. Given that only 1 sampling station is NOT in the Atlantic, please remove all claims that generalize the findings, e.g. "co-stress is the norm rather than the exception" (l. 18) → add "in the Atlantic", if wanting to keep this. Or in line 60f: "simultaneously Fe and P stressed throughout the worlds oceans" – this statement cannot be made with just one station outside the Atlantic.

Change accepted.

Specific Comments

Abstract

The abstract is missing some specificity.

line 19: nitrogenase was most abundant – compared to what? Please rephrase: more abundant than under . . line 22: is confronted by the biophysical limits – when? Under which conditions is it confronted by this? line 24f: be more specific. The last sentence is true for any microbe.

We agree with these points and have updated the text accordingly. On line 24f, we wished to highlight the importance of considering multiple nutrients for *Trichodesmium* specifically, given the historical emphasis on either Fe or P stress.

Abstract. *Trichodesmium* is a globally important marine microbe that provides fixed nitrogen (N) to otherwise N limited ecosystems. In nature, nitrogen fixation is likely regulated by iron or phosphate availability, but the extent and interaction of these controls are unclear. From metaproteomics analyses using established protein biomarkers for iron and phosphate stress, we found that co-stress is the norm rather than the exception for field *Trichodesmium* colonies. Counter-intuitively, the nitrogenase enzyme was more abundant under co-stress than under single nutrient stress, consistent with the idea that *Trichodesmium* has a specific physiological state under nutrient co-stress, as opposed to single nutrient stress. Organic nitrogen uptake was observed to occur simultaneously with nitrogen fixation. Quantification of the phosphate ABC transporter PstA combined with a cellular model of nutrient uptake suggested that *Trichodesmium* is generally confronted by the biophysical limits of membrane space and diffusion rates for iron and phosphate acquisition in the field. Colony formation may benefit nutrient acquisition from particulate and organic nutrient sources, alleviating these pressures. The results highlight that to predict the behavior of *Trichodesmium*, both Fe and P stress must be evaluated and understood simultaneously.

Introduction

Line 36: colloquialism

Change accepted.

Line 58: add . . .Pho box, a regulatory DNA sequence, which is necessary. . .

Change accepted

Line 65f: Fe and P stress were positively associated – only as co-stress? If yes, say so.

They were associated both individually and under co-stressed.

Also say how Fe, P, and N statuses are closely linked.

We suggest these are linked via a currently unknown regulatory network; change accepted.

Methods

Line 119: what does that mean? Which precursors, of what? Was every protein normalized to the top 3 precursor intensities? Make this clear also to a reader who is not familiar with the specifics of proteomics analysis.

Relative abundance was measured by averaging the peptide precursor/MS1 intensities for the 3 most abundant peptides in the protein, then normalizing this value to the total precursor intensity. Text has been updated:

Raw spectra were searched with the Sequest algorithm using a custom-built genomic database (Eng, Fischer, Grossmann, and MacCoss, 2008). The genomic database consisted of a publically available *Trichodesmium* community metagenome available on the JGI IMG platform (IMG ID 2821474806), as well as the entire contents of the CyanoGEBA project genomes (Shih et al., 2013). Protein annotations were derived from the original metagenomes. SequestHT mass tolerances were set at +/- 10ppm (parent) and +/- 0.8 Dalton (fragment). Cysteine modification of +57.022 and methionine modification of +16 were included. Protein identifications were made with Peptide Prophet in Scaffold (Proteome Software) at the 95% protein and peptide identification levels. Relative abundance was measured by averaging the precursor intensity (area under the MS1 peak) of the top 3 most abundant peptides in each protein, then normalizing this value to total precursor ion intensity. Normalization and global false discovery rate (FDR) calculations, which were 0.1% at the peptide level and 1.2% at the protein level, were performed in Scaffold (Proteome Software). FDR was calculated by Scaffold using the probabilistic method by summing the assigned protein or peptide probabilities and dividing by the maximum probability (100%) for each. The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE partner repository with the dataset identifier PXD016225 and 10.6019/PXD016225 (Perez-Riverol et al., 2019). Statistical tests of relationships between proteins were conducted with the scipy stats package (https://docs.scipy.org/doc/scipy/reference/stats.html) using linear Pearson tests when the relationship appeared to be linear and a Spearman rank order test when this was not the case.

Line 120: How is the FDR defined? What does "0.1% peptide" and "1.2% protein" mean?

Text updated (see above paragraph):

FDR was calculated by Scaffold using the probabilistic method by summing the assigned protein probabilities and dividing by the maximum probability (100%) for each. Different FDRs can be assigned for peptides versus proteins depending on which probabilities are used for the calculation.

Line 128: Which peptides were selected?

We have added to the supplemental Tables S6 and S7 which describe the peptides selected for quantitation

Results and Discussion

Line 178: change "most" to "all but one"

Change accepted

Line 232: please rephrase. What exactly is common in marine bacteria. For sure, all bacteria have regulatory networks.

Change accepted. A recent review of regulatory genes found that regulatory networks may be particularly more abundant in marine organisms (Held et al., 2019).

This indicates that the cell's N, P and Fe statuses are linked, perhaps involving a regulatory network which are particularly common in marine bacteria (Figure 5) (Held et al., 2019).

Line 255ff: skip this justification sentence

With respect we prefer to leave this sentence in because exploratory metaproteomics is not yet widely used as an analytical tool in oceanography. We hope this study and others like it will encourage its adoption. We do welcome editorial advice or further discussion on this point.

Section 3.4. Throughout this section, I think the use of the term ligand is not the norm. For ABC transporters, the word "ligand" is typically used for whatever binds to and is transported by the transporter. The part of the transporter binding the substrate is usually called "ligand-binding protein".

We agree with the reviewer that "ligand" is often used to describe a chemical compound, for instance a siderophore, which can be transported by an ABC transporter. However specifically in the uptake kinetics literature "ligand" is used to describe the ABC transporter itself (i.e. the protein that binds the nutrient). To avoid confusion we have updated the text to use the word "transporter" or "protein" instead.

Line 260: change to ". . .required for both iron and phosphate uptake"

Change accepted

Line 305ff: rewrite sentence. Hard to understand.

Updated, and hopefully clearer now!

For a given surface area: volume quotient, we define nutrient limitation to be caused by either membrane crowding or diffusion limitation depending on which model calculated a higher minimum nutrient concentration.

Line 316ff on cylinders: Shouldn't the Trichodesmium filament, instead of a single Trichodesmium cell, be considered for these models? The effective cell surface of a Trichodesmium cell is reduced by its contact to the neighboring cells.

The reviewer is correct that membrane limitation would be exacerbated for cells living in filaments, as the surface area exposed to the surrounding environment would be reduced. We considered modeling filamentous cells but decided to consider only single cells for clarity since this is the most conservative scenario (i.e. the one in which *Trichodesmium* would theoretically have the most exposure to the environment and be the least limited). We have added a sentence to the discussion of the model highlighting the focus of the model (single cells) but mentioning that filamentous cells would have lower surface area.

While this model may be directly applicable to some N_2-fixing cyanobacteria such as Groups B and C, which have roughly spherical cells, *Trichodesmium* cells are not spheres but rather roughly cylindrical (Hynes et al., 2012). Thus, we repeated the model calculations for cylinders with varying radii (r) and heights (2r or 10r) based on previous estimates of *Trichodesmium* cell sizes (Bergman et al., 2013; Hynes et al., 2012). Cylinders have lower surface area: volume quotient than spheres of similar sizes. In addition, the rate constant (k_D) for diffusion, which is a function of cell geometry, is greater. This increases the slope of the diffusion limitation line such that membrane crowding is important across a greater range of cell sizes (Figure 7c-d). *Trichodesmium* cell sizes vary in nature, for instance the cylinder height can be elongated, improving the surface area: volume quotient. However, the impact of cell elongation to radius r and height 10r on both diffusion limitation and membrane crowding is subtle (Figure 7e-f). Furthermore, though not explicitly considered here, cylindrical cells living in filaments would have reduced surface area available for nutrient uptake. Thus, we conclude that in certain scenarios, lack of membrane space could hypothetically limit Fe and perhaps P acquisition by *Trichodesmium*.

Line 361: Reference missing for mucus production being a "hallmark of Trichodesmium colony formation"

Citation added (Eichner et al., 2019)

Line 362f: If mucus acts as a diffusive barrier, it also does the opposite of "protecting them [the cells] from oxygen", namely preventing O2 to diffuse out of the cells during photosynthesis, which was also shown in Eichner et al, 2019.

Note added to this effect:

A key hallmark of *Trichodesmium* colony formation is production of mucus, which can capture particulate matter and concentrate it within the colony (Eichner et al., 2019). In addition to particle entrainment, the mucus layer can benefit cells by protecting them from oxygen and/or concentrating oxygen during photosynthesis, facilitating epibiont associations,

regulating buoyancy, defending against grazers and helping to "stick" trichomes together (Eichner et al., 2019; Lee et al., 2017; Sheridan, 2002). However, these benefits come at a cost because the mucus layer hinders diffusion to the cell surface (Figure 9), reducing contact with the surrounding seawater. Despite this, the benefits of colony formation seem to outweigh the costs, since *Trichodesmium* forms colonies in the field, particularly under stress (Bergman et al., 2013; Capone et al., 1997; Hynes et al., 2012).

Line 384: Which specific regulatory systems should be characterized? What do you mean by chemical phases?

We don't know yet which regulatory systems should be examined! For a review of marine regulatory systems and their often unknown functions see Held et al. 2019. These results suggest that one or more regulatory networks may control Fe, P, and N status in tandem with one another in *Trichodesmium* cells, and we hope this work will stimulate future research on this topic. By chemical phases we meant dissolved versus particulate nutrient sources, since *Trichodesmium* is known to use both – we've clarified this now:

Future studies should aim to characterize the specific regulatory systems, chemical species and phases (i.e. dissolved versus particulate nutrient sources), and symbiotic interactions that underlie *Trichodesmium's* unique behavior and lifestyle.

Figures

Figure 1: Please use the same numbers on the figure and the legend, or at least also add the figure numbering top the legend.

Figure 1 has been updated (see updated figures below).

Figure 2: Please increase the font size on the legend, and add a legend name like "# of times a protein appeared in the same cluster" – consider changing the legend to a percentage. Please also say in the caption what the color legend shows.

Figure 2 and its caption have been updated

Figure 3: Please state in the caption how the protein abundance values were normalized.

Figure 3 caption has been updated:

"Figure 3. (A) Relative abundance of iron stress protein IdiA (A) and phosphate stress protein SphX (B). IdiA and SphX were among the most abundant proteins in the entire dataset. Error bars are one standard deviation on the mean when multiple samples were available. Dashed lines represent average values across the dataset. Proteins abundances were

normalized such that the total MS1 peak area across the entire proteome was the same for each sample. (C) Relative abundance of IdiA (orange) and SphX (blue) overlaid on the sampling locations."

Figure 4: Please adjust font sized throughout panels. How were the dashed lines in c and d defined? Based on what do they denote Fe- or P-stress? And why are they different in c and d?

Figure 4 font sizes have been updated. The dashed lines in (C) and (D) were drawn by hand to help the reader to visually understand the intersection of Fe and P stress. However, given the important discussion about IdiA and SphX abundances raised by the reviewer, we can see how this might be misleading and have removed the dotted lines.

Figure 5: Does not necessarily need to be a figure if wanting to save space.

We welcome editorial advice on this but would advocate for including the figure as we think it illustrates the discussion in section 3.3, particularly for visual learners.

Reviewer # 1 Point by point response

General overview

Held et al., present a metaproteomic analysis of Trichodesmium isolates mostly col- lected in the Tropical and Subtropical Atlantic and use targeted proteomics to quantify putative markers of P and Fe stress to assess correlations with Nitrogenase. Given that Trichodesmium is a dominant fixer of nitrogen in the oceans, mechanisms con- trolling the abundance the major nitrogen fixing enzyme within this cyanobacter is of great interest to ocean biogeochemists (and should be of interest to all). The study involves a survey of Trichodesmium populations and across multiple cruises and dates which supports a more generalizable perspective of Trichodesmium response to low Fe or P. Rather than simply reporting data and immediate findings regarding correlations between proteins between P and Fe levels, the authors make a considerate attempt to extend hypotheses into the biophysical realm by setting up hypothetical scenarios of protein space competition that may lead to reduction in N fixation via membrane overcrowding. Despite lack of experimental evidence and sometime speculative, the authors cite precedent in many cases and have written a thought-provoking manuscript that should lead to testing of alternative hypotheses with regards to how Trichodesmium counters P and Fe limitations at the membrane in addition to regulatory gene expression.

We thank the reviewer for their thoughtful comments on our manuscript, particularly attention paid to the proteomics methodology. Responses to individual comments are provided below in red. Updated text has been provided for major edits specific to this review with changes highlighted in yellow.

Major comments

Abstract line 19, 20 the authors make reference to a '...specific physiological state under nutrient stress'. Given the fact that everything will have a specific physiological state under stress or non-stress conditions, this statement does not carry much impact. Perhaps the authors could be more specific. Did they intend to say that there is a generalized stress response regardless of stressors and that nitrogenase appears to comprise this stress response?

We agree with the reviewer on this important clarification. The intention was to indicate a specific physiological state under co-stress, distinct from that of an organism stressed for just one nutrient. We have added a phrase to clarify this sentence.

Abstract. *Trichodesmium* is a globally important marine microbe that provides fixed nitrogen (N) to otherwise N limited ecosystems. In nature, nitrogen fixation is likely regulated by iron or phosphate availability, but the extent and interaction of these controls are unclear. From metaproteomics analyses using established protein biomarkers for iron and phosphate stress, we found that co-stress is the norm rather than the exception for field *Trichodesmium* colonies. Counter-intuitively, the nitrogenase enzyme was more abundant under co-stress than under single nutrient stress, consistent with the idea that

Trichodesmium has a specific physiological state under nutrient co-stress, as opposed to single nutrient stress. Organic nitrogen uptake was observed to occur simultaneously with nitrogen fixation. Quantification of the phosphate ABC transporter PstA combined with a cellular model of nutrient uptake suggested that *Trichodesmium* is generally confronted by the biophysical limits of membrane space and diffusion rates for iron and phosphate acquisition in the field. Colony formation may benefit nutrient acquisition from particulate and organic nutrient sources, alleviating these pressures. The results highlight that to predict the behavior of *Trichodesmium*, both Fe and P stress must be evaluated and understood simultaneously.

