

# Transcriptomic and proteomic responses of the oceanic diatom *Pseudo-nitzschia granii* to iron limitation

Natalie R Cohen <sup>1,2</sup> Weida Gong <sup>1</sup>  
Dawn M. Moran,<sup>2</sup> Matthew R. McIlvin,<sup>2</sup> Mak A. Saito<sup>2</sup>  
and Adrian Marchetti <sup>1\*</sup>

<sup>1</sup>Department of Marine Sciences, University of North Carolina at Chapel Hill, Chapel Hill, NC, 27514, USA.

<sup>2</sup>Marine Chemistry and Geochemistry Department, Woods Hole Oceanographic Institution, Woods Hole, MA, 02543, USA.

## Summary

Diatoms are a highly successful group of photosynthetic protists that often thrive under adverse environmental conditions. Members of the genus *Pseudo-nitzschia* are ecologically important diatoms which are able to subsist during periods of chronic iron limitation and form dense blooms following iron fertilization events. The cellular strategies within diatoms that orchestrate these physiological responses to variable iron concentrations remain largely uncharacterized. Using a combined transcriptomic and proteomic approach, we explore the exceptional ability of a diatom isolated from the iron-limited Northeast Pacific Ocean to reorganize its intracellular processes as a function of iron. We compared the molecular responses of *Pseudo-nitzschia granii* observed under iron-replete and iron-limited growth conditions to those of other model diatoms. Iron-coordinated molecular responses demonstrated some agreement between gene expression and protein abundance, including iron-starvation-induced-proteins, a putative iron transport system and components of photosynthesis and the Calvin cycle. *Pseudo-nitzschia granii* distinctly differentially expresses genes encoding proteins involved in iron-independent photosynthetic electron transport, urea acquisition and vitamin synthesis. We show that *P. granii* is unique among studied diatoms in its physiology stemming from distinct cellular responses, which may underlie its ability to subsist

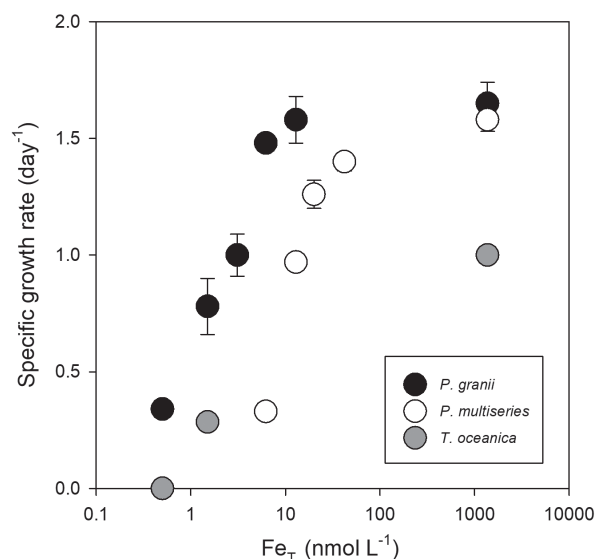
in low iron regions and rapidly bloom to outcompete other diatom taxa following iron enrichment.

## Introduction

In the coastal Northeast (NE) Pacific Ocean, high phytoplankton biomass is fueled by upwelling of nutrient-rich deep ocean waters as well as the merging of complementary water masses where coastal nitrate-poor, iron-rich waters meet oceanic nitrate-rich, iron-poor waters (Ribalet *et al.*, 2010). Further offshore, in the middle of the subpolar Alaskan gyre, chronic iron limitation of phytoplankton growth is well-characterized with occasional blooms occurring when ephemeral iron enters surface waters via atmospheric deposition of continental dust or volcanic activity (Lam *et al.*, 2006; Lam and Bishop, 2008; Olgun *et al.* 2013). Iron fertilization experiments have demonstrated that the introduction of bioavailable iron to these high-nutrient, low-chlorophyll (HNLC) waters creates diatom blooms, often dominated by the pennate diatom genus *Pseudo-nitzschia*, with the potential to transport substantial carbon to the ocean's interior (de Baar *et al.*, 2005; Boyd *et al.*, 2007; Marchetti *et al.*, 2012; Smetacek *et al.*, 2012). Coastal members of this genus are periodically responsible for harmful algal bloom events due to their production of a neurotoxin, domoic acid (Hasle, 2002; Lelong *et al.*, 2012).

The success of *Pseudo-nitzschia* spp. in forming large blooms in the NE Pacific Ocean following iron addition events, and their ability to subsist during long periods of severe iron limitation, has been in part attributed to their competitive growth characteristics. The oceanic diatom *Pseudo-nitzschia granii* is capable of maintaining relatively high growth rates at extremely low iron concentrations. In comparison, the coastal diatom *Pseudo-nitzschia multiseries* cannot survive at the dissolved iron concentrations in which *P. granii* still thrives (Fig. 1). Even when compared to an isolate of the oceanic diatom *Thalassiosira oceanica* from the Sargasso Sea where macronutrients primarily or in combination with iron limit primary productivity (Moore *et al.*, 2006), *P. granii* can continue to grow under low-iron concentrations that do not support growth of *T. oceanica* (Fig. 1).

Received 13 May, 2018; revised 9 August, 2018; accepted 12 August, 2018. \*For correspondence. E-mail amarchetti@unc.edu; Tel. 919-843-3473; Fax (919) 962-1254.



**Fig. 1.** Specific growth rates as a function of log-transformed total iron concentration ( $\text{Fe}_T$ ) in the oceanic diatoms *Pseudo-nitzschia granii* (black) and *Thalassiosira oceanica* (grey), and the coastal diatom *Pseudo-nitzschia multiseriis* (white), regraphed from Marchetti et al. (2009). Error bars for *P. granii* and *P. multiseriis* represent variation among biological triplicates.

Initial investigations into the molecular underpinnings of *Pseudo-nitzschia* iron-limited and iron-enriched growth responses have provided some key insights into its iron-related physiology. *Pseudo-nitzschia granii* highly expresses the gene encoding the iron storage protein, ferritin, under iron-replete conditions (Marchetti et al., 2009), a response not shared by many other examined coastal diatoms (Cohen et al., 2018). This enables *P. granii* to take up and store iron during pulses of high iron availability, and then use these intracellular iron reserves for subsequent cell growth and division once iron in the surrounding environment declines back to low, ambient concentrations. By utilizing ferritin, *P. granii* is able to achieve a high intracellular iron quota of  $300 \mu\text{mol Fe mol C}^{-1}$ , yet can decrease its minimum iron quota to  $< 5 \mu\text{mol Fe mol C}^{-1}$  under low iron conditions, demonstrating a high plasticity in iron quotas when compared to other taxonomically and geographically diverse diatoms (Cohen et al., 2018).

Similar to many other diatoms from Fe-limited regions, *P. granii* also contains a gene encoding for proteorhodopsin, a light-driven proton pump used in the synthesis of ATP, and appears to preferentially use this protein for energy production under iron limitation as an alternative to photosynthesis (Marchetti et al., 2015). These two strategies contribute to the growing understanding of which molecular mechanisms underlie the ecological success of oceanic diatoms such as *P. granii*; however other major metabolic rearrangements occurring in photosynthetic carbon fixation, macronutrient assimilation or

vitamin production as a result of altered iron availability remain poorly resolved. While enrichment studies performed with natural diatom assemblages are useful in identifying large-scale shifts in cellular processes across broad taxonomic groups (Marchetti et al., 2012; Cohen et al., 2017a), it is challenging to assign these specific trends to the species taxonomic level. This species-level resolution may be important in interpreting ecological patterns since differences in phenotypic, transcriptional and proteomic tendencies across species or even strains can be pronounced (Godhe and Rynearson, 2017; Hippman et al. 2017).

Recently there has been growing evidence of the importance of B-vitamins in regulating the growth of phytoplankton (Sañudo-Wilhelmy et al., 2014). In addition to the iron-intensive processes involved in photosynthesis and nitrate assimilation, iron is intricately linked to vitamin production due to the presence of iron–sulphur (Fe–S) cluster-containing enzymes involved in the production of B<sub>1</sub> (thiamine) and B<sub>7</sub> (biotin) (Mihara and Esaki, 2002). It is unknown whether *P. granii* contains the molecular machinery to intracellularly biosynthesize thiamine to fulfill vitamin demands or instead relies on exogenous compounds, and whether the production process is influenced by iron availability as recently demonstrated with B<sub>7</sub> (Cohen et al., 2017b). Although some marine heterokonts appear to be thiamine producers that do not require externally supplied thiamine to grow (Sañudo-Wilhelmy et al., 2014), others may use environmentally-acquired intermediates or precursors and internally convert them to usable forms (Gutowska et al., 2017). Differences in thiamine requirements may additionally occur within a single genus, with an examined strain of *Pseudo-nitzschia multiseriis* not requiring thiamine in culture medium while a closely related strain of *Pseudo-nitzschia pungens* does (Tang et al., 2010; Sañudo-Wilhelmy et al., 2014). Therefore, nutritional requirements cannot be assumed constant across, or even within, a given taxon. Furthermore iron and vitamin B<sub>12</sub> (cobalamin) stress have been linked, with co-limitation observed in the HNLC Southern Ocean (Bertrand et al., 2011, 2013, 2015) and serial B<sub>12</sub> limitation occurring following nitrate and iron co-limitation in regions of the tropical Atlantic (Browning et al. 2017), although a direct effect of iron-limitation on B<sub>12</sub>-dependent processes has not been apparent in the few examined diatom cultures and natural communities (Bertrand et al., 2015; Cohen et al., 2017b). In addition, certain eukaryotic marine algae can convert unusable variants of B<sub>12</sub> (pseudocobalamin) into cobalamin using synthesis and repair proteins encoded by COB genes (Helliwell et al., 2016), and certain COB genes show elevated expression in natural *Pseudo-nitzschia* spp. following iron enrichment in the NE Pacific Ocean (Cohen et al., 2017b). Iron availability may therefore directly influence B<sub>1</sub>

production and perhaps B<sub>12</sub> intermediate transformations in *P. granii*.

Here, we present an assessment of iron-sensitive metabolic features specific to *P. granii*, an ideal model diatom for investigating iron-induced changes in molecular physiology in a low-iron requiring diatom. Through a combined transcriptomic and proteomic approach, we examined a *P. granii* strain originally isolated from the NE Pacific Ocean and identified several cellular mechanisms that are a function of iron status which appear unique to this diatom when compared to the published transcriptomes of two other model diatoms. We furthermore describe iron-coordinated substitutions and reorganizational strategies with strong transcript and protein support. Understanding these molecular responses and their inferred physiological consequences have implications for regional biogeochemistry and community ecology, and is increasingly relevant as we predict how phytoplankton communities will be affected by rapidly warming and acidifying marine environments, which may ultimately lead to changes in the ocean's iron supply (Hutchins and Fu, 2017; McQuaid *et al.*, 2018).

## Results and discussion

### Physiological response to iron limitation

*Pseudo-nitzschia granii* (UNC1102), originally isolated from naturally iron-limited waters of the subarctic North-east Pacific Ocean in June 2011, was grown in separate culture experiments for transcriptomic or proteomic analyses. *P. granii* cultures grown for the transcriptomic analysis achieved a maximum growth rate ( $\mu_{\max}$ ) of  $1.77 \pm 0.03 \text{ day}^{-1}$  under iron-replete conditions and  $0.67 \pm 0.13 \text{ day}^{-1}$  when iron-limited. Cultures grown for proteomic analysis reached a  $\mu_{\max}$  of  $2.16 \pm 0.12 \text{ day}^{-1}$  under iron-replete conditions and  $0.97 \pm 0.03 \text{ day}^{-1}$  when iron-limited. As a result, mean iron-limited growth rates for transcriptomic and proteomic analyses were reduced to 38% and 45% of  $\mu_{\max}$  respectively ( $p < 0.02$ ; Supporting Information Fig. S1). Consistent with the reductions in growth rates,  $F_v:F_m$  values significantly decreased by approximately 10% in the iron-limited cells ( $p < 0.03$ ; Supporting Information Fig. S1). Slight differences in specific growth rates of *P. granii* cells between transcriptomic and proteomic experiments may be attributed to differences in culture age and the associated changes in diatom cell size with time (Hostetter and Hosaw 1972), with the replicate cultures grown approximately 2 years apart during August–December 2012 for the transcriptomic analysis and January–March 2015 for the proteomic analysis.

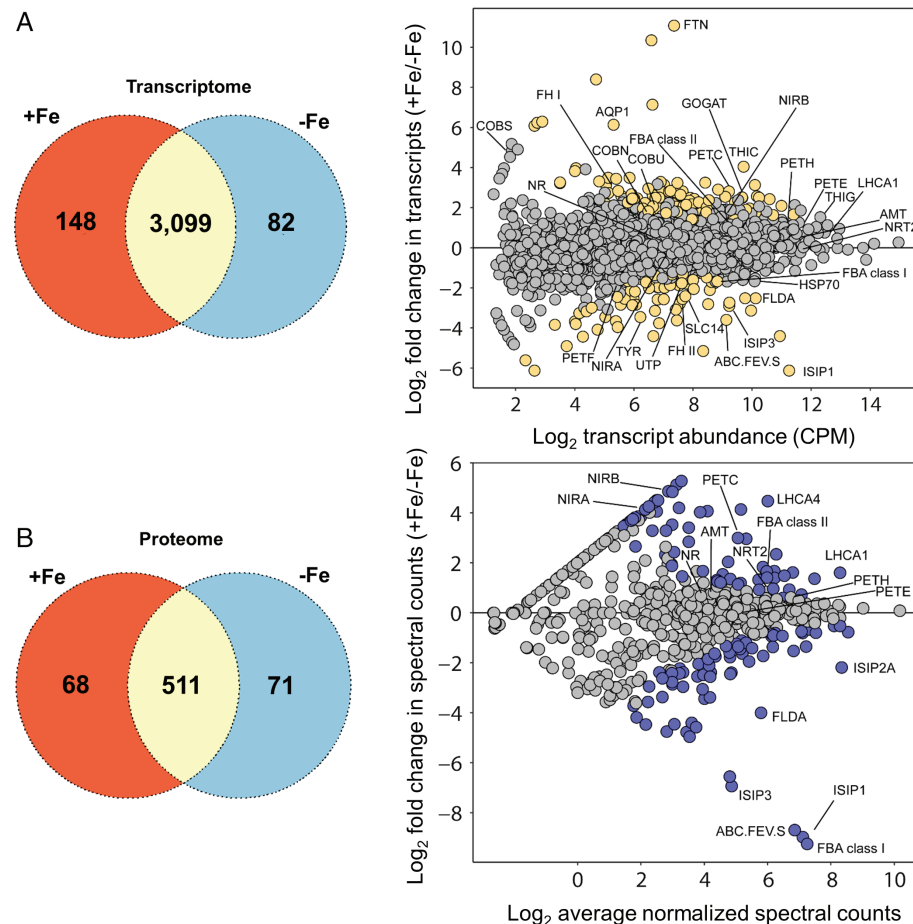
To gain an understanding in how *P. granii* responds to changes in iron bioavailability compared to other characterized model diatoms, a comparative transcriptomic analysis was performed using publicly available data sets

of the oceanic diatom *T. oceanica* and the polar sea ice diatom *Fragilariopsis cylindrus* when grown under iron-replete and iron-limited conditions. Both of these diatoms also grew at approximately 40% of their  $\mu_{\max}$  under iron-limited conditions as established in these other studies (Lommer *et al.*, 2012; Strauss, 2012; Mock *et al.*, 2017).

### Transcriptome sequencing and proteome analyses

High-throughput sequencing of *P. granii* cDNA libraries resulted in 29 million paired-end reads, with an average of  $5.8 \pm 0.7$  million reads per sample (Supporting Information Table S1). Reads assembled into ~70 000 contigs with an N50 of 1199 base pairs. *Pseudo-nitzschia granii* transcriptome completeness was estimated at 66% according to the BUSCO metric (Simão *et al.*, 2015), suggesting sufficient coverage of core genes present within eukaryotic genomes. We detected 3328 unique genes with functional annotations to the KEGG database, of which 7% were differentially expressed between iron-replete and iron-limiting conditions as determined through the edgeR likelihood ratio test (Robinson *et al.*, 2010). Of these, 148 genes were over-represented and 82 were under-represented in the iron-replete condition (false discovery rate [FDR]  $< 0.05$ ; Fig. 2A, Supporting Information Fig. S2 and Table S2). Approximately 1446 KEGG-annotated genes were assigned to the *T. oceanica* transcriptome, and 3589 to the *F. cylindrus* transcriptome. Almost half of these genes (46%) in *F. cylindrus* were differentially expressed (FDR  $< 0.05$ , likelihood ratio test) due to exceptionally low variability among biological replicates. All three diatom species shared transcripts corresponding to 1171 KEGG genes. A number of plastid-encoded genes were detected in the nuclear transcriptome despite targeting mRNA through poly-A selection during sequence library preparation. Because organelle-derived transcripts can contribute to a substantial proportion of the reads contained in RNA-Seq libraries, some being polyadenylated, plastid-encoded genes can be a rich source of expression information and therefore were included in our analysis (Smith, 2013 and references therein). In addition, multiple copies of a given gene may exist in diatoms and each copy could show distinct expression patterns (Bender *et al.*, 2014; Groussman *et al.* 2015). In our analysis, gene expression is a composite of all gene copies to gauge cumulative metabolic responses to iron status, although subsequent analysis of distinct gene copies may be informative.