Intro Line 47 – The use of the term 'established' when describing biomarkers is a bit subjective and loses impact when not followed by multiple citations, which in turn, lends credence to the idea that the protein markers are routinely accepted and utilized by the scientific community. Its reasonable for the authors to state that these proteins were utilized as biomarkers in this study based upon scientific precedent from expression studies, but the discriminatory power of these proteins to classify has not been thor- oughly validated. The rationale provided in line 47-59 is sufficient for inclusion in the study. The word putative or candidate in lieu of 'established' seems a better fit.

We thank the reviewer for highlighting that protein biomarkers are not yet routinely utilized in the community. Change accepted.

There are several ~~established~~ protein biomarkers for Fe and P stress in *Trichodesmium,* all of which are periplasmic binding proteins involved in nutrient acquisition.

Line 151 -157. Point of clarification. The authors state that standard MS2 peak area was linear between 1 amol and 20 fmol per uL. Further that samples were spiked with standard to 10 fmol per uL and 10uL injected. The assumption is that 10uls was also injected for linearity testing of the standards. Please confirm.

Indeed we confirm that 10uL injections were used for linearity testing of the standard peptides. We now clarify this in the text:

The standard mixture was calibrated to establish the exact concentration of the peptides. A known amount (10 fmol μL^{-1}) of the commercially available Pierce standard peptide mixture (Catalog number 88320) and an apomyoglobin digest was spiked into the standard. The ratio of Pierce (isotopically labelled according to JPT standards) or apomyoglobin (light) to heavy standard peptide MS2 peak area was calculated and used to establish the final concentration of the standard peptide mixture (Fu et al., 2016; Milo, 2013). Multiple peptides were used for this calibration and the standard deviation among them was approximately 10%. Finally, the linearity of the peptide standard was tested by generating a dilution curveand ensuring that the concentration of each peptide versus MS2 peak area was linear between 0.001 and 20 fmol μL^{-1} concentration, using 10uL injections consistent with experimental injection volumes.

Methods: There is no mention of transition ions in the PRM section of the methods. For the sake of reproducibility, the authors should make reference to which transition ions were utilized for quantification (can be supplemental). Further the Trichodesmium genome and version should be referenced that was used for searching.

An additional supplemental file (Table S6) providing the transition ions used for PstA quantification will now be provided. The *Trichodesmium* genome used for the initial (DDA) search is referenced in Section 2.4. In PRM mode specific masses were targeted, and the list of targeted ions will now be provided (Table S7). In the process of compiling these tables we noticed a mistake in our prior manuscript – the peptide quantified belongs to PstA, not PstC. Both PstA and PstC are components of the same phosphate transport permease protein, so results and discussion are not affected however it is important to make this distinction, and this naming has been updated in the text.

Line 196 "This clustering indicated direct regulatory links between C and N fixation.." perhaps use suggests in lieu of indicates due to probabilistic nature of the association of protein covariance. Similar strong language should be avoided without follow-up direct experimentation which is beyond the scope of the study.

Agreed and updated, particularly in this section (3.1) which relies on multivariate statistics:

A self-organizing map analysis identified groups of proteins with similar profiles, i.e. proteins whose abundances changed cohesively, suggestive of proteins that may be regulated similarly (Reddy et al., 2016). This revealed the central importance of nitrogen fixation to *Trichodesmium*. The nitrogenase proteins were among the most abundant in the proteome and were located in clusters 1 and 2 (Figure 2 and Table S3). Also in these clusters were nitrogen metabolism proteins including glutamine synthetase, glutamine hydrolyzing guanosine monophosphate (GMP) synthase and glutamate racemase. This is consistent with previous reports finding that N assimilation is synchronized with nitrogen fixation (Carpenter et al., 1992).

Nitrogen fixation was closely linked to carbon fixation. Many photosystem proteins clustered with the nitrogenase proteins, including phycobilisome proteins, photosystem proteins, and the citric acid cycle protein 2-oxoglutarate dehydrogenase. This clustering indicated direct regulatory links between C and N fixation. The nitrogen regulators P-II and NtcA were also present in this cluster and may mediate this association. In non-nitrogen fixing cyanobacteria, high abundance of the nitrogen regulators NtcA and P-II is suggestive of nitrogen stress (Flores and Herrero, 2005; Saito et al., 2014). In diazotrophs, the role of these regulators is unclear because they do not respond to nitrogen compounds such as ammonia as they do in other cyanobacteria (Forchhammer and De Marsac, 1994). Here, clustering of NtcA and P-II with C and N fixation proteins suggests that they play a role in balancing these processes in field populations, though the details of this role has yet to be elucidated.

Line 228 – 229 "This observation contrasts with the current paradigm that Tri- chodesmium down regulates nitrogen fixation when it is Fe or P stressed. . ." This state- ment needs to cite the current published paradigm.

Change accepted:

This observation contrasts with the current paradigm that *Trichodesmium* down regulates nitrogen fixation when it is Fe or P stressed (Frischkorn et al., 2018, Ruoco, et al., 2018, Bergmann et al., 2012, Shi et al., 2007).

Line 267 – 278 when considering membrane protein space, the authors make a con- siderate attempt. Space and physics are often ignored; however, because the authors measured whole cell protein abundance and did not attempt to isolate the plasma mem- brane fraction, the % occupancy estimates are bound to be overestimates. A statement providing limitations of the estimate are needed here. Limitation of space and crowding on a membrane makes sense. Assigning the protein number to the membrane alone is not accurate.

This is an important point and a sentence has been added to the text to indicate this limitation:

"The calculation assumes that 100% of the PstC protein quantified is present in the plasma membrane, however it should be recognized that some fraction is likely present in the cytoplasm, leading to the possibility of over-estimation of membrane space occupied."

Line 321 "Thus, we conclude that in certain scenarios, lack of membrane space could indeed limit Fe and perhaps P acquisition by Trichodesmium." There is no disagree- ment with the rationale and calculations that led to this statement, but the statement is hypothetical in nature and the term 'hypothetical' should be included. The authors do a nice job in addressing model limitations in the next paragraph and go on to describe an artificial scenario where space limitation could produce further nutrient uptake lim- itation as additional proteins are made and transported to the membrane. One has to assume uptake activity does not change or is influenced by intracellular events or membrane compositional changes that lead to conformational changes; however, the idea that a generalized stress response to Fe or P could lead to a negative effect on uptake and more limitation due to space limits is fun to ponder.

We thank the reviewer for their positive comment and interest in this section. We have updated Line 321 to highlight uncertainty in the statement:

Thus, we conclude that in certain scenarios, lack of membrane space could hypothetically limit Fe and perhaps P acquisition by *Trichodesmium*.

Line 369 – 385, Conclusions. The conclusion section is written as a perspective which is fine given that the conclusions linked specifically to the data are stated within the results and discussion. In line 372,

the use of the word 'norm' is understandable given the common phrase 'the norm rather than the exception', but this might be contentious because normal is being assigned to all Trichodesmium based on 16 sampling sites mainly focused in the Atlantic. Fe and P stress may be more common that previously accepted or realized. If the authors feel strongly about this phrasing and believe the audience will be receptive and not over-interpret, then it is fine. Otherwise, perhaps it can be a bit more tempered. The same phrase was used in the Abstract, but due to space limitations the fact that authors were inferring norm based on their samples seems reasonable.

We thank the reviewer for this comment. We have altered our wording in this section to avoid over-reach (similar changes will also be made elsewhere in the paper):

Trichodesmium's colonial lifestyle likely produces challenges for dissolved Fe and P acquisition, which must be compensated for by production of multiple nutrient transport systems, such as for particulate iron and organic phosphorous, at a considerable cost. While laboratory studies have largely focused on single nutrient stresses in free filaments, these metaproteomic observations and accompanying nutrient uptake model demonstrate that Fe and P co-stress may be the norm rather than the exception, particularly in the North Atlantic ocean. This means that the emphasis on single limiting nutrients in culture studies and biological models may not capture the complexities of *Trichodesmium's* physiology in situ. Thus, biogeochemical models should consider incorporating Fe and P co-stress conditions. Specifically, in this study and in others there is evidence that nitrogen fixation is optimal under co-limited or co-stressed conditions, implying that an input of either Fe or P could counter-intuitively decrease N_2 driven new production (Garcia et al., 2015; Walworth et al., 2016).

Conclusion Line 378-385, membrane space limitation is likely to be confronted by all cells, not only Trichodesmium, but the idea is understood. The authors make a case for including co-stress based on protein observations and if nutrient stress is truly oc- curring (which can be difficult to define to everyone's liking and measure in situ; hence use of protein markers) then including these parameters in biological models makes sense.

We agree that membrane space is likely confronted by all cells, though few oceanographic studies have demonstrated or discussed this. In the interest of not over-reaching our argument we leave the reviewer's point out of our main text but hope that this work will stimulate future investigations.

Figure 3. This is a very informative and nice figure. If possible it would be great to see Figure S2 incorporated for ease of reading and comparison.

Both reviewer #1 and reviewer #2 suggested this change, so we now include the Figure S2 plots in the main text as part of Figure 3. Figure S2 now provides scatter plots of IdiA relating dFe and SphX versus phosphate concentrations. Per these changes and the comments from the reviewers the discussion in Section 3.2 has been updated as below.

3.2 *Trichodesmium* is simultaneously iron and phosphate stressed throughout its habitat

A surprising emergent observation from the *Trichodesmium* metaproteomes was the co-occurrence of the iron (IdiA) and phosphate (SphX) stress biomarkers across the samples. The ubiquitous and highly abundant presence of these proteins relative to total protein implied that co-stress may be the norm rather than the exception for *Trichodesmium* colonies in the field, particularly in the North Atlantic. Even though low-level basal expression of IdiA and SphX has been observed, it was clear that the colonies were devoting a large fraction of their cellular resources to Fe and P uptake, respectively (see Tables S8 and S9) (Webb et al., 2001, Webb et al., 2007, Chappell et al., 2010, Orchard et al., 2010, Snow et al., 2015, Walworth et al., 2016, Frischkorn et al., 2019). This, combined with the responsiveness of IdiA and SphX to nutrient availability in laboratory experiments on *Trichodesmium* filaments laboratory, indicated that co-stress was occurring.

Interestingly, biomarker abundance was not necessarily associated with nutrient concentrations in the surface ocean, suggesting that the colonies were experiencing stress despite variation in nutrient availability (Figure 3 C-D). SphX abundance varied up to 7.5 fold and may have been negatively associated with dissolved phosphate concentrations, though analytical differences across the field expeditions may have forced this relationship (Figure S2). Oceanographically, SphX was most abundant in the P-deplete, summer-stratified North Atlantic gyre (JC150 expedition) compared with winter waters near the Amazon river plume (Tricolim expedition) or at station ALOHA, where phosphate concentrations were greater (Hynes et al., 2009; Sañudo-Wilhelmy et al., 2001; Wu et al., 2000). IdiA varied up to 8 fold but there was no observable relationship with dFe concentrations at the surface. Instead, IdiA may be responsive to other factors such as the varying iron requirements of the populations/species examined here. For instance, it should be highlighted that in this study only *Trichodesmium* colonies were examined, so factors such as colony size may affect iron availability and biomarker expression. Additionally, because the surface ocean iron inventory is low, transient inputs such as from the Sahara desert can dramatically impact iron availability on short time scales, and the time scale of these inputs relative to changes in biomarker abundance is not well understood (Kunde et al., 2019). Carefully calibrated datasets relating IdiA and SphX abundance to nutrient-limited growth rates of *Trichodesmium* in both the filamentous and colonial forms would facilitate further interpretation of this data.

Figure 9 is unnecessary, but certainly would be welcomed by visual learners if space is not an issue. Otherwise the concepts are described in the discussion.

As the first author is a visual learner and space is not an issue, she would prefer to leave this in, but welcomes editorial advice on the matter!

Minor comments

Line 63-64"..suggesting nutrient stress was driven not only by biogeochemical gradi- ents but also by Trichodesmium's inherent physiology". The term 'inherent physiology' is very broad and does not add

substance to the sentence. Trichodesmium is respond- ing to stress in the study and saying something like ". . ..but also by Trichodesmium's response to stress" puts the sentence in a category that doesn't contain every bio- chemical reaction in the organism.

Rephrased: "suggesting nutrient stress was driven not only by biogeochemical gradients but also by *Trichodesmium's* response to nutrient depletion"

Line 75. Table S1 does not correspond to the supplemental table indicated. Please correct the designation.

Change accepted – should read Table S2

Line 100 ". . ...vacuum centrifugation to 1 100 µg µL-1 concentration." Assume this was estimated based on starting concentration of protein and not actually measured?

Yes, this is estimated from starting concentration and will now be noted in the text:

The resulting peptide mixture was concentrated by vacuum centrifugation to 1 µg μL^{-1} concentration estimated from the starting protein concentration.

Line 107 "C18 columns packed in house." Please add column size, diameter, c18 particle size and supplier.

This section has been updated with more details per the reviewer:

2.3 Sample acquisition

The global proteomes were analysed by online comprehensive active-modulation two-dimensional liquid chromatography (LC x LC-MS) using high and low pH reverse phase chromatography with inline PLRP-S (200µm x 150mm, 3µm bead size, 300A pore size, NanoLCMS Solutions) and C18 columns packed in house (100 mm x 150 mm, 3 µm particle size, 120 Å pore size, C18 Reprosil-Gold, Dr. Maisch GmbH packed in a New Objective PicoFrit column). The first dimension utilized an 8 hour pH = 10 gradient (10mM ammonium formate and 10mM ammonium formate in in 90% acetonitrile), and was trapped every 30min on alternating dual traps, then eluted at 500nL/min onto the C18 column with a 30 min gradient (0.1% formic acid and 0.1% formic acid in 99.9% acetonitrile). 10 µg of protein was injected per run directly onto the first column using a Thermo Dionex Ultimate3000 RSLCnano system (Waltham, MA), and an additional RSLCnano pump was used for the second dimension gradient. The samples were then analyzed on a Thermo Orbitrap Fusion mass spectrometer with a Thermo Flex ion source (Waltham, MA). MS1 scans were monitored between 380-1580 m/z, with a 1.6 m/z MS2 isolation window (CID mode), 50 millisecond maximum injection time and 5 second dynamic exclusion time.

Line 110 A little more detail regarding the parameters would be useful. Based on the search parameters the instrument was likely operated in orbitrap/ion trap mode with HCD? This would be of interest to include and assume more details are in located in Pride.

Details are located in Pride but the reviewer is right that more should be included in the main text.

Added sentence: "MS1 scans were monitored between 380-1580 m/z, with a 1.6 m/z MS2 isolation window (CID mode), 50 millisecond maximum injection time and 5 second dynamic exclusion time."

Line 120 should include the term "local FDR" if local FDR was used.

These are global FDRs (will now be noted).

Line 179-180 is repeated from methods section. Can be deleted.