Proteomic analyses detected 326,585 mass spectra from four biological samples, resulting in 1204 protein identifications at a corresponding protein false discovery rate of 1.0% (Supporting Information Table S3). Approximately 67%, or 805, of these proteins had KEGG annotations and 649 corresponded to unique functional gene annotations. Of these annotated proteins, 139 were



**Fig. 2.** Venn diagrams and MA (log ratio vs. mean average) plots displaying the number of KEGG-annotated genes differentially expressed in the transcriptome (A) or differentially abundant in the proteome (B) under iron-replete (+Fe) and iron-limited (−Fe) conditions. The MA plots show differential abundance between treatments. Each point on the plot corresponds to a unique KEGG gene. The y-axis represents the  $\log_2$  fold change in expression between treatments, while the x-axis represents  $\log_2$  normalized transcript abundance in counts per million (CPM) (A), or the average normalized spectral counting score (NSAF) across the two treatments (B). Points are coloured if differentially expressed or abundant between treatments (edgeR likelihood ratio test or PLGEM, FDR/FPR < 0.05). Proteins with no spectral counts in the numerator or denominator of either treatment were manually assigned a score of 0.5 to be displayed on the plot. Only genes or proteins with KEGG functional annotations are shown. *PETC*: cytochrome  $b_6f$  complex iron–sulphur subunit; *LHCA*: light-harvesting complex I chlorophyll a/b binding protein. See Fig. 4 for full list of gene names.

differentially abundant between treatments with 68 over-represented and 71 under-represented in the iron-replete condition (estimated FPR < 0.05, power law global error model [PLGEM]; Fig. 2B).

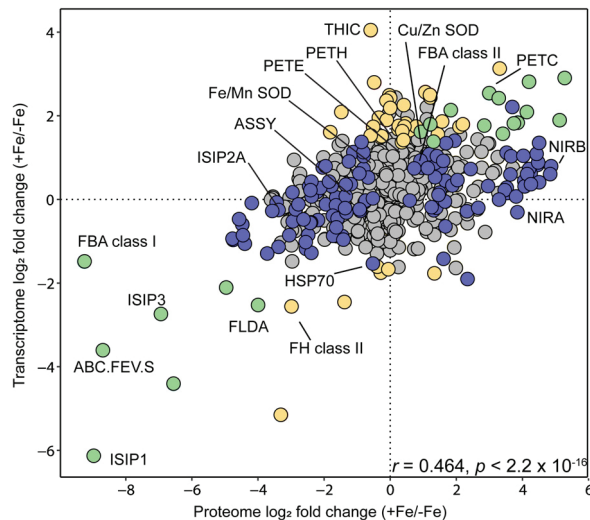
The transcriptome and proteome of *P. granii* demonstrated substantial agreement in terms of metal transporters and components of iron-related photosynthesis between iron-replete and iron-limited conditions (+Fe/−Fe), and demonstrated an overall positive Pearson correlation coefficient ( $r$ ) of 0.464 ( $p = 2.2 \times 10^{-16}$ ; Fig. 3). Of the 3327 unique KEGG genes detected in the *P. granii* transcriptome, 650 (20%) were also detected in the proteome using the analytical approach applied in this study. Only 21 shared KEGG genes and proteins were similarly differentially expressed or abundant (FDR/FPR < 0.05), while 40 and 118 were exclusively differentially expressed in the transcriptome or differentially abundant in the proteome respectively (Supporting Information Table S4).

#### Iron-limited molecular responses of *P. granii*

Genes and proteins with congruent responses to iron status include components of photosynthesis, metal transport and the Calvin cycle (Fig. 3, Supporting Information

Fig. S3). The gene encoding an iron-independent photosynthetic electron acceptor flavodoxin (*FLDA*), which can functionally replace iron-dependent ferredoxin (*PETF*), increased in gene expression by 5.7-fold under iron limitation (FDR =  $5.3 \times 10^{-5}$ ). Similarly, *FLDA* protein abundance was 16-fold higher under iron limitation ( $p = 0$ , FPR < 0.05). Interestingly, the substitution of *FLDA* with *PETF* under iron-replete conditions, as observed in natural diatom communities (LaRoche *et al.*, 1995; Doucette *et al.*, 1996; Erdner and Anderson, 1999; McKay *et al.*, 1999; Allen *et al.*, 2008), was not apparent in *P. granii*. Instead *PETF* transcript abundance was appreciably lower than *FLDA* regardless of iron status (Supporting Information Fig. S2). As a result, the use of flavodoxin as an iron stress marker independent of ferredoxin dynamics may be more appropriate, as observed in non-diazotrophic marine cyanobacteria (Saito *et al.*, 2014; Mackey *et al.*, 2015). Consistent with our findings, a global survey of gene expression in marine phytoplankton has demonstrated oceanic members of the diatom lineage strongly express flavodoxin over ferredoxin, while certain coastal diatoms, likely adapted to environments experiencing more frequent and larger fluctuations in iron supply, more highly express ferredoxin under iron-replete conditions (Carradec *et al.*, 2018).





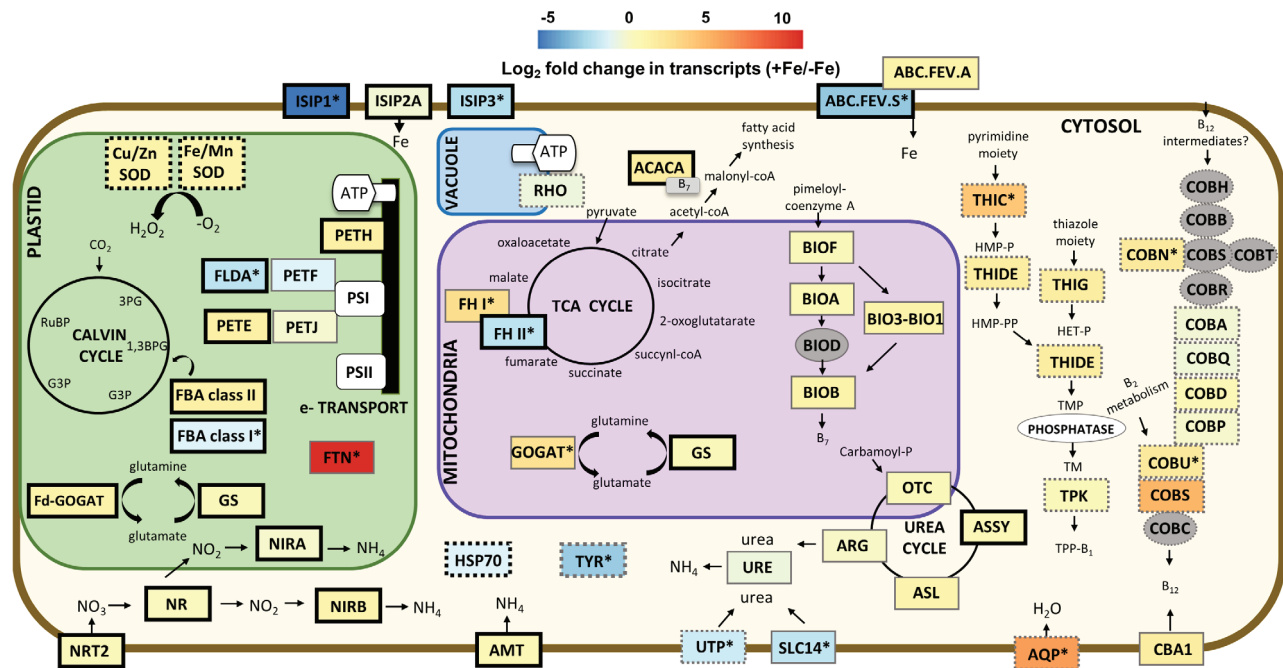
**Fig. 3.** Comparison of  $\log_2$  fold changes in gene (y-axis) and protein (x-axis) expression under iron-replete relative to iron-limited conditions (+Fe/-Fe). Only genes and proteins with KEGG gene annotations that were detected in both the transcriptome and proteome are shown. Genes differentially expressed only in the transcriptome are shown in yellow (edgeR's likelihood ratio test, FDR < 0.05), while those differentially abundant only in the proteome are shown in blue (PLGEM, estimated FPR < 0.05). Genes and proteins both differentially expressed or abundant are shown in green, and those not differentially expressed or abundant are shown in grey. The strength of the  $\log_2$  fold change relationship between the transcriptome and proteome was assessed using the Pearson correlation coefficient ( $r$ ). See Fig. 4 for full list of gene names.

Apart from photosynthetic carbon fixation, ferredoxins are known to participate in metabolic reactions involved in nutrient assimilation, chlorophyll synthesis and redox cycling (Hanke and Mulo, 2012), and they may serve an important cellular role under iron limitation. In *P. granii*, *PETF* gene expression increased by threefold when cells were iron-limited (Fig. 2A) in contrast to the responses observed in three diatoms from the genus *Thalassiosira* (Whitney *et al.*, 2011; Lommer *et al.*, 2012). An amino acid alignment of *PETF* from model diatoms demonstrates that *P. granii* ferredoxin resembles that of the HNLC Southern Ocean diatom *F. cylindrus* (71% similarity; Supporting Information Fig. S4A) with a Phyre2-modelled secondary structure consistent with the 2Fe-2S ferredoxin-related family (Kelley *et al.*, 2015). These results paradoxically support a distinct role of ferredoxin under iron-limited conditions in *P. granii* and perhaps *F. cylindrus*. As ferredoxin proteins are additionally hypothesized to play a role in detoxifying reactive oxygen species under heat stress (Lin *et al.* 2013) and are expressed in one *T. oceanica* strain under copper limitation (Hippman *et al.* 2017), ferredoxin in *P. granii* could be used to neutralize elevated reactive oxygen species accumulation under iron stress (Luo *et al.* 2015). Alternative detoxifiers to ferredoxin such as the heme-dependent ascorbate peroxidase are conversely overexpressed in the

transcriptome under iron-replete conditions in *P. granii* ( $p = 9.2 \times 10^{-4}$ ; Supporting Information Table S2). Further investigations are required to determine if differences in ferredoxin amino acid sequences are suggestive of distinct cellular functions in marine diatoms.

Differential gene expression and protein abundance of iron transport pathways was additionally evident. *Pseudo-nitzschia granii* likely possesses the ability to import iron using a variety of transport mechanisms (Fig. 4, Supporting Information Fig. S5). The gene encoding a substrate-binding protein involved in an ABC-type iron transport complex identified in cyanobacteria (ABC.FEV.S) (Katoh *et al.*, 2001) increased in gene expression under iron limitation by 12-fold (FDR =  $1.7 \times 10^{-9}$ ) while high protein spectral counts were only detected in the iron-limited treatment (represented by a protein abundance increase by 388-fold [ $p = 0$ , FPR < 0.05] due to no spectral counts detected under iron-replete conditions and a trivial spectral count value of 0.5 being artificially assigned to allow for a fold-change estimation). These ABC.FEV.S responses are consistent with observations of higher gene expression in natural populations of *Thalassiosira* under iron stress (Cohen *et al.*, 2017a). In field studies with *Pseudo-nitzschia*, ABC.FEV.S was found to be similarly overexpressed in potentially iron-limited natural diatom communities within the Bransfield Strait of the Antarctic Peninsula (Pearson *et al.*, 2015), while low-iron Northeast Pacific and coastal California upwelling *Pseudo-nitzschia* populations increased ABC.FEV.S gene expression following iron resupply (Cohen *et al.*, 2017a). Many of these studies imply a large increase in ABC.FEV.S transporters to iron scarcity; however, there are likely complex environmental contextual factors that may result in these transport systems serving different roles across diatom species and distinct physiochemical conditions, which warrants further investigation. We conclude here that *P. granii* relies upon this ABC transport complex under steady-state iron limitation given the strong transcriptional and proteomic support observed here. The discovered tryptic peptides from this protein and example mass spectra are shown in Supporting Information Table S5 and Fig. S6 to enable future targeted metaproteomic applications for this potential *Pseudo-nitzschia* iron stress biomarker.

A group of abundant and iron-responsive proteins involved in iron transport were examined (Fig. 4). Iron-starvation-induced-protein 1 (*ISIP1*), involved in the non-reductive uptake of iron bound to siderophores through endocytosis in marine diatoms (Kazamia *et al.*, 2018), belongs to the same gene cluster as the copper-containing enzyme tyrosinase (*TYR*; Allen *et al.*, 2012) with both highly expressed under iron limitation in *P. granii* (Fig. 4). *ISIP1* transcripts were abundant ( $\log_2$  CPM = 11.3) and it was one of the most differentially



**Fig. 4.** Cell schematic displaying the log<sub>2</sub> fold change in gene expression between iron-replete (+Fe) and iron-limited (–Fe) treatments in *P. granii*. On the continuous colour scale, red represents higher expression under iron-replete conditions, and blue represents higher expression under iron-limited conditions. Grey indicates that transcripts were not detected. Genes encoding proteins that were also detected in the proteome are outlined with a black border. A dashed outline indicates protein location in the cell is speculative. Asterisks (\*) denote a significant difference in gene expression between treatments as determined through edgeR's likelihood ratio test (FDR < 0.05). The B<sub>12</sub> pathway, urea cycle, TCA cycle, nitrate and urea transporters, and localization of ISIPs, SOD and GOGATs were constructed based on information from Alexander *et al.* (2015), Allen *et al.* (2011), Bender *et al.* (2014), Bertrand *et al.* (2012), Gutowska *et al.* (2017), Hockin *et al.* (2012), Lommer *et al.* (2012), McRose *et al.* (2014), Paerl *et al.* (2016), Pearson *et al.* (2015), Pourcel *et al.* (2013), Smith *et al.*, (2016) and Wolfe-Simon *et al.* (2006). [Cytosol] *ISIP1*: iron starvation-induced protein 1; *ISIP2A*: iron starvation-induced protein 2A; *ISIP3*: iron starvation-induced protein 3; *ABC.FEV.S*: iron complex transport system substrate-binding protein; *ABC.FEV.A*: Iron (III) dicitrate transport ATP-binding protein; *ACACA*: biotin carboxylase; *HSP70*: heat shock protein 70 kDa 1/8; *TYR*: tyrosinase; *NRT2*: nitrate transporter; *AMT*: ammonium transporter; *NR*: nitrate reductase; *NIRB*: NADPH-nitrite reductase; *UTP*: urea transporter; *SLC14*: solute carrier family 14 (urea transporter); *URE*: urease; *ARG*: arginase; *ASL*: argininosuccinate lyase; *ASSY*: argininosuccinate synthase; *AQP*: aquaporin; *THIC*: phosphomethylpyrimidine synthase; *THIDE*: hydroxymethylpyrimidine kinase; *THIG*: thiazole synthase; *TPK*: thiamine pyrophosphokinase; *COBH*: precorrin-8X/cobalt-precorrin-8 methylmutase; *COBB*: cobyrinic acid a,c-diamide synthase; *COBN*, *COBS*, *COBT*: cobaltochelatase; *COBR*: cob(II)yrinic acid a,c-diamide reductase; *COBA*: cob(II)alamin adenosyltransferase; *COBU*: adenosylcobyrinic acid synthase; *COBD*: adenosylcobinamide-phosphate synthase; *COBP*: adenosylcobinamide kinase; *COBC*: nicotinate-nucleotide-dimethylbenzimidazole phosphoribosyltransferase; *COBC*: alpha-ribazole phosphatase; *COBS*: adenosylcobinamide-GDP ribazoletransferase; *CBA1*: cobalamin acquisition protein 1. [Plastid] *Cu/Zn SOD*: Copper/zinc superoxide dismutase; *Fe/Mn SOD*: Iron/manganese superoxide dismutase; *FBA class I*: fructose biphosphate aldolase class I; *FBA class II*: fructose biphosphate aldolase class II; *Fd-GOGAT*: ferredoxin-dependent glutamate synthase; *GS*: glutamine synthetase; *FTN*: ferritin; *PETE*: plastocyanin; *PETJ*: cytochrome c<sub>6</sub>; *FLDA*: flavodoxin; *PETF*: ferredoxin; *PETH*: ferredoxin-NADP<sup>+</sup> reductase. [Mitochondria] *FH I*: fumarate hydratase class I; *FH II*: fumarate hydratase class II; *GOGAT*: NADPH-dependent glutamate synthase; *NIRA*: ferredoxin-dependent nitrite reductase; *OTC*: ornithine carbamoyltransferase; *BIOF*: 8-amino-7-oxononanoate synthase; *BIOA*: adenosylmethionine-8-amino-7-oxononanoate aminotransferase; *BIOD*: dethiobiotin synthetase; *BIO3-BIO1*: bifunctional dethiobiotin synthetase; *BIOB*: biotin synthase. [Vacuole] *RHO*: proteorhodopsin.

expressed genes under iron limitation with a 64-fold change in gene expression (FDR =  $4.2 \times 10^{-17}$ ; Fig. 2A). *ISIP1* protein expression followed a similar trend and increased 478-fold under iron limitation (not detected under iron-replete conditions;  $p = 0$ , FPR < 0.05; Fig. 2B). *ISIP3* gene expression likewise increased by sixfold under iron limitation (FDR =  $3.8 \times 10^{-6}$ ), and protein abundance by 119-fold (not detected under iron-replete conditions;  $p = 0$ , FPR < 0.05). Both *ISIP1* and *ISIP3* have been observed to increase under iron stress in numerous diatom isolates as well as in natural communities (Allen *et al.*, 2008; Marchetti *et al.*, 2012; Bender *et al.*, 2014; Cohen *et al.*, 2017a). Lastly, *ISIP2A*, which

encodes a recently characterized phytotransferrin that binds ferric iron directly and is internalized through endocytosis (Morrissey *et al.*, 2015; McQuaid *et al.*, 2018), was weakly expressed under iron limitation by 1.4-fold (FDR = 0.76) while protein abundance was 4.5-fold higher under iron limitation ( $p = 0$ , FPR < 0.05). Taken together, given their high transcript and protein abundance levels and sensitivity to iron status observed within both laboratory and field analyses, these ISIP proteins are thought to play critical roles in facilitating both labile inorganic and ligand-bound iron uptake (Kazamia *et al.*, 2018; McQuaid *et al.*, 2018) and in maintaining *Pseudo-nitzschia* growth under iron-limiting conditions.