Change accepted

Line 185 ribosomal and phycobilisomal

Change accepted

Line 237 – 257 This is quite fascinating although not the focus of the study. urtA substrate specificity is poorly defined outside of urea. Curious if the authors also found urease protein elevation in stressed samples.

Unfortunately urease protein abundances were very patchy (and were generally in low abundance) so we were unable to draw any specific conclusions. We did make an effort to note that increased abundance of the urea ABC transporter indicates general use of organic N sources including urea but also possibly TMA or other compounds.

Figure S2 "Note that the phosphate concentrations from the Tricolim cruise were not measured at the. . ." Do the authors mean to say Tricolim_13?

Per the reviewers we now include the data Figure S2 in the main text as part of Figure 3, and have updated the caption accordingly.

Supplemental figures. Please read through legends for spelling errors.

Change accepted.

Co-occurrence of Fe and P stress in natural populations of the marine diazotroph *Trichodesmium*

Noelle A. Held[1,2], Eric A. Webb[3], Matthew M. McIlvin[1], David A. Hutchins[3], Natalie R. Cohen[1], Dawn M. Moran[1], Korinna Kunde[4], Maeve C. Lohan[4], Claire M. Mahaffey[5], E. Malcolm S. Woodward[6], Mak A. Saito[1*]

[1]Department of Marine Chemistry and Geochemistry, Woods Hole Oceanographic Institution, Woods Hole, MA 02543 USA
[2]Department of Earth, Atmospheric, and Planetary Sciences, Massachusetts Institute of Technology, Cambridge, MA. 02139 USA
[3]Marine and Environmental Biology, Department of Biological Sciences, University of Southern California, Los Angeles, CA, 90089 USA
[4]Ocean and Earth Science, National Oceanography Centre, University of Southampton, Southampton, UK
[5]Department of Earth, Ocean and Ecological Sciences, University of Liverpool, Liverpool, UK
[6]Plymouth Marine Laboratory, Plymouth, UK

Correspondence to: Mak Saito (msaito@whoi.edu)

Abstract. *Trichodesmium* is a globally important marine microbe that provides fixed nitrogen (N) to otherwise N limited ecosystems. In nature, nitrogen fixation is likely regulated by iron or phosphate availability, but the extent and interaction of these controls are unclear. From metaproteomics analyses using established protein biomarkers for iron and phosphate stress, we found that co-stress is the norm rather than the exception for field *Trichodesmium* colonies. Counter-intuitively, the nitrogenase enzyme was more abundant under co-stress as opposed to single nutrient stress. This is consistent with the idea that *Trichodesmium* has a specific physiological state during nutrient co-stress. Organic nitrogen uptake was observed and occurred simultaneously with nitrogen fixation. Quantification of the phosphate ABC transporter PstA combined with a cellular model of nutrient uptake suggested that *Trichodesmium* is generally confronted by the biophysical limits of membrane space and diffusion rates for iron and phosphate acquisition in the field. Colony formation may benefit nutrient acquisition from particulate and organic nutrient sources, alleviating these pressures. The results highlight that to predict the behavior of *Trichodesmium*, both Fe and P stress must be evaluated simultaneously.

1. Introduction

The diazotrophic cyanobacterium *Trichodesmium* plays an important ecological and biogeochemical role in the tropical and subtropical oceans globally. By providing bioavailable nitrogen (N) to otherwise N-limited ecosystems, it supports basin-scale food webs, increasing primary productivity and carbon flux from the surface ocean (Capone, 1997; Carpenter and Romans, 1991; Coles et al., 2004; Deutsch et al., 2007; Sohm et al., 2011). Nitrogen fixation is energetically and nutritionally expensive, so it typically occurs when other sources of N are unavailable, i.e. in N-starved environments (Karl et al., 2002). However, nitrogen availability is not the sole control on nitrogen fixation, which must be balanced against

the cell's overall nutritional status. Because it can access a theoretically unlimited supply of atmospheric N_2, *Trichodesmium* often becomes phosphorus (P) limited (Frischkorn et al., 2018; Hynes et al., 2009; Orchard, 2010; Sañudo-Wilhelmy et al., 2001; Wu et al., 2000). It also has a tendency to experience iron (Fe) limitation because the nitrogenase enzyme is iron-demanding (Bergman et al., 2013; Chappell et al., 2012; Rouco et al., 2018; Sunda, 2012; Walworth et al., 2016).

There is uncertainty about when and where *Trichodesmium* is Fe and P stressed and how this impacts nitrogen fixation in nature. Some reports suggest that *Trichodesmium* is primarily phosphate stressed in the North Atlantic, and primarily Fe stressed in the Pacific, owing to relative Fe and P availability in these regions (Bergman et al., 2013; Chappell et al., 2012; Frischkorn et al., 2018; Hynes et al., 2009; Orchard, 2010; Sañudo-Wilhelmy et al., 2001). However, others have suggested that Fe and P can be co-limiting to *Trichodesmium*; one incubation study found two examples of Fe/P co-limitation in the field (Mills et al., 2004). Even less clear is how Fe and/or P stress impacts nitrogen fixation. For instance, despite the intuitive suggestion that nitrogen fixation is limited by Fe or P availability, laboratory evidence indicated that *Trichodesmium* is specifically adapted to co-limited conditions, with higher growth and N_2-fixation rates under co-limitation than under single nutrient limitation (Garcia et al., 2015; Walworth et al., 2016).

There are several ~~established~~ protein biomarkers for Fe and P stress in *Trichodesmium,* all of which are periplasmic binding proteins involved in nutrient acquisition. For Fe, this includes the IdiA and IsiB proteins and for phosphorus, specifically phosphate, the PstS and SphX proteins (see Table S1). In *Trichodesmium,* IsiB, a flavodoxin, and IdiA, an ABC transport protein, are expressed under Fe limiting conditions, and both are conserved across species with high sequence identity (Chappell et al., 2012; Webb et al., 2007). Transcriptomic and proteomic studies have shown that they are more abundant under Fe stress conditions, though there is low-level basal level expression (Chappell et al., 2012; Snow et al., 2015; Walworth et al., 2016). In this dataset, IsiB and IdiA were both highly abundant and correlated to one another (Figure S1). IdiA is used as the molecular biomarker of Fe stress in the following discussion, but the same conclusions could be drawn from IsiB distributions. Like IdiA and IsiB, SphX and PstS are conserved across diverse *Trichodesmium* species (Chappell et al., 2012; Walworth et al., 2016). SphX is abundant at the transcript and protein level under phosphate limitation (Orchard et al., 2009; Orchard, 2010). PstS, a homologous protein located a few genes downstream of SphX, responds less clearly to phosphate stress. In *Trichodesmium*, the reason may be that PstS is not preceded by a Pho box, a regulatory DNA sequence which is necessary for P based regulation (Orchard et al., 2009). Thus, in this study we focused on SphX as a measure of phosphate stress and IdiA as a marker of Fe stress.

Here, we present evidence based on field metaproteomes that *Trichodesmium* colonies were simultaneously Fe and P stressed, particularly in the tropical and subtropical Atlantic. While Fe/P stress has been suggested before, this study provides molecular evidence for co-stress in a broad geographical and temporal survey. This co-stress occurred across significant gradients in Fe and P concentration, suggesting nutrient stress was driven not only by biogeochemical gradients but also by *Trichodesmium's* response to nutrient depletion; we explore possible biophysical and biochemical mechanisms behind this. Fe and P stress were positively associated with nitrogen fixation and organic nitrogen uptake, suggesting that *Trichodesmium's* Fe, P, and N statuses are linked, perhaps via a regulatory network.

2. Materials and methods

2.1 Sample acquisition

A total of 37 samples were examined in this study. Samples were acquired by the authors on various research expeditions and most exist in biological duplicate or triplicate (Table S2). *Trichodesmium* colonies were hand-picked from 200 μm or 130 μm surface plankton net tows, rinsed thrice in 0.2 μm filtered trace metal clean surface seawater into trace metal clean LDPE bottles, decanted onto 0.2-5 μm filters, and frozen until protein extraction. The samples were of mixed puff and tuff morphology types, depending on the natural diversity present at the sampling location. The majority of samples considered in this study were taken in the early morning pre-dawn hours. Details such as filter size, morphology, location, cruise, date, and time of sampling are provided in Table S2.

2.2 Sample acquisition

Proteins were extracted by a detergent based method following Saito et al. (2014) and Lu et al. (2005). To reduce protein loss and contamination, all tubes were ethanol rinsed and dried prior to use and all water and organic solvents used were LC/MS grade. Sample filters were placed in a tube with 1-2 mL 1% sodium dodecyl sulfate (SDS) extraction buffer (1% SDS, 0.1 M Tris/HCL pH 7.5, 10 mM EDTA) and incubated for 10 min at 95°C with shaking, then for one hour at room temperature with shaking. The protein extract was decanted and clarified by centrifugation (14100xg) at room temperature. The crude protein extracts were quantified with the colormetric BCA protein concentration assay with bovine serum albumin as a standard (Pierce catalog number 23225). Extracts were concentrated by 5 kD membrane centrifugation (Vivaspin spin columns, GE Healthcare). The protein extracts were purified by organic precipitation in 0.5 mM HCl made in 50% methanol and 50% acetone at -20 °C for at least one week, then collected by centrifugation at 14100xg for 30 min at 4 °C, decanted and dried by vacuum concentration for 10min. The protein pellets were re-suspended in a minimum amount of 1% SDS extraction buffer, and re-quantified by BCA protein concentration assay to assess extraction efficiency.

The proteins were embedded in a 500 μL final volume acrylamide gel, which was then cut up into 1 mm pieces to maximize surface area and rinsed in 50:50 acetonitrile: 25 mM ammonium bicarbonate overnight at room temperature. The next morning, the rinse solution was replaced and the rinse repeated for 1 hour. Gels were dehydrated thrice in acetonitrile, dried by vacuum centrifugation, and rehydrated in 10 mM dithiothrietol (DTT) in 25 mM ammonium bicarbonate, then incubated for one hour at 56 °C with shaking. Unabsorbed DTT solution was removed and the volume recorded, allowing for calculation of the total gel volume. Gels were washed in 25 mM ammonium bicarbonate, then incubated in 55 mM iodacetamide for one hour at room temperature in the dark. Gels were again dehydrated thrice in acetonitrile. Trypsin (Promega Gold) was added at a ratio of 1:20 μg total protein in 25 mM ammonium bicarbonate in a volume sufficient to barely cover the gel pieces. Proteins were digested overnight at 37 °C with shaking. Any unabsorbed solution was then removed to a new tube and 50μL of peptide extraction buffer (50% acetonitrile, 5% formic acid in water) was added and incubated for 20 min at room temperature. The supernatant as then decanted and combined with the unabsorbed solution, and

the step then repeated. The resulting peptide mixture was concentrated by vacuum centrifugation to 1 μg μL^{-1} concentration estimated from the starting protein concentration. Finally, the peptides were clarified by centrifugation at room temperature, taking the top 90% of the volume to reduce the carry over of gel debris.

2.3 Data acquisition

The global proteomes were analysed by online comprehensive active-modulation two-dimensional liquid chromatography (LC x LC-MS) using high and low pH reverse phase chromatography with inline PLRP-S (200μm x 150mm, 3μm bead size, 300A pore size, NanoLCMS Solutions) and C18 columns packed in house (100 mm x 150 mm, 3 μm particle size, 120 Å pore size, C18 Reprosil-Gold, Dr. Maisch GmbH packed in a New Objective PicoFrit column). The first dimension utilized an 8 hour pH = 10 gradient (10mM ammonium formate and 10mM ammonium formate in in 90% acetonitrile), and was trapped every 30min on alternating dual traps, then eluted at 500nL/min onto the C18 column with a 30 min gradient (0.1% formic acid and 0.1% formic acid in 99.9% acetonitrile). 10 μg of protein was injected per run directly onto the first column using a Thermo Dionex Ultimate3000 RSLCnano system (Waltham, MA), and an additional RSLCnano pump was used for the second dimension gradient. The samples were then analyzed on a Thermo Orbitrap Fusion mass spectrometer with a Thermo Flex ion source (Waltham, MA). MS1 scans were monitored between 380-1580 m/z, with a 1.6 m/z MS2 isolation window (CID mode), 50 millisecond maximum injection time and 5 second dynamic exclusion time.

2.4 Relative quantitation of peptides and proteins

Raw spectra were searched with the Sequest algorithm using a custom-built genomic database (Eng, Fischer, Grossmann, and MacCoss, 2008). The genomic database consisted of a publically available *Trichodesmium* community metagenome available on the JGI IMG platform (IMG ID 2821474806), as well as the entire contents of the CyanoGEBA project genomes (Shih et al., 2013). Protein annotations were derived from the original metagenomes. SequestHT mass tolerances were set at +/- 10ppm (parent) and +/- 0.8 Dalton (fragment). Cysteine modification of +57.022 and methionine modification of +16 were included. Protein identifications were made with Peptide Prophet in Scaffold (Proteome Software) at the 95% protein and peptide identification levels. Relative abundance was measured by averaging the precursor intensity (area under the MS1 peak) of the top 3 most abundant peptides in each protein, then normalizing this value to total precursor ion intensity. Normalization and global false discovery rate (FDR) calculations, which were 0.1% at the peptide level and 1.2% at the protein level, were performed in Scaffold (Proteome Software). FDR was calculated by Scaffold using the probabilistic method by summing the assigned protein or peptide probabilities and dividing by the maximum probability (100%) for each. The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE partner repository with the dataset identifier PXD016225 and 10.6019/PXD016225 (Perez-Riverol et al., 2019). Statistical tests of relationships between proteins were conducted with the scipy stats package

(https://docs.scipy.org/doc/scipy/reference/stats.html) using linear Pearson tests when the relationship appeared to be linear and a Spearman rank order test when this was not the case.

2.5 Absolute quantitation of peptides

A small number of peptides were selected for absolute quantitation using a modified heterologous expression system. The peptides were ensured to be specific to *Trichodesmium* species based on sequence identity compared to over 300 marine bacteria genomes, three metagenomes, and 956 specialized assemblies (see www.metatryp.whoi.edu) (Saito et al., 2015). A custom plasmid was designed that contained the *Escherichia coli* K12 optimized reverse translation sequences for peptides of interest separated by tryptic spacers (protein sequence = TPELFR). The peptides and transition ions included are provided in Table S7. To avoid repetition of the spacer nucleotide sequence, twelve different codons were utilized to encode the spacer. Six equine apomyoglobin and three peptides from the commercially available Pierce peptide retention time calibration mixture (product number 88320) were also included. The sequence was inserted into a pet(30a)+ plasmid using the BAMH1 5' and XhoI 3' restriction sites.