Several coordinated protein substitutions further demonstrate *P. granii*'s extensive metabolic reorganization when adapting to changes in iron availability. Two classes of fructose biphosphate aldolases (FBA) functioning in the Calvin cycle, gluconeogenesis and glycolysis, were identified with class II and class I forms differentially expressed as a function of iron status (Fig. 4). FBA class I uses Schiff-based catalysis, not requiring a metal cofactor, while class II depends on metal catalysis (Horecker *et al.*, 1972). Class II FBA transcripts were 2.5-fold (FDR = 0.08), and proteins 3.2-fold ( $p = 0.007$ , FPR < 0.05) more abundant under iron-replete conditions, while class I FBA transcripts were 2.8-fold (FDR = 0.04) and proteins 588-fold ( $p = 0$ , FPR < 0.05) more abundant under iron limitation (Fig. 3). The patterns we observed are consistent with laboratory cultures of *T. oceanica* and *P. tricornutum* grown under varying iron states (Allen *et al.* 2011; Lommer *et al.*, 2012). Similar to the FBAs, fumarate hydratase (FH) plays a critical role in the tricarboxylic acid (TCA) cycle and can exist in two different forms, with class I (FH I) containing iron–sulphur clusters (Picaud *et al.* 2011). FH I transcripts were sixfold (FDR =  $6 \times 10^{-3}$ ) more abundant under iron-replete conditions, while FH II transcripts were sixfold (FDR =  $5.1 \times 10^{-5}$ ) and proteins 7.9-fold ( $p = 0.08$ ) more abundant under iron limitation (Figs 2 and 4). The iron-dependent protein substitutions and acquisition strategies presented here collectively aid in our understanding of how *P. granii* maintains growth under iron stress and efficiently reorganizes its metabolism under iron-replete conditions.

Previous combined transcriptomic and proteomic studies have also shown consistency with regards to important nutrient stress biomarkers in diatoms; both transcripts and proteins involved in phosphate metabolism and B<sub>12</sub> acquisition displayed coordinated patterns under phosphate-limited and B<sub>12</sub>-limited conditions respectively (Bertrand *et al.*, 2012; Dyhrman *et al.*, 2012). We provide here additional agreement between gene expression and protein abundance of photosynthetic components and iron transporters in *P. granii* as a function of iron status. While FLDA and ISIP proteins are already established bioindicators of iron stress in *Thalassiosira* (Lommer *et al.*, 2012; Chappell *et al.* 2015; Marchetti *et al.* 2017), a protein involved in the cytochrome b<sub>6</sub>f complex (PETC) and FBA are additional reliable molecular markers (Fig. 3) and should also be considered when assessing iron bioavailability in natural *Pseudo-nitzschia* populations.

Conversely, a number of genes and proteins involved in critical metabolic processes lacked consistent patterns between the transcriptome and proteome, including a number that demonstrated opposite responses to iron-replete conditions (quadrants II and IV in Fig. 3). Two different forms of superoxide dismutase enzymes were

identified, with one requiring copper and zinc (Cu/Zn SOD) and the other either iron or manganese (Fe/Mn SOD). Both SODs in the transcriptome increased expression under iron-replete conditions (by 2.3-fold and 1.9-fold respectively) possibly as a result of iron-induced photosynthetic production of superoxide radicals (Asada, 2006). Although Cu/Zn SOD protein abundance increased by twofold under iron-replete conditions, Fe/Mn SOD protein abundance exhibited an opposite pattern and increased by 1.9-fold under iron limitation (Supporting Information Fig. 3). It is difficult to determine whether this Fe/Mn SOD enzyme relies upon Fe or Mn as the metal cofactor based on protein sequence alone (Hunter *et al.*, 2017), although we hypothesize this isoform utilizes Mn in its active site to reduce overall cellular iron demand. Future biochemical analyses are needed to resolve its metal specificity in *P. granii*.

Divergent trends were furthermore identified in nitrogen acquisition pathways, with the nitrate transporter NRT2 and assimilation components ferredoxin-dependent and NADPH-dependent nitrite reductases (NIRA and NIRB) not differentially expressed at the gene level, while protein abundances were appreciably higher under iron-replete conditions. NIRA and NIRB were two of the most highly differentially abundant proteins detected under iron-replete conditions ( $p = 0.04$  and  $0.01$ , respectively, FPR < 0.05), with NRT2 also 1.8-fold ( $p = 0.05$ ) more abundant under iron-replete conditions (Supporting Information Fig. S3). There are a number of possible explanations for these discrepancies, including post-transcriptional and post-translational regulation, and/or differential mRNA and protein stabilities and their influence on RNA and protein inventories (Walworth *et al.*, submitted), highlighting the usefulness of examining both transcripts and proteins whenever possible in order to fully interpret metabolic responses of diatoms to environmental gradients.

#### *Distinct molecular physiology of P. granii compared to other model diatoms*

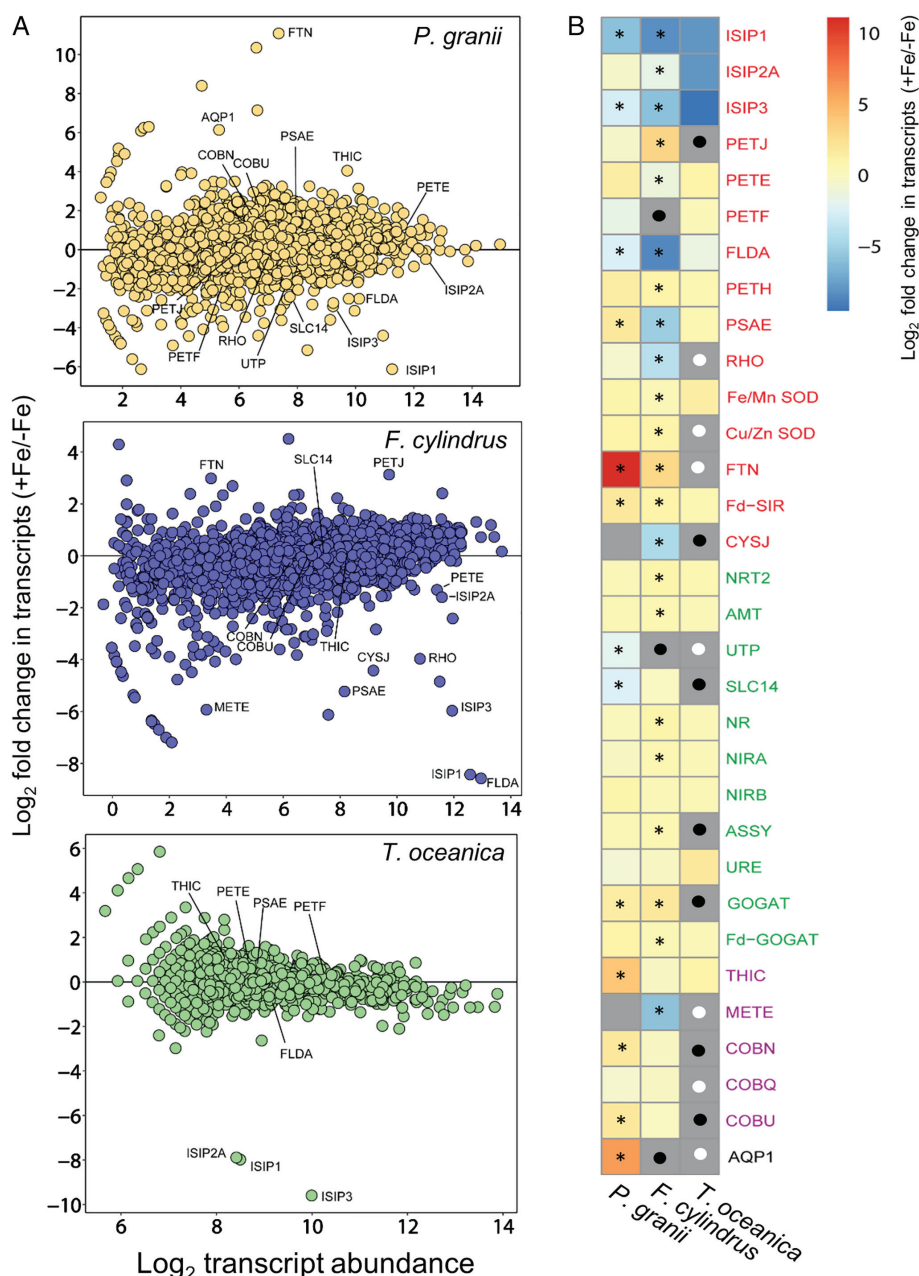
Several genes involved in photosynthesis, energy (i.e., ATP) production, nitrogen assimilation and vitamin production demonstrated distinctive gene expression patterns to iron availability in *P. granii* with responses not uniform across the other model diatoms examined. Transcripts for the gene encoding cytochrome c<sub>6</sub> (PETJ), an iron-containing protein also involved in photosynthetic electron transport, were weakly abundant in both iron treatments in *P. granii* while transcripts of the gene encoding the copper-containing functional replacement for this protein, plastocyanin (PETE) (Peers and Price, 2006) were highly abundant, increasing by 2.4-fold under iron-replete conditions (Fig. 5). In contrast,



*F. cylindrus* demonstrates the opposite response and more highly expresses the genes *PETJ* under iron-replete conditions by 8.7-fold (FDR =  $6.5 \times 10^{-85}$ ) and *PETE* by 2.5-fold under iron-limited conditions (FDR =  $2.7 \times 10^{-22}$ ). Our results suggest *P. granii* is particularly adapted to its chronically iron-limited environment and constitutively utilizes PETE, rather than as a temporary replacement for PETJ under iron limitation as observed in *F. cylindrus*.

Additional differences in photosynthetic gene expression patterns were identified among diatoms despite growing at a similar degree of iron limitation. Transcripts

for proteorhodopsin (*RHO*), a putative iron-independent light-driven proton pump capable of generating energy in marine microbes (Beja *et al.*, 2000), were more abundant under iron-limitation in both *P. granii* (1.7-fold), and *F. cylindrus* (16-fold, FDR =  $4.1 \times 10^{-34}$ , Fig. 5). Previous reports show *RHO* gene expression 10-fold higher in iron-limited *P. granii* as determined through qPCR techniques, though no qPCR expression data for *Fragilariopsis* is available for comparison (Marchetti *et al.*, 2015). This larger degree of *RHO* regulation in *F. cylindrus* may indicate differences in rhodopsin mRNA/protein stabilities between the two groups, though it is also possible that



**Fig. 5.** MA (log ratio vs. mean average) plots showing differential gene expression between acclimated iron-replete (+Fe) and iron-limited (-Fe) treatments in *P. granii* compared to the diatoms *Fragilariopsis cylindrus* and *Thalassiosira oceanica* (A). Heat-map displays the log<sub>2</sub> fold change in gene expression between iron treatments for genes of interest encoding proteins involved in iron (red), nitrogen (green), vitamin (purple) or other (black) metabolic processes (B). Asterisks (\*) denote a significant difference in gene expression between treatments in *P. granii* and *F. cylindrus* as determined through edgeR's likelihood ratio test (FDR < 0.05). Differential expression was not assessed in *T. oceanica* due to lack of biological replication. Grey-filled boxes indicate transcripts were not detected. For *T. oceanica* and *F. cylindrus*, a black circle indicates the gene was detected in the genome, while a white circle indicates it was absent in the genome. PSAE: photosystem I subunit IV; CYSJ: NADPH-sulphite reductase. See Fig. 4 for full list of gene names. *F. cylindrus* raw reads were obtained from Mock *et al.*, 2017, and *T. oceanica* raw reads and assembly from Lommer *et al.*, 2012.



light-limited Southern Ocean diatoms depend more heavily on rhodopsins than temperate diatoms such as *P. granii*.

The gene encoding PSI subunit IV (*PSAE*), involved in the docking of ferredoxin to PSI (Jeanjean *et al.*, 2008), was more highly expressed by 3.7-fold (FDR = 0.04) in *P. granii* under iron-replete conditions and proteins were more abundant by 35-fold ( $p = 0.008$ , FPR < 0.05), yet transcripts were more highly expressed by 38-fold under iron limitation in *F. cylindrus* (FDR =  $6 \times 10^{-193}$ ; Fig. 5B). The pattern in *P. granii* is consistent with low iron-adapted temperate diatoms decreasing their iron-rich PSI relative to PSII ratios (Strzepek *et al.*, 2004) and may indicate that the open ocean diatom *P. granii* more effectively rearranges its photosynthetic machinery when growing under chronic iron limitation as compared to a polar sea ice diatom, which appears to undergo a different re-arrangement of photosynthetic proteins under iron-limiting conditions. Finally, both *P. granii* and *F. cylindrus* more highly expressed ferredoxin-dependent sulphite reductase (*Fd-SIR*) under iron-replete conditions, yet only *F. cylindrus* increased gene expression of an alternative form of this enzyme, NADPH-dependent sulphite reductase (*CYSJ*), by 21-fold under iron limitation (FDR =  $1 \times 10^{-86}$ ; Fig. 5).

Although both the NE Pacific Ocean and Southern Ocean experience iron limitation, *F. cylindrus* was isolated from the Weddell Sea (64°S, 48°W) where total dissolved iron concentrations in the mixed layer can be 20-fold higher than those typical of the HNLC Northeast Pacific (1 nmol L<sup>-1</sup> vs. 0.05 nmol L<sup>-1</sup>; Westerlund and Öhman, 1991; Boyd and Harrison, 1999). The physiochemical environments between the two regions are furthermore vastly different with distinctive light levels, ranges in temperature, and micronutrient availability all influencing microbial community functional capabilities. Even within a particular ecosystem, species can be adapted to specific microenvironments, a notable example being *F. cylindrus* outcompeting other phytoplankton under high and constant light levels (Arrigo *et al.*, 2010). We therefore hypothesize that the dissimilarities in *PETE*, *PETJ*, *RHO*, *PSAE* and *CYSJ* gene expression patterns between *P. granii* and *F. cylindrus* represent pennate diatom ecotypes that have evolved to their distinct environments.

Nitrate assimilation is an iron-intensive process with several proteins involved directly requiring iron as a metal cofactor (Rueler and Ades, 1987). Several key nitrogen uptake and assimilation genes were not differentially expressed under iron-replete conditions in *P. granii*, including nitrate reductase (*NR*), ferredoxin-nitrite reductase (*NIRA*) and a nitrate transporter (*NRT2*; Fig. 4). Since the transcript abundance of these assimilation genes increased following iron enrichment in natural

diatom assemblages of the HLNC Northeast Pacific Ocean (Marchetti *et al.*, 2012; Cohen *et al.*, 2017a), these differences may be a result of steady-state versus non-steady state responses to changes in iron status. We speculate that transcript levels initially may rise following pulse iron additions, but later decrease under steady-state iron-replete conditions once high mRNA levels and protein pools are established. Under this scenario, natural communities of iron-limited *P. granii* would immediately undergo selective partitioning of iron resources once iron is resupplied, in which newly acquired iron is preferentially shunted into iron-requiring pathways that were previously constrained by low iron availability.

Diatoms utilize the urea cycle for a source of reduced nitrogen compounds when nitrate is not available (Allen *et al.*, 2012), and *P. granii* contains the full pathway in addition to several urea transporters (Fig. 4). A urea transporter (*UTP*) and solute-family carrier 14 (*SLC14*) were both more highly expressed in *P. granii* under iron limitation by fourfold (FDR =  $2 \times 10^{-3}$ ) and 5.3-fold (FDR =  $3 \times 10^{-3}$ ) respectively. Although *SLC14* is found in both the *T. oceanica* and *F. cylindrus* genomes, it was not detected in the *T. oceanica* transcriptome or differentially expressed in *F. cylindrus*. *UTP* was similarly not detected in the *T. oceanica* genome or *F. cylindrus* transcriptome (Fig. 5). Urea is converted into ammonium via the enzyme urease (*URE*), and this gene was more highly expressed by 1.7-fold under iron limitation. This is in contrast to *URE* gene expression in *T. oceanica*, which is 3.2-fold higher under iron-replete conditions (Fig. 5). The only urea cycle component identified in the proteome was argininosuccinate synthase (*ASSY*), and proteins were more abundant by 2.3-fold under iron-limiting conditions ( $p = 0.04$ , FDR < 0.05), although it was not differentially expressed in the transcriptome (Supporting Information Table S4). *P. granii* appears to be rather unique among the examined diatoms in its urea transport and degradation under iron limitation, perhaps as a source of reduced nitrogen to conserve intracellular iron otherwise needed for iron-rich nitrate reduction.

#### Vitamin synthesis and cobalamin remodelling

*Pseudo-nitzschia granii* highly expresses components of vitamin synthesis and salvage pathways with such gene expression responses comparatively reduced or absent in the diatoms *F. cylindrus* and *T. oceanica* (Fig. 5). The thiamine biosynthesis gene *THIC* (phosphomethylpyrimidine synthase), encoding a 4Fe-4S cluster-containing enzyme which converts purine intermediates into HMP-P (4-amino-5-hydroxymethyl-2-methylpyrimidine), was one of the most differentially expressed genes in *P. granii* with expression 16-fold higher under iron-replete conditions (FDR =  $1.9 \times 10^{-5}$ ; Fig. 5). The thiamine

biosynthesis module was collectively more highly expressed by 16-fold under iron-replete conditions rendering it one of the most differentially expressed modules identified (Supporting Information Fig. S5). *THIC* was conversely not differentially expressed in *F. cylindrus*, though was in *T. oceanica* by 2.3-fold, suggesting oceanic diatoms may have a greater need for the gene product HMP-P following iron enrichment compared to sea ice diatoms (Fig. 4). The continued expression of thiamine biosynthesis genes despite thiamine being added to the culture medium suggests *P. granii* does not have a mechanism to take up HMP-P from the environment, and continues to biosynthesize products even when thiamine is present externally. The *THIC* gene has not been identified in diverse and globally abundant haptophytes with available transcriptomes and genomes, consistent with observations of haptophyte members utilizing exogenously-acquired intermediates to fulfill thiamine demands (Gutowska et al., 2017). These findings support substantial differences in vitamin utilization strategies between haptophytes and diatoms, and even within the diatom lineage, likely influencing niche differentiation in natural surface plankton communities, micronutrient recycling patterns and regional microbial community dynamics.