The plasmid was transformed into competent tuner(DE3)pLys *E.coli* cells and grown on kanamycin amended LB agar plates to ensure plasmid incorporation. A single colony was used to inoculate a small amount of kanamycin containing ^{15}N labelled S.O.C. media (Cambridge Isotope Laboratories) as a starter culture. These cells were grown overnight and then used to inoculate 10 mL of ^{15}N labeled, kanamycin-containing SOC media. Cells were grown to approximately OD600 0.6, then induced with 1 mM isopropyl β-D-1-thiogalactopyranoside (IPTG), incubated in the overexpression phase overnight at room temperature and harvested by centrifugation.

Cells were lysed with BugBuster detergent with added benzonase nuclease. The extracts were centrifuged and a large pellet of insoluble cellular material remained. Because the plasmid protein was large, this pellet contained a large number of inclusion bodies containing nearly pure protein. The inclusion bodies were solubilized in 6 M urea at 4 °C overnight. The protein was reduced, alkylated, and trypsin digested in solution to generate a standard peptide mixture.

The standard mixture was calibrated to establish the exact concentration of the peptides. A known amount (10 fmol μL^{-1}) of the commercially available Pierce standard peptide mixture (Catalog number 88320) and an apomyoglobin digest was spiked into the standard. The ratio of Pierce (isotopically labelled according to JPT standards) or apomyoglobin (light) to heavy standard peptide MS2 peak area was calculated and used to establish the final concentration of the standard peptide mixture (Fu et al., 2016; Milo, 2013). Multiple peptides were used for this calibration and the standard deviation among them was approximately 10%. Finally, the linearity of the peptide standard was tested by generating a dilution curveand ensuring that the concentration of each peptide versus MS2 peak area was linear between 0.001 and 20 fmol μL^{-1} concentration, using 10uL injections consistent with experimental injection volumes.

The sample was prepared at 0.2 µg μL^{-1} concentration, with 10 µL injected to give a total of 2 µg total protein analyzed. The heavy labelled standard peptide mixture was spiked into each sample at a concentration of 10 fmol μL^{-1}. The concentration of the light peptide was calculated as the ratio of the MS2 area of the light:heavy peptide multiplied by 10 µg

μL^{-1}. A correction was applied for protein recovery before and after purification, and the result was the absolute concentration of the peptide in fmol μg^{-1} total protein.

The percent of the membrane occupied by the ABC transporter PstA was calculated by converting the absolute protein concentration to molecules per *Trichodesmium* cell, using average values for *Trichodesmium* cell volume (Hynes et al., 2012), carbon content per volume (Strathmann, 1967), protein content per g carbon (Rouwenhorst, et al., 1991), and the cross sectional area of a calcium ATPase (Hudson et al., 1992) (see Table S3).

2.6 Self-organizing map analyses

Self-organizing maps were used to reduce the dimensionality of the data and explore relationships among co-varying proteins of interest. Only *Trichodesmium* proteins were considered. Analyses were conducted in Python 3.0 and fully reproducible code is available at https://github.com/naheld/self_organizing_map_tricho_metaP.

The input data consisted of a table of protein names (rows) and samples (columns) such that the input vectors contained 2818 features. To eliminate effects of scaling, the data was unit normalized with the Scikit-learn pre-processing algorithm. The input vectors were used to initialize a 100 output node (10x10) self-organizing map using the SOMPY Python library (https://github.com/sevamoo/SOMPY). The output nodes were then clustered using a k-means clustering algorithm (k = 10) implemented in scikit learn. The input nodes (proteins) assigned to each map node were then retrieved and the entire process repeated 10,000 times. Proteins were considered in the same cluster if they appeared in the same cluster of output nodes more than 99.99% of the time.

3. Results and Discussion

3.1 Proteome overview

This study presents 36 field metaproteomes of colonial *Trichodesmium* populations collected at sixteen locations on four expeditions (Table S2). All but one location were in the subtropical and tropical Atlantic; most samples were collected in the early morning hours to avoid changes occurring on the diel cycle (Figure 1 and Table S2). ~~At each location, *Trichodesmium* colonies were hand-picked from plankton net tows, rinsed in filtered seawater, collected onto filters, and immediately frozen.~~ The metaproteomes were analyzed with a two-dimensional LC-MS/MS workflow that provided deep coverage of the proteome. This resulted in 4478 protein identifications, of which 2944 were *Trichodesmium* proteins. The remaining proteins were from colony-associated epibionts, and will be discussed in a future publication. Protein abundance is presented as precursor (MS1) intensity of the three most abundant peptides for each protein, normalized to total protein in the sample. Thus, changes in protein abundance were interpreted as changes in the fraction of the proteome devoted to that protein. The most abundant were GroEL, ribosomal, and phycobilisome proteins.

A self-organizing map analysis identified groups of proteins with similar profiles, i.e. proteins whose abundances changed cohesively, suggestive of proteins that may be regulated similarly (Reddy et al., 2016). This revealed the central importance of nitrogen fixation to *Trichodesmium*. The nitrogenase proteins were among the most abundant in the proteome

and were located in clusters 1 and 2 (Figure 2 and Table S3). Also in these clusters were nitrogen metabolism proteins including glutamine synthetase, glutamine hydrolyzing guanosine monophosphate (GMP) synthase and glutamate racemase. This is consistent with previous reports finding that N assimilation is synchronized with nitrogen fixation (Carpenter et al., 1992).

Nitrogen fixation was closely linked to carbon fixation. Many photosystem proteins clustered with the nitrogenase proteins, including phycobilisome proteins, photosystem proteins, and the citric acid cycle protein 2-oxoglutarate dehydrogenase. This clustering indicated direct regulatory links between C and N fixation. The nitrogen regulators P-II and NtcA were also present in this cluster and may mediate this association. In non-nitrogen fixing cyanobacteria, high abundance of the nitrogen regulators NtcA and P-II is suggestive of nitrogen stress (Flores and Herrero, 2005; Saito et al., 2014). In diazotrophs, the role of these regulators is unclear because they do not respond to nitrogen compounds such as ammonia as they do in other cyanobacteria (Forchhammer and De Marsac, 1994). Here, clustering of NtcA and P-II with C and N fixation proteins suggests that they play a role in balancing these processes in field populations, though the details of this role has yet to be elucidated.

In addition to identifying links between C and N fixation, the self-organizing map analysis demonstrated that field populations of *Trichodesmium* invest heavily in macro- and micro-nutrient acquisition. There were clusters of proteins involved in trace metal acquisition and management, including Fe, zinc, and metal transport proteins, with the latter including proteins likely involved in Ni and Mo uptake (protein IDs TCCM_0270.00000020 and TCCM_0481.00000160). We also noted clusters of proteins involved in phosphate acquisition. Importantly, SphX and PstS appear in separate clusters, highlighting differential regulation of these functionally similar proteins.

3.2 *Trichodesmium* is simultaneously iron and phosphate stressed throughout the North Atlantic

A surprising emergent observation from the *Trichodesmium* metaproteomes was the co-occurrence of the iron (IdiA) and phosphate (SphX) stress biomarkers across the samples. The ubiquitous and highly abundant presence of these proteins relative to total protein implied that co-stress may be the norm rather than the exception for *Trichodesmium* colonies in the field, particularly in the North Atlantic. Even though low-level basal expression of IdiA and SphX has been observed, it was clear that the colonies were devoting a large fraction of their cellular resources to Fe and P uptake, respectively (see Tables S8 and S9) (Webb et al., 2001, Webb et al., 2007, Chappell et al., 2010, Orchard et al., 2010, Snow et al., 2015, Walworth et al., 2016, Frischkorn et al., 2019). This, combined with the responsiveness of IdiA and SphX to nutrient availability in *Trichodesmium* filaments in the laboratory, indicated that co-stress was occurring.

Interestingly, biomarker abundance was not necessarily associated with nutrient concentrations in the surface ocean, suggesting that the colonies were experiencing stress despite variation in nutrient availability (Figure 3 C-D). SphX abundance varied up to 7.5 fold and was negatively associated with dissolved phosphate concentrations, though analytical differences across the field expeditions may have forced this relationship (Figure S2). Oceanographically, SphX was most abundant in the P-deplete, summer-stratified North Atlantic gyre (JC150 expedition) compared with winter waters near the

Amazon river plume (Tricolim expedition) or at station ALOHA, where phosphate concentrations were greater (Hynes et al., 2009; Sañudo-Wilhelmy et al., 2001; Wu et al., 2000). IdiA varied up to 8 fold but there was no observable relationship with dFe concentrations at the surface. Instead, IdiA may be responsive to other factors such as the varying iron requirements of the populations/species examined. It should be highlighted that in this study only *Trichodesmium* colonies were considered, so factors such as colony size may have affected iron availability and biomarker expression. Additionally, because the surface ocean iron inventory was low, transient inputs such as from the Sahara desert could dramatically impact iron availability on short time scales, and the time scale of these inputs relative to changes in biomarker abundance is not well understood (Kunde et al., 2019). Carefully calibrated datasets relating IdiA and SphX abundance to nutrient-limited growth rates of *Trichodesmium* in both the filamentous and colonial forms would facilitate quantitative interpretation of this data.

3.3 The intersection of Fe, P and N stress

The metaproteomes enabled the relationship between Fe and P stress and overall cellular metabolism to be explored. Nitrogenase protein abundance was positively correlated with both IdiA and SphX, and was in fact highest at the intersection of high Fe and P stress (Figure 4). This observation contrasts with the current paradigm that *Trichodesmium* down regulates nitrogen fixation when it is Fe or P stressed (Frischkorn et al., 2018, Ruoco, et al., 2018, Bergmann et al., 2012, Shi et al., 2007). Instead, it is consistent with the idea that the nutritional demands of nitrogen fixation could drive the organism to Fe and P stress, thereby initiating an increase in Fe and P acquisition proteins including IdiA and SphX. This indicates that the cell's N, P and Fe statuses are linked, perhaps involving one or more regulatory networks, which are particularly common in marine bacteria (Figure 5) (Held et al., 2019). This network may regulate a specific physiological adaptation to nutrient co-stress. For instance, Fe and P co-limited *Trichodesmium* cells may reduce their cell size to optimize their surface area: volume quotient for nutrient uptake. However, a putative cell size biomarker Tery_1090, while abundant in co-limited cells in culture, was not identified in these metaproteomes despite bioinformatic efforts to target it, likely because it is a low abundance protein (Walworth et al., 2016).

Nitrogen fixation is not the only way that *Trichodesmium* can acquire fixed N (Dyhrman et al., 2006; Küpper et al., 2008; Mills et al., 2004; Sañudo-Wilhelmy et al., 2001). In culture, *Trichodesmium* can be grown on multiple nitrogen sources including urea; in fact, it has been reported that nitrogen fixation provides less than 20% of the fixed N demand of cells, and a revised nitrogen fixation model suggests that *Trichodesmium* takes up fixed nitrogen in the field (McGillicuddy , 2014; Mulholland and Capone, 1999). In this dataset, a urea ABC transporter was abundant, indicating that urea could be an important source of fixed nitrogen to colonies (Figure 6a). The transporter is unambiguously attributed to *Trichodesmium* rather than a member of the epibiont community. Of course, this does not rule out the possibility that urea or other organic nitrogen sources such as trimethylamine (TMA) are also utilized by epibionts, although no such epibiont transporters were identified in the metaproteomes.

Typically, elevated urea concentration decreases or eliminates nitrogen fixation in colonies (Ohki, et al., 1991). However, in laboratory studies urea exposure must be unrealistically high (often over 20 μM) for this to occur, compared

with natural concentrations which are much lower (Ohki et al., 1991; Wang et al., 2000). In the field, urea utilization and nitrogen fixation seem to occur simultaneously, with a urea uptake protein positively correlated to nitrogenase abundance (Figure 6b). Urea and other organic nitrogen sources such as trimethylamine (TMA) could be sources of nitrogen for *Trichodesmium*, and the relationship to nitrogenase abundance may indicate a general N stress signature driving both organic nitrogen uptake and nitrogen fixation (Walworth et al., 2018). Alternatively, urea uptake could be a colony-specific behavior, since colonies were sampled here as opposed to laboratory cultures that typically grow as single filaments. For instance, urea could be used for recycling of fixed N within the colony, or there could be heterogeneity in nitrogen fixation, with some cells taking up organic nitrogen and others fixing it. These unexpected observations of co-occurring nitrogen fixation and organic nitrogen transport show the value of exploratory metaproteomics, which does not require targeting of a specific protein based on a prior hypothesis.

3.4 Mechanisms of simultaneous iron and phosphate stress – membrane crowding

ABC transporters are multi-unit, trans-membrane protein complexes that use ATP to shuttle substrates across membranes. Specific ABC transporters are required for both iron versus phosphate uptake (Chappell et al., 2012; Orchard et al., 2009). Nutrient transport rates can be modulated by changing the number of uptake proteins installed on the cell membrane or the efficiency of the uptake proteins through expression of assisting proteins such as IdiA and SphX, which bind Fe or P respectively in the periplasm and shuttle the elements to their respective membrane transport complexes (Hudson and Morel, 1992). The high abundance of proteins involved in ABC transport suggested that nutrient transport rates could limit the amount of Fe and P *Trichodesmium* can acquire. Thus, we explored whether membrane crowding, i.e. lack of membrane space, can constrain nutrient acquisition by *Trichodesmium*.

To investigate this, we quantified the absolute concentration of the phosphate ABC transporter PstA, which interacts with the phosphate stress biomarkers SphX and PstS. This analysis is distinct from the above global metaproteomes, which allowed patterns to be identified but did not allow for absolute quantitation of the proteins. The analysis was performed similar to an isotope dilution experiment where labelled peptide standards are used to control for analytical biases. The analysis was performed for three Tricolim and six JC150 stations. Briefly, ^{15}N labelled peptide standards were prepared and spiked into the samples prior to PRM LC-MS/MS analysis. The concentration of the peptide in fmol μg^{-1} total protein was calculated using the ratio of product ion intensities for the heavy (spike) and light (sample) peptide and converted to PstA molecules per cell (Table 1 and see also Table S4). The peptide used for quantitation of PstA was specific to *Trichodesmium* species. Based on these calculations, on average up to 19 to 36% of the membrane was occupied by the PstA transporter. In one population (JC150 expedition, Station 7), up to 83% of the membrane was occupied by PstA alone. While these are first estimates, it is clear that the majority of *Trichodesmium* cells devoted a large fraction of their membrane surface area to phosphate uptake.