Several cobalamin synthesis and repair proteins contained within the KEGG cobalamin synthesis module were identified within the *P. granii* transcriptome (Fig. 4). Sequence homology analyses indicate these *COB* genes match annotated and unannotated genes within the *F. cylindrus* genome as well as other eukaryotic protists available within NCBI's non-redundant protein database ( $E$ -value <  $10^{-3}$ ). Noticeably transcripts corresponding to genes encoding proteins that synthesize the corrin ring structure of the B<sub>12</sub> molecule were absent, while most of the genes involved in the salvage and repair portion of the pathway were detected (Fig. 4), including cobalt insertion, adenosylation and  $\alpha$ -ribazole attachment. These genes have also been observed in the metagenomes of other non-B<sub>12</sub>-producing organisms (Rodionov et al., 2003; Bertrand et al., 2015; Helliwell et al., 2016; Heal et al., 2017). Although most genes were not differentially expressed as a function of iron status, the gene *COBN* (cobaltochelatease), encoding a cobalt insertion protein and the final synthesis genes *COBU* (nicotinate-nucleotide--dimethylbenzimidazole phosphoribosyltransferase) and *COBS* (adenosylcobinamide-GDP ribazole-transferase) were more highly expressed under iron-replete conditions by 2.5-fold, (FDR = 0.01), 3.6-fold (FDR =  $4 \times 10^{-3}$ ) and 28-fold (FDR = 0.1) respectively. *COBN* and *COBU* were similarly detected in the *T. oceanica* genome and *F. cylindrus* transcriptome, though in contrast to *P. granii*, were not differentially expressed under iron-replete conditions. In addition to these cobalamin synthesis and repair genes, all three

diatoms contain a cobalamin acquisition protein (*CBA1*) which binds external cobalamin and allows for intracellular transport (Bertrand et al., 2012, 2015). *CBA1* gene expression was similarly threefold higher under iron-replete conditions compared to iron limitation in *P. granii* (FDR = 0.06; Fig. 4; Supporting Information Table S2).

Faster growth rates in *P. granii* under iron-replete conditions could be resulting in elevated *COB* synthesis/repair and *CBA1* gene expression in order to take up cobalamin directly or process its intermediates from the environment, which is then primarily used as a cofactor in the cobalamin-dependent enzyme methionine synthase (*METH*) enzyme to synthesize the amino acid methionine. In contrast, *F. cylindrus* utilizes a cobalamin-independent version of methionine synthase (*METE*) under iron limitation despite cobalamin being supplied to growth media (Fig. 5; Mock et al., 2017; Paaanen et al., 2017), with *METE* absent in both the *P. granii* transcriptome and *T. oceanica* genome. *METE* gene expression is strongly regulated by cobalamin availability in a variety of diatoms (Bertrand et al., 2012, 2013, 2015; Ellis et al., 2017) with iron having little to no influence on gene expression in laboratory cultures of diatoms *Grammonema cf. islandica*, also isolated from the NE Pacific Ocean (Cohen et al., 2017b), *Phaeodactylum tricornutum* (Bertrand et al., 2012) or in natural populations of *Fragilariopsis* from the Southern Ocean (Bertrand et al., 2015). Given these gene expression responses, it is possible *P. granii* uses a cobalamin remodelling and/or repair of degraded cobalamin approach to fulfill increased vitamin demands under conditions suited to achieving faster growth rates, while *F. cylindrus* instead utilizes *METE* under low-iron conditions to meet its methionine requirements. Future work is required to determine whether diatoms such as *Pseudo-nitzschia* can use their *COB* genes to convert pseudocobalamin to cobalamin when the necessary organic components are supplied, as has been shown for other eukaryotic microalgae (Helliwell et al., 2016), and to clarify the biochemical advantage of relying upon an energetically costly (Bertrand et al., 2013), yet cobalamin-independent, isoform of the methionine synthase enzyme (*METE*) under low iron conditions, as occurs in *F. cylindrus*.

In summary, using a combined transcriptomic and proteomic approach, we have identified several strategies utilized by the low-iron adapted diatom *P. granii* when reorganizing metabolic processes as a function of iron status. These gene expression and protein abundance patterns contribute to our understanding of how *P. granii* sustains relatively fast growth during periods of chronic iron limitation and rapidly restructures its cellular machinery in response to iron resupply to quickly become a dominate member of an iron-stimulated phytoplankton community.

## Experimental procedures

### Iron-replete and iron-limited culture conditions

The pennate diatom *Pseudo-nitzschia granii* (UNC1102) used in both transcriptomic and proteomic analyses was isolated from naturally iron-limited waters of the North-east Pacific Ocean at Ocean Station Papa (OSP; 50°N, 145°W) in 2011. The isolate was identified through sequencing of the 18S rDNA and ITS1 regions (accession numbers KJ866907 and EU051654 respectively). Growth data from *P. granii* UWOSP36, the oceanic diatom *T. oceanica* CCMP 1003 and coastal diatom *Pseudo-nitzschia multiseries* CLN-47 were retrieved from Marchetti *et al.* (2009), grown under the same culture conditions to serve as comparisons of growth rates under varying iron availability (Fig. 1).

Cultures were maintained in the artificial seawater medium Aquil using trace metal clean (TMC) techniques according to Marchetti *et al.* (2006). For media preparation, 2 liter (L) aliquots of Aquil were placed into acid-cleaned, Milli-Q (18.2 MΩ·cm) H<sub>2</sub>O-rinsed polycarbonate bottles. Dispensed seawater was supplemented with filter-sterilized (0.2-μm Acrodisc) ethylenediaminetetraacetic acid (EDTA)-trace metals (without iron) and vitamins (cobalamin, thiamine and biotin) according to Price *et al.* (1989). Trace metal concentrations were buffered using 100 μmol L<sup>-1</sup> of EDTA. Aquil medium macronutrient concentrations were 300 μmol L<sup>-1</sup> nitrate, 10 μmol L<sup>-1</sup> phosphate and 100 μmol L<sup>-1</sup> silicic acid.

For transcriptomic and proteomic analyses, *P. granii* cultures were grown in Aquil medium within 4 L acid-cleaned, Milli-Q H<sub>2</sub>O-rinsed modified polycarbonate flasks that permitted TMC subsampling. Iron-replete treatments (+Fe) were prepared by adding 1370 nmol L<sup>-1</sup> of total iron (Fe<sub>T</sub>) in a 1:1 Fe:EDTA solution to Aquil medium, corresponding to a dissolved inorganic iron concentration of 2.7 nmol L<sup>-1</sup> (Marchetti *et al.*, 2009). To growth limit *P. granii* by iron (−Fe), 200 nmol L<sup>-1</sup> of the iron chelator desferrioxamine B (DFB) was added to 1.5 nmol L<sup>-1</sup> Fe<sub>T</sub>, yielding an Fe concentration close to 0.02 pmol L<sup>-1</sup> (Strzepek *et al.*, 2012). The Fe:DFB solution was allowed to equilibrate overnight before being added and the Aquil medium was also left to equilibrate overnight before use. The large cultures were grown under a continuous, saturating photon flux density of 110 μmol photons m<sup>-2</sup> s<sup>-1</sup>. The length of experiments was typically between 5 and 15 days, depending on the iron treatment. *P. granii* cultures were maintained at 12°C and stirred for the duration of the experimental periods to prevent cells from settling. Subsamples were withdrawn daily under a TMC laminar flow hood to monitor growth rates and changes in cell density, as well as to ensure cells remained in the exponential growth phase throughout the experimental period.

Iron status was determined in the *P. granii* experiments through measuring growth and maximum photochemical yields of PSII ( $F_v:F_m$ ). Growth rates were calculated by measuring changes in the natural log of *in vivo* fluorescence over time using a Turner Designs model 10-AU fluorometer, as described in (Brand *et al.*, 1981).  $F_v:F_m$  was obtained by measuring fluorescence induction with a Satlantic FRe fluorometer (Gorbunov and Falkowski 2005). Before each measurement, a subsample (5 ml) of each culture was placed in the dark for 20 min. A short, saturating pulse of blue light (450 nm) was applied to dark-acclimated phytoplankton for a 100 μs duration to measure fluorescence induction from minimum ( $F_o$ ) to maximum ( $F_m$ ) fluorescence yields.  $F_o$  and  $F_m$  were used to estimate  $F_v:F_m$  from  $(F_m - F_o) F_m^{-1}$ . Values were calculated for each culture upon harvesting during the mid-exponential growth phase. Measurements were performed using multiple-turnover flash (MTF).

### Transcriptomic analysis

*Pseudo-nitzschia granii* cells in 1 L of culture medium were harvested for RNA analysis from both iron-replete and iron-limited conditions while in late-exponential phase. Cells were directly collected onto 3 μm polycarbonate filters (Millipore; 25 mm) using vacuum filtration and immediately stored at -80°C until RNA extractions were performed using the RNAqueous 4-PCR kit (Ambion) following manufacturer's protocols. To remove DNA contamination, RNA was incubated with deoxyribonuclease (DNase) I at 37°C for 45 min and purified with DNase inactivation reagent (Ambion). RNA concentrations were measured using a NanoDrop spectrophotometer, and samples containing concentrations < 100 ng μl<sup>-1</sup> were concentrated using the RNeasy MinElute Cleanup Kit (Qiagen). Quantitative PCR (qPCR) with the RNA samples was performed to test for DNA contamination levels, and samples with amplification of DNA products were given an additional round of the DNase incubation.