To examine whether membrane crowding can indeed cause nutrient stress or limitation, we developed a model of cellular nutrient uptake in *Trichodesmium*. The model identifies the concentration at which free Fe or phosphate limits the

growth of *Trichodesmium* cells. This is distinct from nutrient stress, which changes the cell's physiological state but does not necessarily impact growth. In the model, nutrient limitation occurs when the daily cellular requirement is greater than the uptake rate, a function of the cell's growth rate and elemental quota. Following the example of Hudson and Morel (1992), the model assumes that intake of nutrients once bound to the ABC transporter protein is instantaneous, i.e. that nutrient uptake is limited by formation of the metal-transporter complex at the cell surface. This is an idealized scenario, because if intake is the slow step, for instance in a high affinity transport system, the uptake rate would be slower and nutrient limitation exacerbated (discussed below).

We considered two types of nutrient limitation in the model (Table S5). First, we considered a diffusion-limited case, in which the rate of uptake is determined by diffusion of the nutrient to the cell's boundary layer ($\mu*Q = {}^{2}/_{3}k_D[\text{nutrient}]$, where μ = the cell growth rate, Q = the cell nutrient quota, and k_D = the diffusion rate constant, dependent on the surface area and diffusion coefficient of the nutrient in seawater). Based on empirical evidence provided by Hudson and Morel (1992), limitation occurs when the cell quota is greater than ${}^{2}/_{3}$ the diffusive-limited flux because beyond this, depletion of the nutrient in the boundary layer occurs. In the second case, membrane crowding limitation, the rate of uptake is determined by the rate of transporter-metal complex formation ($\mu*Q = k_f[\text{transport protein}][\text{nutrient}]$, where k_f = the rate of ligand-nutrient complex formation). Here, up to 50% of the membrane can be occupied by the transport protein following the example of and Morel (1992). This is within the range of the above estimates of membrane occupation by phosphate transporter PstA. The model uses conservative estimates for diffusion coefficients, cell quotas, growth rates, and membrane space occupation to identify the lowest concentration threshold for nutrient limitation; as a result it is likely that *Trichodesmium* becomes limited at higher nutrient concentrations than the model suggests. At this time, the model can only consider labile dissolved Fe and inorganic phosphate, though *Trichodesmium* can also acquire particulate iron, organic phosphorus, phosphite, and phosphonates (Dyhrman et al., 2006; Frischkorn et al., 2018; Polyviou et al., 2015; Poorvin et al., 2004; Rubin et al., 2011).

We first considered a spherical cell, where the surface area: volume quotient decreases as cell radius increases (Figure 7). As the cell grows in size, higher nutrient concentrations are required to sustain growth. This is consistent with the general understanding that larger microbial cells with lower surface area: volume quotient are less competitive in nutrient uptake (Chisholm, 1992; Hudson and Morel, 1992). For a given surface area: volume quotient, the mechanism driving nutrient limitation is whichever model (diffusion or membrane crowding) results in a higher minimum nutrient concentration below which limitation occurs. For a spherical cell, Fe limitation is driven by diffusion when the cell is large and the surface area: volume quotient is low (Figure 7a). However, when cells are smaller and the surface area: volume quotient is high, membrane crowding drives nutrient limitation, meaning that the number of ligands, and not diffusion from the surrounding environment, is the primary control on nutrient uptake. For phosphate, diffusion is almost always the driver of nutrient limitation owing to the higher rate of transporter-nutrient complex formation (k_f) for phosphate, which causes very fast membrane transport rates and relieves membrane-crowding pressures across all cell sizes (Figure 7b) (Froelich et al., 1982).

While this model may be directly applicable to some N_2-fixing cyanobacteria such as Groups B and C, which have roughly spherical cells, *Trichodesmium* cells are not spheres but rather roughly cylindrical (Hynes et al., 2012). Thus, we

repeated the model calculations for cylinders with varying radii (r) and heights (2r or 10r) based on previous estimates of *Trichodesmium* cell sizes (Bergman et al., 2013; Hynes et al., 2012). Cylinders have lower surface area: volume quotient than spheres of similar sizes. In addition, the rate constant (k_D) for diffusion, which is a function of cell geometry, is greater. This increases the slope of the diffusion limitation line such that membrane crowding is important across a greater range of cell sizes (Figure 7c-d). *Trichodesmium* cell sizes vary in nature, for instance the cylinder height can be elongated, improving the surface area: volume quotient. However, the impact of cell elongation to radius r and height 10r on both diffusion limitation and membrane crowding is subtle (Figure 7e-f). Furthermore, though not explicitly considered here, cylindrical cells living in filaments would have reduced surface area available for nutrient uptake. Thus, we conclude that in certain scenarios, lack of membrane space could hypothetically limit Fe and perhaps P acquisition by *Trichodesmium*, particularly when the cells live in filaments or colonies as occurs in nature.

A key assumption of the model is that uptake rates are instantaneous. In the above calculations, we use the dissociation kinetics of Fe from water and phosphorus with common seawater cations as the best case (i.e. fastest possible) kinetic scenario for nutrient acquisition. The model does not account for delays caused by internalization kinetics, which would exacerbate nutrient limitation. For instance it does not consider nutrient speciation, which could affect internalization rates, particularly for Fe (Hudson and Morel, 1992). Furthermore, the involvement of the periplasmic binding proteins IdiA and SphX suggest that uptake is not simultaneous; their participation is likely associated with a kinetic rate of binding and dissociation from the periplasmic proteins in addition to any rate of ABC transport.

Membrane crowding could produce real cellular challenges, leading to the observation of Fe and P co-stress across the field populations examined. The above model explicitly allows 50% of the cell surface area to be occupied by any one type of transporter, consistent with our estimate of cell surface area occupied by the PstA transporter. If 50% of the membrane is occupied by phosphate transporters, and another 50% for Fe transporters, this would leave no room for other essential membrane proteins and even the membrane lipids themselves. The problem is further exacerbated if the cell installs transporters for nitrogen compounds such as urea, as the metaproteomes suggest. Thus, installation of transporters for any one nutrient must be balanced against transporters for other nutrients. This interpretation is inconsistent with Liebig's law of nutrient limitation, which assumes that nutrients are independent (Liebig, 1855; Saito et al., 2008). In an oligotrophic environment, membrane crowding could explicitly link cellular Fe, P, and N uptake status, driving the cell to be co-stressed for multiple nutrients.

3.5 Advantages of the colonial form

Living in a colony has specific advantages and disadvantages for a *Trichodesmium* cell. Colonies may be able to access nutrient sources that would be infeasible for use by single cells or filaments. For instance, *Trichodesmium* colonies have a remarkable ability to entrain dust particles and can move these particles into the center of said colony (Basu et al., 2019; Basu and Shaked, 2018; Poorvin et al., 2004; Rubin et al., 2011). In this study, which focused on *Trichodesmium* colonies, the chemotaxis response regulator CheY was very abundant, particularly in populations sampled near the Amazon

and Orinoco river plumes. CheY was positively correlated with Fe stress biomarker IdiA, but not with phosphate stress biomarker SphX, suggesting that chemotactic movement is involved in entrainment of trace metals including from particulate sources (Figure 8).

The metaproteomes and nutrient uptake model presented in this paper support the growing understanding that *Trichodesmium* must be able to access particulate and organic matter. Living in a colony can be advantageous because such substrates can be concentrated, improving the viability of extracellular nutrient uptake systems. *Trichodesmium's* epibiont community produces siderophores, which assist in Fe uptake, particularly from particulate organic matter (Chappell and Webb, 2010; Lee et al., 2018). Siderophore production is energetically and nutritionally expensive, so it is most advantageous when resource concentrations are high and loss is low, as would occur in the center of a colony (Leventhal et al., 2019). Colonies may similarly enjoy advantages for phosphate acquisition, particularly when the excreted enzyme alkaline phosphatase is utilized to access organic sources (Frischkorn et al., 2018; Elizabeth Duncan Orchard, 2010; Orcutt et al., 2013; Yamaguchi et al., 2016; Yentsch et al., 1970). Additionally, the concentration of cells in a colony means that the products of nitrogen fixation, including urea, can be recycled and are less likely to be lost to the environment. By increasing effective size and concentrating deterrent toxins, colony formation may also protect against grazing (Hawser et al., 1992).

A key hallmark of *Trichodesmium* colony formation is production of mucus, which can capture particulate matter and concentrate it within the colony (Eichner et al., 2019). In addition to particle entrainment, the mucus layer can benefit cells by inhibiting oxygen diffusion, facilitating epibiont associations, regulating buoyancy, defending against grazers and helping to "stick" trichomes together (Eichner et al., 2019; Lee et al., 2017; Sheridan, 2002). However, these benefits come at a cost because the mucus layer hinders diffusion to the cell surface (Figure 9), reducing contact with the surrounding seawater. Despite this, the benefits of colony formation seem to outweigh the costs, since *Trichodesmium* forms colonies in the field, particularly under stress (Bergman et al., 2013; Capone et al., 1997; Hynes et al., 2012).

4. Conclusions

Trichodesmium's colonial lifestyle likely produces challenges for dissolved Fe and P acquisition, which must be compensated for by production of multiple nutrient transport systems, such as for particulate iron and organic phosphorous, at a considerable cost. While laboratory studies have largely focused on single nutrient stresses in free filaments, these metaproteomic observations and accompanying nutrient uptake model demonstrate that Fe and P co-stress is the norm rather than the exception, particularly in the North Atlantic ocean. This means that the emphasis on single limiting nutrients in culture studies and biological models may not capture the complexities of *Trichodesmium's* physiology in situ. Thus, biogeochemical models should consider incorporating Fe and P co-stress conditions. Specifically, in this study and in others there is evidence that nitrogen fixation is optimal under co-limited or co-stressed conditions, implying that an input of either Fe or P could counter-intuitively decrease N_2 driven new production (Garcia et al., 2015; Walworth et al., 2016).

These data demonstrate that *Trichodesmium* cells are confronted by the biophysical limits of membrane space and diffusion rates for their Fe, P, and possibly urea, acquisition systems. This means that there is little room available for

systems that interact with other resources such as light, CO_2, Ni, and other trace metals, providing a mechanism by which nutrient stress could compromise acquisition of other supplies. The cell membrane could be a key link allowing *Trichodesmium* to optimize its physiology in response to multiple environmental stimuli. This is particularly important in an ocean where nutrient availability is sporadic and unpredictable. Future studies should aim to characterize the specific regulatory systems, chemical species and phases (i.e. dissolved versus particulate nutrient sources), and symbiotic interactions that underlie *Trichodesmium's* unique behavior and lifestyle.

Data Availability

All new data is provided in the supplementary material. The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE partner repository with the dataset identifier PXD016225 and 10.6019/PXD016225 (Perez-Riverol et al., 2019).

Supplement

Supplementary information is provided in a separate file (Figure S1, Table S1, Table S2, Table S6), with Tables S3, S4, and S5 provided separately due to their large sizes.

Author Contributions

N.A. and M.S. conceptualized the study. D.H. and E.W. lead the Tricolim expedition. C.M. and M.L. lead the JC150 expedition. N.C., M.W., and K.K. measured nutrient distributions on the Tricolim and JC150 expeditions. D.M. and M.M. helped with proteomics analyses. N.A. prepared the manuscript with contributions from all co-authors.

Competing Interests

The authors declare no competing interests.

Acknowledgements

We acknowledge Elena Cerdan Garcia, Asa Conover, Joanna Harley, Despo Polyviou, and Petroc Shelley for assistance with sampling and nutrient measurements while at sea, in addition to the entire crew of the JC150 and Tricolim expeditions. We thank Ben Van Mooy for insightful discussions regarding this work. This work was supported by a National Science Foundation Graduate Research Fellowship under grant 1122274 [N.Held], Gordon and Betty Moore Foundation grant number 3782 [M.Saito], and National Science Foundation grants OCE-1657755 and EarthCube-1639714 [M.Saito]. We also acknowledge funding from the UK Natural Environment Research Council (NERC) grants awarded to CM (NE/N001079/1) and ML (NE/N001125/1). NRC was supported by grant 544236 from the Simons Foundation.

References

Basu, S., Gledhill, M., de Beer, D., Prabhu Matondkar, S. G., and Shaked, Y.: Colonies of marine cyanobacteria *Trichodesmium* interact with associated bacteria to acquire iron from dust, Communications Biology, 2(1), 1–8, https://doi.org/10.1038/s42003-019-0534-z, 2019.

Basu, S., and Shaked, Y.: Mineral iron utilization by natural and cultured *Trichodesmium* and associated bacteria, Limnology and Oceanography, 63(6), 2307–2320, https://doi.org/10.1002/lno.10939, 2018.

Bergman, B., Sandh, G., Lin, S., Larsson, J., and Carpenter, E. J.: *Trichodesmium*--a widespread marine cyanobacterium with unusual nitrogen fixation properties, FEMS Microbiology Reviews, 37(3), 286–302, https://doi.org/10.1111/j.1574-6976.2012.00352.x, 2013.

Capone, Douglas G., Zehr, J. P., Paerl, H. W., Bergman, B., and Carpenter, E. J.: *Trichodesmium*, a globally significant marine cyanobacterium, Science, 276(5316), 1221–1229, https://doi.org/10.1126/science.276.5316.1221, 1997.

Carpenter, E J, Bergman, B., Dawson, R., Siddiqui, P. J., Söderbäck, E., and Capone, D. G.: Glutamine synthetase and nitrogen cycling in colonies of the marine diazotrophic cyanobacteria *Trichodesmium* spp., Applied and Environmental Microbiology, 58(9), 3122–3129, 1992.

Carpenter, Edward J, and Romans, K.: Major Role of the Cyanobacterium *Trichodesmium* in Nutrient Cycling in the North Atlantic Ocean, Science, 254(1989), 1989–1992., 1991.

Chappell, P. Dreux, and Webb, E. A.: A molecular assessment of the iron stress response in the two phylogenetic clades of *Trichodesmium*, Environmental Microbiology, 12(1), 13–27, https://doi.org/10.1111/j.1462-2920.2009.02026.x, 2010.

Chappell, Phoebe Dreux, Moffett, J. W., Hynes, A. M., and Webb, E. A.: Molecular evidence of iron limitation and availability in the global diazotroph *Trichodesmium*, The ISME Journal, 6(9), 1728–1739, https://doi.org/10.1038/ismej.2012.13, 2012.

Chisholm, S. W.: Phytoplankton Size, in: Primary Productivity and Biogeochemical Cycles in the Sea, edited by: Falkowski P.G., Woodhead A.D., and Vivirito, K., Springer, Boston, US, 213-237, https://doi.org/10.1007/978-1-4899-0762-2_12, 1992.

Coles, V. J., Hood, R. R., Pascual, M., and Capone, D. G.: Modeling the impact of *Trichodesmium* and nitrogen fixation in the Atlantic ocean, Journal of Geophysical Research : Oceans, 109(6), 1–17, https://doi.org/10.1029/2002JC001754, 2004.

Deutsch, C., Sarmiento, J. L., Sigman, D. M., Gruber, N., and Dunne, J. P.: Spatial coupling of nitrogen inputs and losses in the ocean, Nature, 445, 163–167, https://doi.org/10.1038/nature05392, 2007.

Dyhrman, S. T., Chappell, P. D., Haley, S. T., Moffett, J. W., Orchard, E. D., Waterbury, J. B., and Webb, E. A.: Phosphonate utilization by the globally important marine diazotroph *Trichodesmium*, Nature, 439(7072), 68–71, https://doi.org/10.1038/nature04203, 2006.

Eichner, M., Thoms, S., Rost, B., Mohr, W., Ahmerkamp, S., Ploug, H., Kuypers, M.M.M., de Beer, D.: N_2 fixation in free-floating filaments of *Trichodesmium* is higher than in transiently suboxic colony microenvironments, New Phytologist, 222(2), 852–863, https://doi.org/10.1111/nph.15621, 2019.

Eng, J. K., Fischer, B., Grossmann, J., and MacCoss, M. J.: A fast SEQUEST cross correlation algorithm, Journal of Proteome Research, 7(10), 4598–4602, https://doi.org/10.1021/pr800420s, 2008.

Flores, E., and Herrero, A.: Nitrogen assimilation and nitrogen control in cyanobacteria, Biochemical Society Transactions, 33(1), 164–167, https://doi.org/10.1042/BST0330164, 2005.

Forchhammer, K., and De Marsac, N. T.: The P(II) protein in the cyanobacterium Synechococcus sp. strain PCC 7942 is modified by serine phosphorylation and signals the cellular N-status, Journal of Bacteriology, 176(1), 84–91, 1994.

Frischkorn, K. R., Krupke, A., Guieu, C., Louis, J., and Rouco, M.: *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, 2018.

Fu, Q., Chen, Z., Zhang, S., Parker, S. J., Fu, Z., Tin, A., Liu, X., Van Eyk, J. E.: Multiple and Selective Reaction Monitoring Using Triple Quadrupole Mass Spectrometer: Preclinical Large Cohort Analysis, Quantitative Proteomics by Mass Spectrometry (Vol. 1410), https://doi.org/10.1007/978-1-4939-3524-6_15, 2016.

Garcia, N. S., Fu, F., Sedwick, P. N., and Hutchins, D. A.: Iron deficiency increases growth and nitrogen-fixation rates of phosphorus-deficient marine cyanobacteria, The ISME Journal, 9(1), 238–245, https://doi.org/10.1038/ismej.2014.104, 2015.

Hawser, S. P., O'Neil, J. M., Roman, M. R., and Codd, G. A.: Toxicity of blooms of the cyanobacterium *Trichodesmium* to zooplankton, Journal of Applied Phycology, 4(1), 79–86, https://doi.org/10.1007/BF00003963, 1992.

Held, N. A., Mcilvin, M. R., Moran, D. M., Laub, M. T., and Saito, M. A.: Unique Patterns and Biogeochemical Relevance of Two-Component Sensing in Marine Bacteria, MSystems, 1–16, https://doi.org/10.1128/mSystems.00317-18, 2019.

Hudson, R.J., Morel, F. M. M.: Trace metal transport by marine microorganisms: implications of metal coordination kinetics, Deep Sea Research Part I: Oceanographic Research Papers, 40(1), 129–150, 1992.

Hynes, A. M., Chappell, P. D., Dyhrman, S. T., Doney, S. C., and Webb, E. A.: Cross-basin comparison of phosphorus stress and nitrogen fixation in *Trichodesmium*, Limnology and Oceanography, 54(5), 1438–1448, https://doi.org/10.4319/lo.2009.54.5.1438, 2009.

Hynes, A. M., Webb, E. A., Doney, S. C., and Waterbury, J. B.: Comparison of cultured *Trichodesmium* (Cyanophyceae) with species characterized from the field, J.Phycol, 48, 196–210, https://doi.org/10.1111/j.1529-8817.2011.01096.x, 2012.

Kunde, K., Wyatt, N. J., González-Santana, D., Tagliabue, A., Mahaffey, C., and Lohan, M. C.: Iron Distribution in the Subtropical North Atlantic: The Pivotal Role of Colloidal Iron, Global Biogeochemical Cycles, 2019GB006326, https://doi.org/10.1029/2019GB006326, 2019.

Karl, D., Michaels, A., Bergman, B., Capone, D., Carpenter, E., Letelier, R., Lipschultz, F., Paerl, H., Sigman, D., Stal, L.: Dinitrogen fixation in the world's oceans, Biogeochemistry, 57–58, 47 98, https://doi.org/10.1023/A:1015798105851, 2002.

Küpper, H., Šetlík, I., Seibert, S., Prášil, O., Šetlikova, E., Strittmatter, M., Levitan, O., Lohscheider, J., Adamska I.,

Berman-Frank, I.: Iron limitation in the marine cyanobacterium *Trichodesmium* reveals new insights into regulation of photosynthesis and nitrogen fixation, New Phytologist, 179, 784–798, 2008.

Lee, M. D., Webb, E. A., Walworth, N. G., Fu, F. X., Held, N. A., Saito, M. A., and Hutchins, D. A.: Transcriptional activities of the microbial consortium living with the marine nitrogenfixing cyanobacterium *Trichodesmium* reveal potential roles in community-level nitrogen cycling, Applied and Environmental Microbiology, 84(1), https://doi.org/10.1128/AEM.02026-17, 2018.

Leventhal, G. E., Ackermann, M., and Schiessl, K. T.: Why microbes secrete molecules to modify their environment: The case of iron-chelating siderophores, Journal of the Royal Society Interface, 16(150), https://doi.org/10.1098/rsif.2018.0674, 2019.

Liebig, J. V.: Principles of agricultural chemistry with special reference to the late researches made in England. Dowden, Hutchinson, and Ross., 1855.

Lu, X., and Zhu, H.: Tube-Gel Digestion: A Novel Proteomic Approach for High Throughput Analysis of Membrane Proteins, Mol Cell Proteomics, 4(12), 1948–1958, https://doi.org/10.1074/mcp.M500138-MCP200, 2005.

McGillicuddy, D. J.: Do *Trichodesmium* spp. populations in the North Atlantic export most of the nitrogen they fix?, Global Biogeochemical Cycles, 28, 103–114, https://doi.org/10.1002/2013GB004652, 2014.

Mills, M. M., Ridame, C., Davey, M., La Roche, J., and Geider, R. J.: Iron and phosphorus co-limit nitrogen fixation in the eastern tropical North Atlantic, Nature, 429(6989), 292–294. https://doi.org/10.1038/nature02550, 2004

Milo, R.: What is the total number of protein molecules per cell volume? A call to rethink some published values, BioEssays, 35(12), 1050–1055, https://doi.org/10.1002/bies.201300066, 2013.

Mulholland, M. R., and Capone, D. G.: Nitrogen utilization and metabolism relative to patterns of N_2 fixation in populations of *Trichodesmium* from the North Atlantic Ocean and Caribbean Sea, Marine Ecology Progress Series, 188, 33–49, 1999.

Ohki, K., Zehr, J. P., Falkowski, P. G., and Fujita, Y.: Regulation of nitrogen-fixation by different nitrogen sources in the marine non-heterocystous cyanobacterium *Trichodesmium* sp. NIBB1067, Archives of Microbiology, 156(5), 335–337, https://doi.org/10.1007/BF00248706, 1991.

Orchard, Elizabeth D., Webb, E. A., and Dyhrman, S. T.: Molecular analysis of the phosphorus starvation response in *Trichodesmium* spp., Environmental Microbiology, 11(9), 2400–2411, https://doi.org/10.1111/j.1462-2920.2009.01968.x, 2009.

Orchard, Elizabeth Duncan.: Phosphorus physiology of the marine Cyanobacterium *Trichodesmium*, Massachusetts Institute of Technology, 130, https://doi.org/10.1575/1912/3366, 2010.

Orcutt, K. M., Gundersen, K., and Ammerman, J. W.: Intense ectoenzyme activities associated with *Trichodesmium* colonies in the Sargasso Sea, Marine Ecology Progress Series, 478(March), 101–113. https://doi.org/10.3354/meps10153, 2013

P.N. Froelich, M.L. Bender, N.A. Luedtke, G.R. Heath, T. D.: The Marine Phosphorus Cycle, American Journal of Science, 282, 464–511, 1982.

Perez-Riverol, Y., Csordas, A., Bai, J., Bernal-Llinares, M., Hewapathirana, S., Kundu, D. J. Inuganti, A., Griss, J., Mayer G., Eisenacher, M., Perez, E., Uszkoriet, J., Pfeuffer, J., Sachsenberg, T., Yilmaz, S., Tiwary, S., Cox, J., Audain, E.,Walzer, M., Januczak, A.F. Ternent, T., Brazma, A., Vizcaíno, J. A.: The PRIDE database and related tools and resources in 2019: Improving support for quantification data, Nucleic Acids Research, 47(D1), D442–D450, https://doi.org/10.1093/nar/gky1106, 2019.

Polyviou, D., Hitchcock, A., Baylay, A. J., Moore, C. M., and Bibby, T. S.: Phosphite utilization by the globally important marine diazotroph *Trichodesmium*, Environmental Microbiology Reports, 7(6), 824–830, https://doi.org/10.1111/1758-2229.12308, 2015.

Poorvin, L., Rinta-kanto, J. M., Hutchins, D. A., and Wilhelm, S. W.: Viral release of iron and its bioavailability to marine plankton, Limnol. Oceanogr 49(5), 1734–1741, 2004.

Reddy, R. J., Gajadhar, A. S., Swenson, E. J., Rothenberg, D. A., Curran, T. G., and White, F. M. Early signaling dynamics of the epidermal growth factor receptor, Proceedings of the National Academy of Sciences of the United States of America, 113(11), 201521288, https://doi.org/10.1073/pnas.1521288113, 2016.

Rouco, M., Frischkorn, K. R., Haley, S. T., and Alexander, H.: Transcriptional patterns identify resource controls on the diazotroph *Trichodesmium* in the Atlantic and Paci fi c oceans, The ISME Journal, 1486–1495, https://doi.org/10.1038/s41396-018-0087-z, 2018.

Rouwenhorst, R. J., Frank Jzn, J., Scheffers, W. A., and van Dijken, J. P.: Determination of protein concentration by total organic carbon analysis, Journal of Biochemical and Biophysical Methods, 22(2), 119–128, https://doi.org/10.1016/0165-022X(91)90024-Q, 1991.

Rubin, M., Berman-Frank, I., and Shaked, Y.: Dust-and mineral-iron utilization by the marine dinitrogen-fixer *Trichodesmium*, Nature Geoscience, 4(8), 529–534, https://doi.org/10.1038/ngeo1181, 2011.

Saito, M. A.., Dorsk, A., Post, A. F., McIlvin, M. R., Rappé, M. S., DiTullio, G. R., and Moran, D. M.: Needles in the blue sea: Sub-species specificity in targeted protein biomarker analyses within the vast oceanic microbial metaproteome, Proteomics, https://doi.org/10.1002/pmic.201400630, 2015.

Saito, M. A.., Goepfert, T. J., and Ritt, J. T.: Some thoughts on the concept of colimitation: Three definitions and the importance of bioavailability, Limnology and Oceanography, 53(1), 276–290, https://doi.org/10.4319/lo.2008.53.1.0276, 2008.

Saito, M. A., McIlvin, M. R., Moran, D. M., Goepfert, T. J., DiTullio, G. R., Post, A. F., and Lamborg, C. H.: Multiple nutrient stresses at intersecting Pacific Ocean biomes detected by protein biomarkers, Science (New York, N.Y.), 345(6201), 1173–1177, https://doi.org/10.1126/science.1256450, 2014.

Sañudo-Wilhelmy, S. A., Kustka, A. B., Gobler, C. J., Hutchins, D. a., Yang, M., Lwiza, K., Burns, J., Capone, D.G., Raven, J.A., Carpenter, E. J.: Phosphorus limitation of nitrogen fixation by *Trichodesmium* in the central Atlantic Ocean, Nature, 411(6833), 66–69, https://doi.org/10.1038/35075041, 2001.

Sheridan, C. C.: The microbial and metazoan community associated with colonies of *Trichodesmium* spp.: a quantitative

survey. Journal of Plankton Research, 24(9), 913–922, https://doi.org/10.1093/plankt/24.9.913, 2002.

Shi T, Sun Y & Falkowski P.G.: Effects of iron limitation on the expression of metabolic genes in the marine cyanobacterium *Trichodesmium erythraeum* IMS101. Environ Microbiol 9: 2945–2956, 2007.

Shih, P. M., Wu, D., Latifi, A., Axen, S. D., Fewer, D. P., Talla, E., Calteau, A., Cai, F., Tandeau de Marsac, N., Rippka, R., Herdman, M., Sivonen, K., Coursin, T, Laurent, T., Goodwin, L., Nolan, M., Davenport, K.W., Han, C.S., Rubin, E.M., Eisen, J.A., Woyke, T., Gugger, M., Kerfeld, C. A.: Improving the coverage of the cyanobacterial phylum using diversity-driven genome sequencing, Proceedings of the National Academy of Sciences of the United States of America, 110(3), 1053–1058, https://doi.org/10.1073/pnas.1217107110, 2013.

Snow, J. T., Polyviou, D., Skipp, P., Chrismas, N. A. M., Hitchcock, A., Geider, R., Moore, C.M., Bibby, T. S.: Quantifying integrated proteomic responses to iron stress in the globally important marine diazotroph *Trichodesmium*, PLoS ONE, 10(11), 1–24, https://doi.org/10.1371/journal.pone.0142626, 2015.

Sohm, J. A., Webb, E. A., and Capone, D. G.: Emerging patterns of marine nitrogen fixation, Nature Reviews Microbiology, 9(7), 499–508, https://doi.org/10.1038/nrmicro2594, 2011.

Strathmann, R.R.: Estimating Organic Carbon Content of Phytoplankton from Cell Volume or Plasma Volume, Limnology and Oceanography, 12(3), 411-418, 1967.

Sunda, W. G.: Feedback interactions between trace metal nutrients and phytoplankton in the ocean, Frontiers in Microbiology, 3, 1–22, https://doi.org/10.3389/fmicb.2012.00204, 2012.

Walworth, N. G., Fu, F.-X., Webb, E. A., Saito, M. A., Moran, D., Mcllvin, M. R., Lee, M.D., Hutchins, D. A.: Mechanisms of increased *Trichodesmium* fitness under iron and phosphorus co-limitation in the present and future ocean, Nat Commun, 7(May), 1–11, https://doi.org/10.1038/ncomms12081, 2016.

Walworth, N. G., Fu, F. X., Lee, M. D., Cai, X., Saito, M. A., Webb, E. A., and Hutchins, D. A.: Nutrient-colimited *Trichodesmium* as a nitrogen source or sink in a future ocean, Applied and Environmental Microbiology, 84(3), 1–14, https://doi.org/10.1128/AEM.02137-17, 2018.

Walworth, N. G., Lee, M. D., Fu, F.-X., Hutchins, D. A., and Webb, E. A.: Molecular and physiological evidence of genetic assimilation to high CO_2 in the marine nitrogen fixer *Trichodesmium*, Proceedings of the National Academy of Sciences, (3), 201605202, https://doi.org/10.1073/pnas.1605202113, 2016.

Wang, Q., Li, H., and Post, A. F.: Nitrate assimilation genes of the marine diazotrophic, filamentous cyanobacterium *Trichodesmium* sp. strain WH9601, Journal of Bacteriology, 182(6), 1764–1767, https://doi.org/10.1128/JB.182.6.1764-1767. 2000.

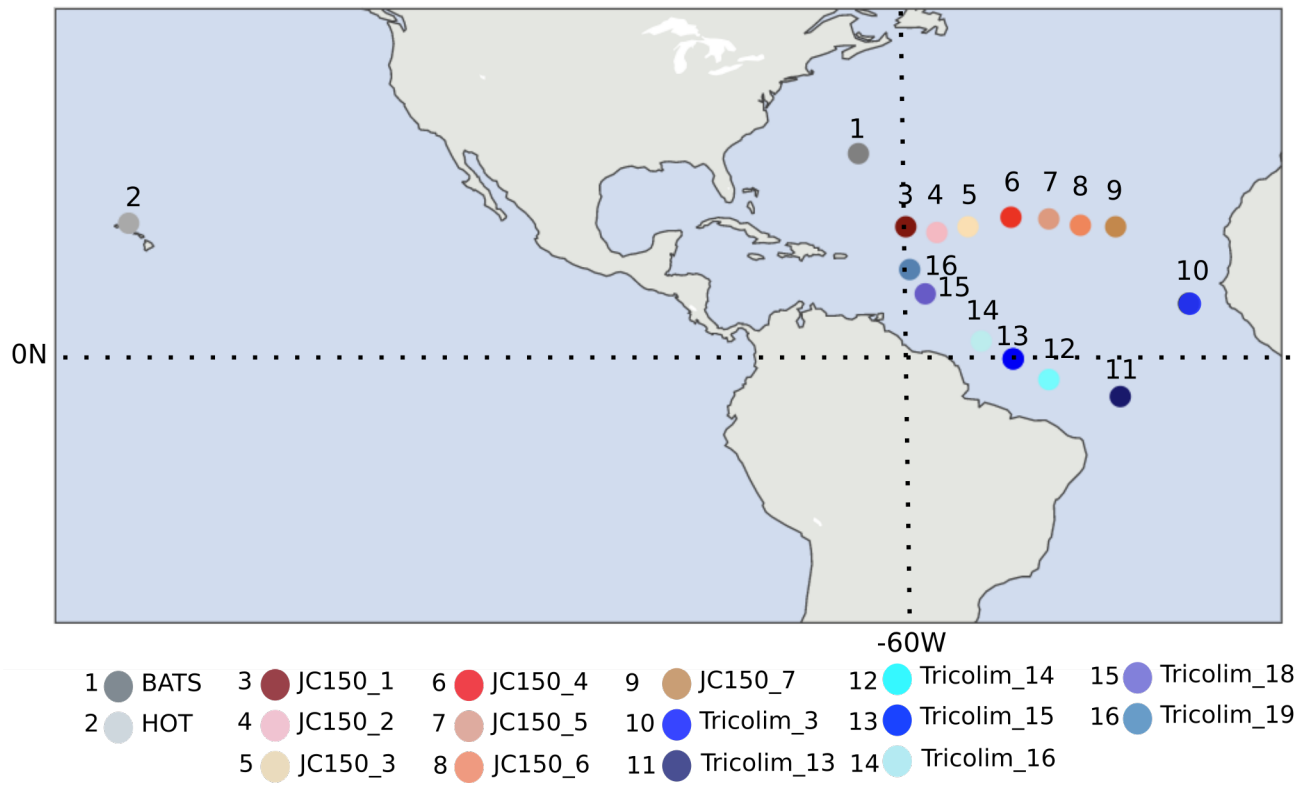
Webb, E. A., Jakuba, R. W., Moffett, J. W., and Dyhrman, S. T.: Molecular assessment of phosphorus and iron physiology in *Trichodesmium* populations from the western Central and western South Atlantic, Limnology and Oceanography, 52(5), 2221–2232, https://doi.org/10.4319/lo.2007.52.5.2221, 2007.

Wu, J. F., Sunda, W., Boyle, E. A., and Karl, D. M.: Phosphate depletion in the western North Atlantic Ocean, Science, 289(5480), 759–762, https://doi.org/Doi 10.1126/Science.289.5480.759, 2000.

Yamaguchi, T., Furuya, K., Sato, M., Takahashi, K.: Phosphate release due to excess alkaline phosphatase activity in *Trichodesmium erythraeum*, Plankton Benthos Res, 11(1), 29–36, https://doi.org/10.3800/pbr.11.29, 2016.

Yentsch, C.,M., Yentsch, C.S. , Perras, J.P.: Alkaline phosphatase activity in the tropical marine blue-green alga *Trichodesmium*, Limnol. Oceanogr., 17, 72–774, https://doi.org/10.4319/lo.1972.17.5.0772, 1972.

Figures



1160

Figure 1. Sampling locations. Red/pink colors indicate JC150 stations; blue colors indicate Tricolim stations, dark grey indicates the Bermuda Atlantic Time Series (BATS) and light grey indicates Hawaii Ocean Time Series (HOT). Most samples exist in duplicate or triplicate; see Table S2 for detailed information.

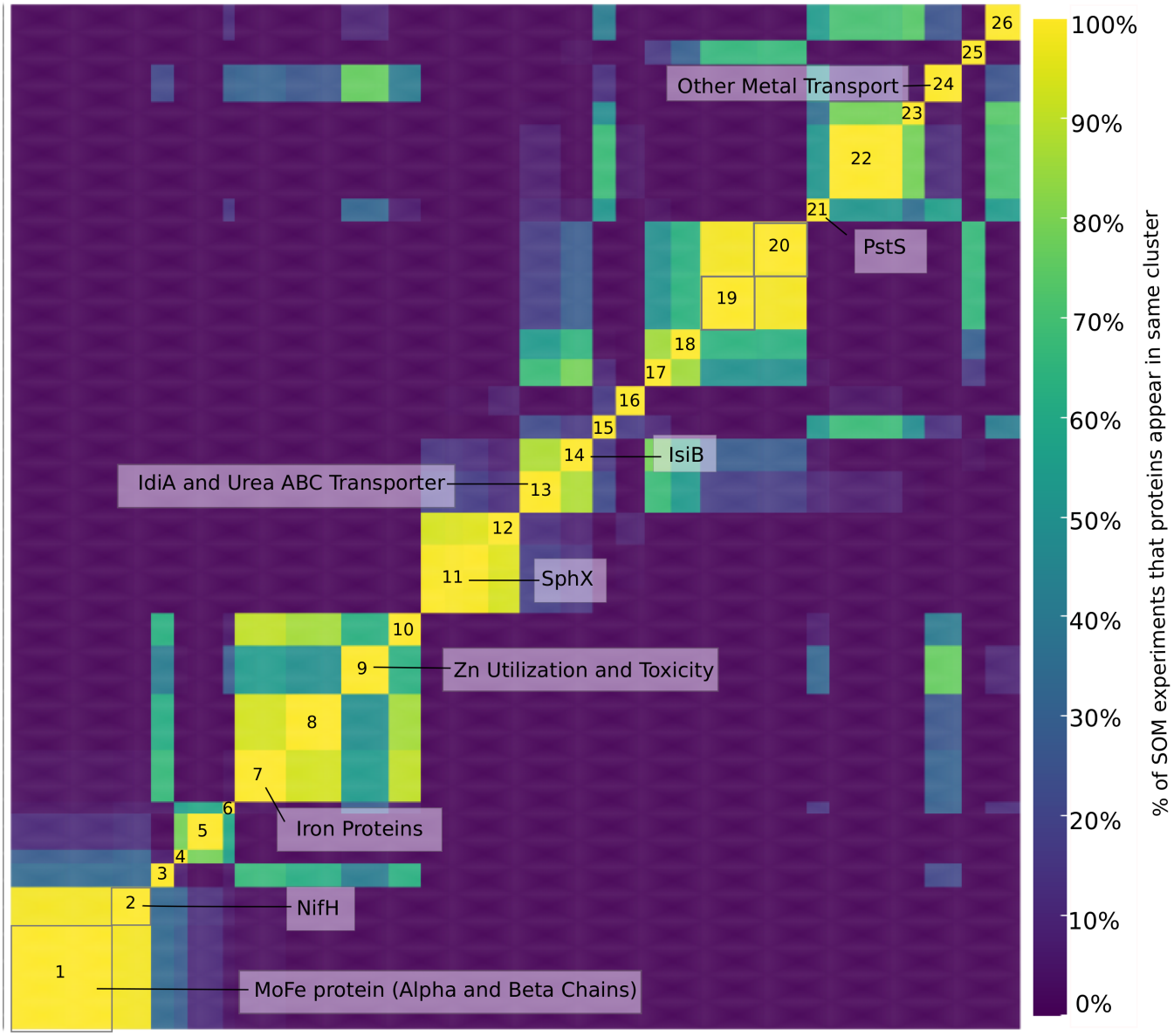


Figure 2. Heatmap displaying results of self-organizing map analysis. Each protein was mapped to a self-organizing map grid, and the grids subsequently clustered by a k-means clustering algorithm. The process was repeated 10,000 times and the results displayed here as a heatmap with warm colors representing proteins that appear in the same cluster. The color bar indicates the percent of SOM experiments in which two proteins appear in the same cluster. Only the top 500 most abundant proteins are displayed. Dark yellow indicates proteins that appear in the same cluster 99.99% of the time. Clusters # 1 and 2 contain nitrogen fixation, carbon fixation, and nitrogen assimilation

proteins as well as the regulatory systems NtcA and P-II. The cluster assignments for the proteins are available in Table S4.

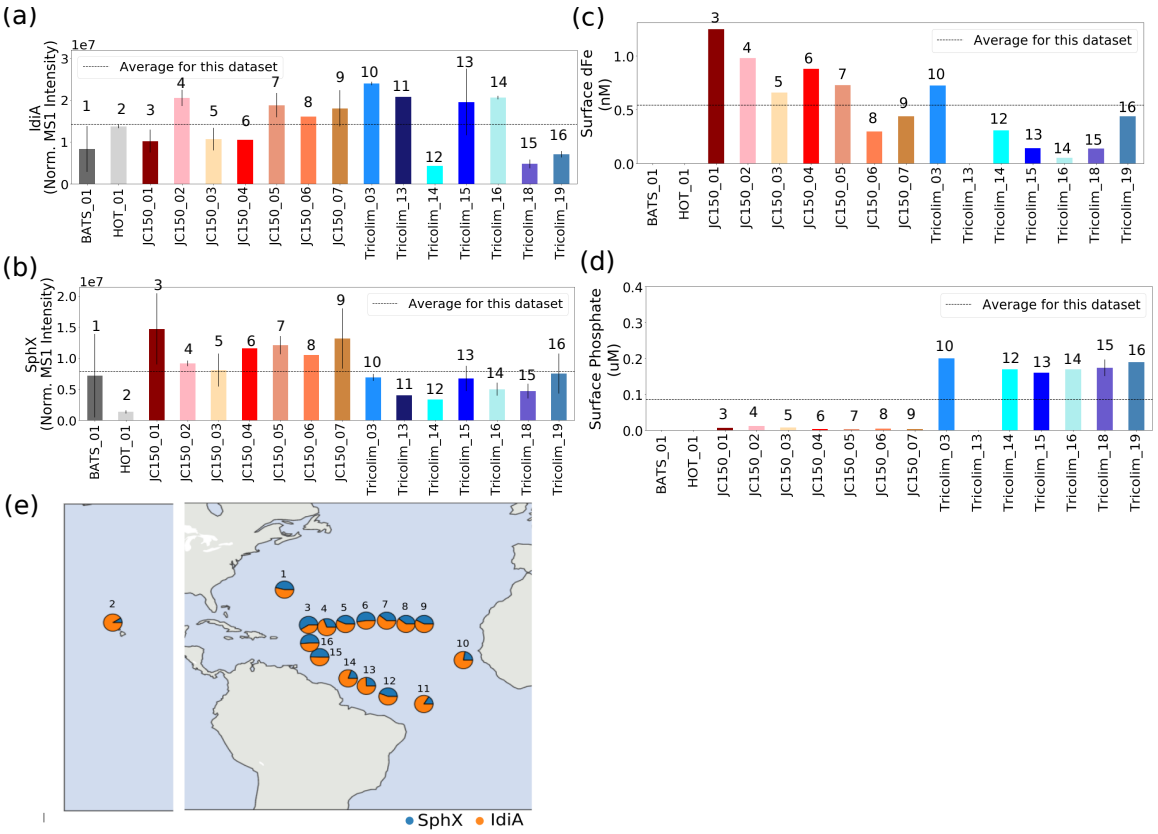


Figure 3. (A) Relative abundance of iron stress protein IdiA (A) and phosphate stress protein SphX (B). IdiA and SphX were among the most abundant proteins in the entire dataset. Error bars are one standard deviation from the mean. Dashed lines represent average values across the dataset. Proteins abundances were normalized such that the total MS1 peak area across the entire proteome was the same for each sample. (C) and (D) concentrations of dFe and dissolved phosphate nutrients. (E) Relative abundance of IdiA (orange) and SphX (blue) overlaid on the sampling locations.

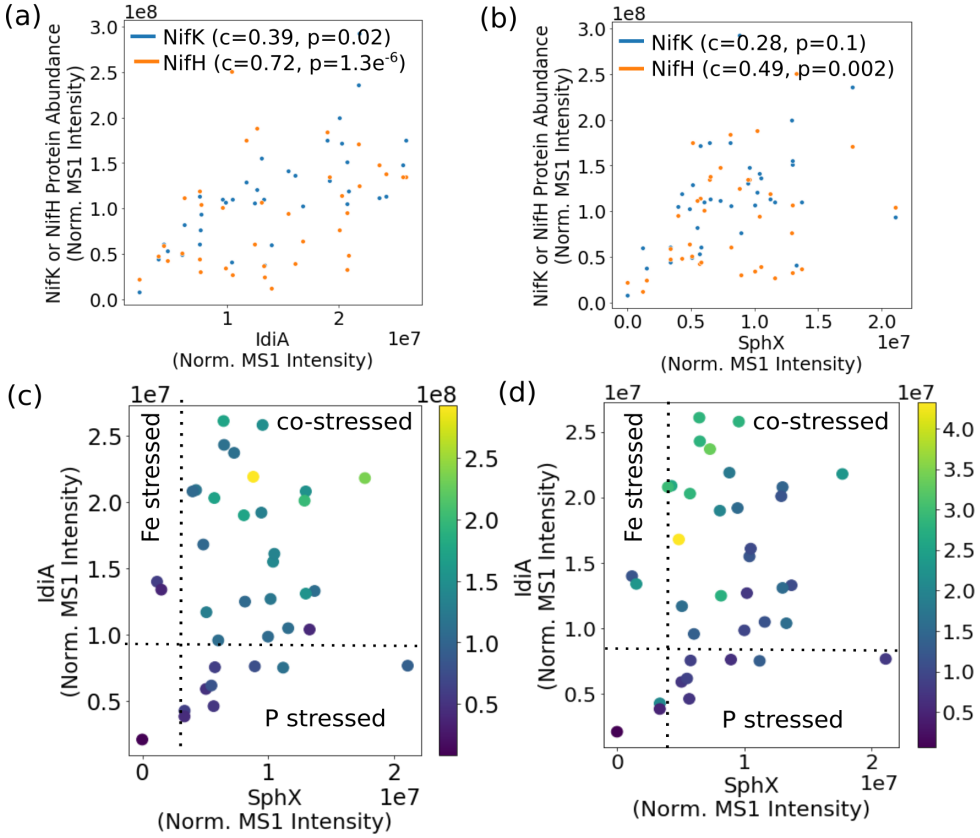


Figure 4. Nitrogenase abundance is highest at the intersection of high iron and phosphate stress. A) IdiA and B) SphX abundance is positively related to nitrogenase MoFe and Fe protein abundance (c = Spearman rank-order correlation coefficient, p = Spearman p-value). Effects of combined iron and phosphate stress biomarkers on nitrogenase abundance. Marker colors represent abundance of NifK (panel C) and NifH (panel D).

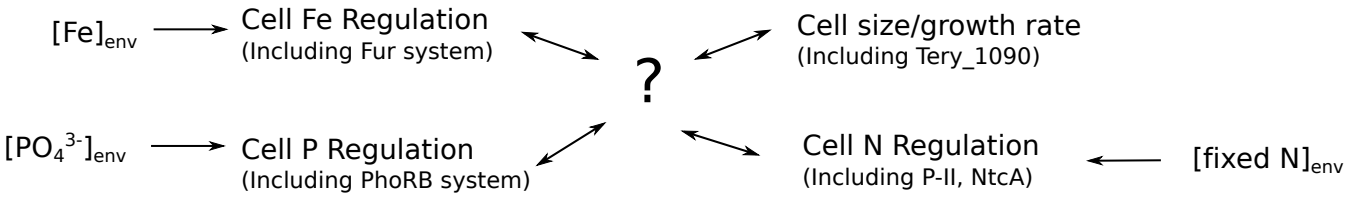
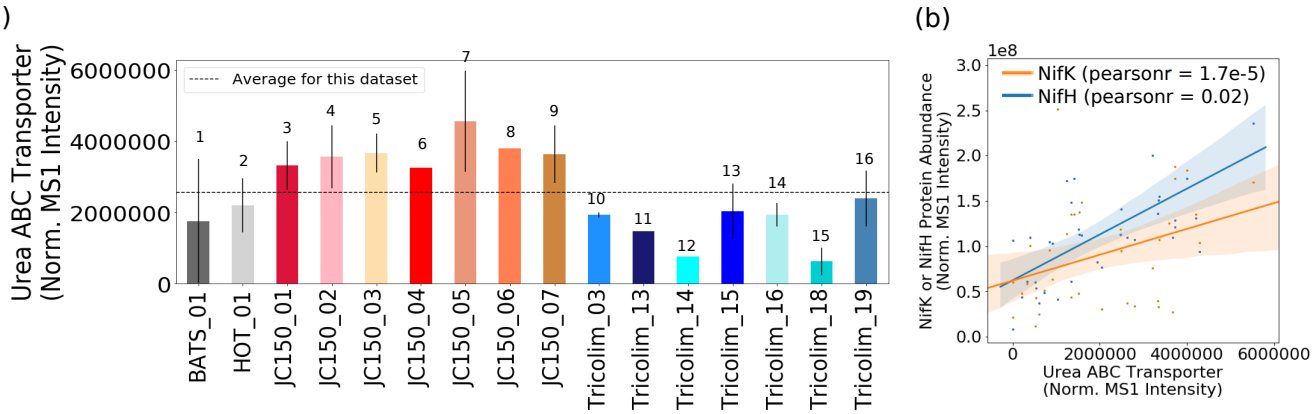


Figure 5. The metaproteomes suggest that there is a currently unknown regulatory link between cellular Fe, P, and N regulation. Key: Fur = ferric uptake regulator, PhoRB = phosphate two component sensory system, Tery_1090 = putative cell size regulator, P-II/NtcA = nitrogen regulatory proteins.



1205

Figure 6. A) Relative abundance of the *Trichodesmium* urea ABC transporter. B) The abundance of the urea ABC transporter is positively correlated with NifH and NifK abundance. Pearson linear correlation coefficients (r values) are provided (p value for NifK = $1.7e^{-5}$, NifH = 0.02). Shaded bars indicate 95% confidence intervals.

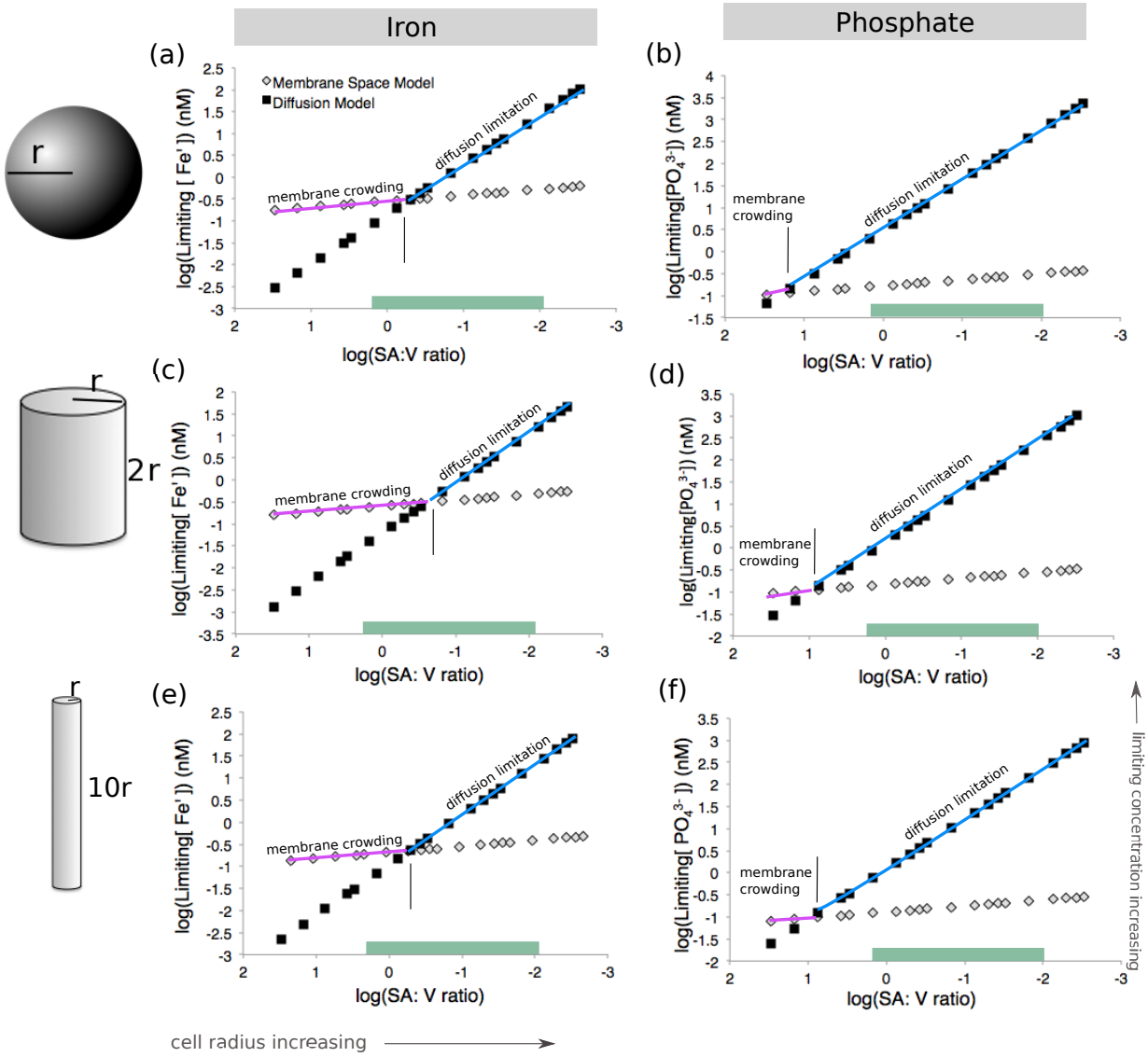


Figure 7. Model calculations for membrane space and diffusion based nutrient limitation reveal that membrane crowding could drive *Trichodesmium* to iron or phosphate stress, particularly when cells are small. Two cell morphologies (sphere and cylinder) were modeled for both iron and phosphate limitation. Calculations are detailed in Table S5. As the cell radius increases and the surface area: volume quotient decreases, the limiting concentration increases. This is concurrent with the current understanding that the low surface area: volume quotient of large cells leads to limitation. Green bars represent common SA: V ratio quotients for *T. Theibautii*.(Hynes et al., 2012) **(A-B). Membrane crowding (purple) occurs if the limiting nutrient concentration is greater than in the diffusion limitation model (blue). Membrane crowding is more significant for cylindrical cells in particular (C-D); altering the length of the cylinder minimally affects the model (E-F).**

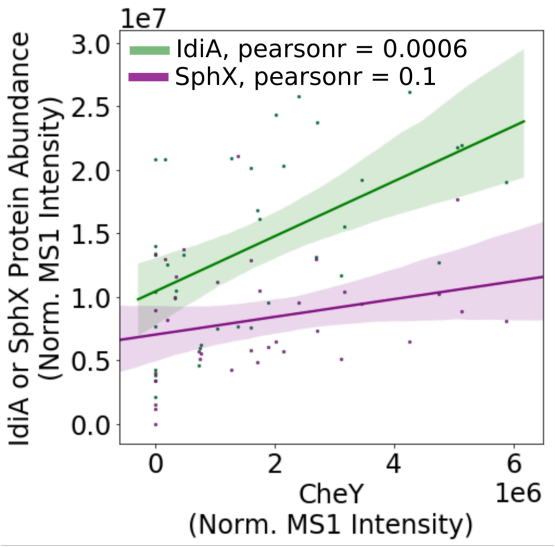


Figure 8. CheY is positively correlated with the iron stress biomarker IdiA, but has a weaker association with phosphate stress biomarker SphX. This suggests that it might be involved in iron acquisition, for instance by helping colonies to move dust particles to the colony center. Pearson linear correlation coefficients (r values) are provided (p value for IdiA = $6e^{-4}$, SphX = 0.1). Shaded bars indicate 95% confidence intervals.

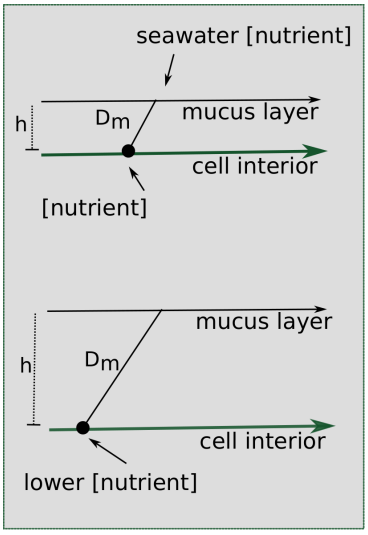


Figure 9. Scheme for the effect of a mucus layer on nutrient diffusion. h = height of the mucus membrane, D_m = diffusion coefficient of the mucus. Assuming some diffusion constant for the nutrient through the mucus and the same starting seawater nutrient concentration, a thicker layer of mucus surrounding a cell in a colony would result in a lower concentration of nutrient experienced at the cell surface.

1235 **Table 1. Quantification of the PstA ABC transporter and estimation of membrane space occupied**

Station	[Pst] in fmol/ug total protein (replicate average)	standard deviation replicates (if available)	Pst molecules per cell assuming 30% w/w protein content*	% surface area occupied assuming 30% w/w^	Pst molecules per cell assuming 55% w/w protein content**	% surface area occupied assuming 55% w/w^
Tricolim_18	13.0	1.8	3.8E+05	3.6	7.0E+05	6.6
Tricolim_15	11.2	3.4	3.3E+05	3.1	6.1E+05	5.7
Tricolim_16	89.1	123.1	2.6E+06	24.5	4.8E+06	45.0
JC150_3	38.7	63.3	1.1E+06	10.7	2.1E+06	19.5
JC150_4	89.6	14.7	2.6E+06	24.7	4.9E+06	45.2
JC150_5	74.2	36.4	2.2E+06	20.4	4.0E+06	37.5
JC150_6	61.6	40.1	1.8E+06	17.0	3.3E+06	31.1
JC150_7	165.7		4.9E+06	45.6	9.0E+06	83.6
JC150_1	106.1		3.1E+06	29.2	5.7E+06	53.5
average				19.9		36.4
stdev				13.4		24.6

* calculated using *Trichodesmium* cell volume of 3000um3 (Berman-Frank et al., 2001), cell volume to carbon conversion logC = 0.716log(V)-0.314 (Strathman, 1967), protein content of a cyanobacterium 30% w/w (Gonzalez Lopez et al., 2010), carbon to total protein conversion 0.53 g C/ g total protein (Rowenhoerst et al., 1991). **calculated as in (*) but with protein content of a cyanobacterium 55% w/w (Gonzalez Lopez et al., 2010). ^calculated using cross sectional area of an Ca ATPase of 0.0000167 um2 (Hudson and Morel 1992)