For transcriptomic sequencing, total RNA was prepared with the Illumina TruSeq Stranded mRNA Library Preparation Kit. RNA from three replicate cultures of iron-replete treatments and two replicate cultures of iron-limited treatments were separately prepared for sequencing. Barcoded samples were sequenced on a single shared lane of Illumina MiSeq (150 bp, paired-end). Sequencing statistics are provided in Supplemental Table S1. The *P. granii* RNA-seq data reported here has been deposited in the National Center for Biotechnology (NCBI) sequence read archive under accession number SRP131929.

Raw reads were trimmed of poor quality bases and adapters using Trimmomatic v0.32 (paired-end mode, adaptive quality trim with 40 bp target length and

strictness of 0.6, minimum length of 36 bp; Bolger *et al.*, 2014). Trimmed reads from both treatments were assembled into a single *de novo* transcriptome using Trinity v2.2.0 (Grabherr *et al.*, 2011). Read counts were obtained by mapping raw reads to assembled contigs with bowtie2 v2.2.6 (end-to-end alignment; Langmead and Salzberg, 2012). Protein-coding regions were predicted using GeneMark-S-T (Tang *et al.*, 2015). Functional annotations were assigned based on sequence homology to proteins in the Kyoto Encyclopedia of Genes and Genomes database (KEGG; Release 75) via BLASTp v2.3.0 with an Expect (*E*)-value cutoff of  $10^{-3}$  (Altschul *et al.*, 1990).

All read counts corresponding to identical functional genes, or KEGG Orthology (KO) entries, were summed together. For genes without KO assignments but possessing an annotated gene definition according to KEGG (e.g., *ISIPs*, *RHO*), read counts corresponding to these definitions were summed. For pathway-level analysis, read counts corresponding to KOs contained within each KEGG module were summed. EdgeR v3.12.0 was used to calculate normalized counts, fold change and counts-per-million (CPM; estimated transcript abundance) from pairwise comparisons between iron-replete and iron-limited samples. A general linear model was used to estimate the tag-wise common dispersion followed by a likelihood ratio test to assess differential expression (Robinson and Smyth, 2008; Robinson *et al.*, 2010). Genes determined to be differentially expressed ( $p < 0.05$ ) were adjusted for false positives using the Benjamini-Hochberg multiple correction test in edgeR (FDR < 0.05).

The *P. granii* transcriptomes under iron-replete and iron-limited conditions were compared to those of *Thalassiosira oceanica* CCMP1005 and *Fragilariopsis cylindrus* CCMP1102. The publicly available iron-replete *T. oceanica* assembled transcriptome, generated using Roche 454 pyrosequencing, was obtained from the NCBI (bioproject accession number PRJNA73029; Lommer *et al.*, 2012). *T. oceanica* raw reads corresponding to iron-replete and iron-limited cultures were obtained from NCBI using accession numbers SRX096487 and SRX096488. The *F. cylindrus* transcriptome was assembled using publicly available raw reads from iron-replete and iron-limited culture experiments, generated using Illumina HiSeq 2000, and obtained through ArrayExpress (sample number E-MTAB-5024; Mock *et al.*, 2017). Similar to *P. granii*, *de novo* assemblies were created with *F. cylindrus* reads by first trimming with Trimmomatic v.0.32 and later assembling using Trinity v2.2.0. To generate quantitative count information for *T. oceanica* and *F. cylindrus*, raw reads from both iron treatments were mapped to respective transcriptomes using the aligner bowtie2. Transcriptomes were assigned functional

annotation based on sequence homology to proteins in the KEGG database via BLASTx v2.3.0 with an *E*-value cutoff of  $10^{-3}$ . Differential expression in *F. cylindrus* was assessed using an identical method as used with *P. granii*: the generalized linear model approach with estimated tagwise dispersions and a likelihood ratio test. Statistical significance in differential expression between iron treatments in *T. oceanica* was not estimated due to lack of biological replication and therefore only a qualitative analysis was performed. R packages (R Foundation for Statistical Computing, 2016) used for processing of bioinformatic data and visualization include SeqKit (Shen *et al.*, 2016), caroline, VennDiagram (Chen and Boutros, 2011), ggplot2 (Wickham, 2009), pheatmap, RColorBrewer (Neuwirth, 2014) and viridis. *Pseudo-nitzschia granii* assembled contigs, read counts and KEGG annotations are available at marchettilab.web.unc.edu/data.

#### Global proteomic approach

*Pseudo-nitzschia granii* cultures were grown in duplicate 4 L acid-cleaned polycarbonate bottles under acclimated iron-replete and iron-limited conditions using the culture methodology described previously. After reaching mid to late exponential growth phase, all cells were filtered onto 0.45 µm polyethersulphone filters (0.45 µm, 47 mm), re-suspended in 5 ml of Aquil filtrate, and centrifuged (12 000 rpm) into pellets that were immediately frozen at  $-80^{\circ}\text{C}$  until analysis.

Protein was extracted from cell pellets in 4 ml of SDS Lysis Buffer (1% SDS, 0.1 M Tris/HCL (pH 7.5), 10 mM EDTA) and heated at  $95^{\circ}\text{C}$  for 10 min. Cell debris was pelleted at 2057 rcf for 30 min and transferred supernatant was concentrated to 400 µl using 5000 MWCO Vivaspinn filtration units (Sartorius, UK). Concentrated proteins were precipitated overnight at  $-20^{\circ}\text{C}$  in 50% MeOH, 50% acetone and 0.5 mM HCl at a ratio of 1:4 protein: MeOH/acetone. Precipitated proteins were centrifuged at 14 100 rcf, dried for 10 min in a Thermo DNA 110 SpeedVac, and resuspended in SDS lysis buffer. Protein yield was determined using the BioRad DC Protein Assay. Protein digestion was performed using a modified tube gel method according to Bender *et al.*, 2018. Briefly, protein was reduced, alkylated and digested with a 1:20 trypsin: protein ratio. Tryptic peptides were analysed via liquid chromatography tandem mass spectrometry (LC/MS/MS) using a Michrom Advance HPLC system with reverse phase chromatography coupled to a Thermo Scientific Q-Exactive Orbitrap ion-trap mass spectrometer with a Michrom Advance CaptiveSpray source (1 µl min<sup>-1</sup> flowrate, C18 column). Each sample was concentrated onto a trap column (0.2 x 10 mm ID, 5 µm particle size, 120 Å pore size, C18 Reprosil-Gold, Dr. Maisch GmbH) and rinsed with 100 µl 0.1% formic acid, 2% acetonitrile



(ACN), 97.9% water before gradient elution through a reverse phase C18 column (0.1 x 250 mm ID, 3 µm particle size, 120 Å pore size, C18 Reprosil-Gold, Dr. Maisch GmbH) at a flow rate of 500 nL min<sup>-1</sup>. The chromatography consisted of a nonlinear 150 min gradient from 5% to 95% buffer B, where A was 0.1% formic acid in water and B was 0.1% formic acid in ACN (all solvents being Fisher Optima grade). The mass spectrometer monitored MS1 scans from 380 to 1580 m/z at 70 K resolution. MS2 scans were performed on the top 10 ions with an isolation window of 2.0 m/z and a 15 s exclusion time.

Mass spectra were searched against the *P. granii* translated transcriptome using Proteome Discoverer's SEQUEST algorithm (Thermo) with a fragment tolerance of 0.02 Da and parent tolerance of 10 ppm. Identification criteria consisted of a peptide threshold of 91% and protein threshold of 99% (two peptides minimum), corresponding to a false discovery rate (protein FDR) of 1.0%. Visualization and processing of MS/MS proteomics data were performed using Scaffold 3 (version 4.8.4, Proteome Software). In cases of nonunique reference contigs where protein sequences completely overlap, Scaffold assigns peptides and associated spectral counts to only one protein. Spectral counts were divided by protein length to yield normalized spectral abundance factor scores and were subsequently normalized by total counts per sample in Scaffold (NSAF). Similar to the transcriptomic analysis, all NSAF scores associated with contigs having identical KO annotations were summed together. The *P. granii* mass spectrometry proteomics data has been deposited to the ProteomeXchange Consortium through the PRIDE repository (Vizcaíno *et al.*, 2016) under accession number PXD009698.

Differential abundance of proteins between iron treatments was determined using the PLGEM R package in step-by-step mode with normalized spectral counts (NSAF; Pavelka *et al.*, 2004, 2008). PLGEM describes the relationship between the standard deviation and mean of the normalized spectral counts, and in our analysis the model fit the data moderately well with a slope of 0.61, adjusted  $r^2$  of 0.96 and Pearson correlation of 0.73. Proteins were considered significantly differentially abundant between treatments if estimated false positive rates, based on  $p$ -values calculated from signal-to-noise ratios, were below 0.05. In our study with 805 proteins considered, an estimated FPR of 0.05 may result in up to ~40 proteins falsely labelled differentially abundant due to chance (Pavelka *et al.*, 2008). This method has previously been applied to marine laboratory proteomic experiments (Walworth *et al.*, 2016; Saunders *et al.*, 2017). KEGG-annotated proteins (KOs) with normalized spectral counts of zero were artificially increased to a trivial value of 0.5 to allow for fold-change estimations between treatments.

The one dimensional chromatographic approach applied to this study was suitable for examining changes in protein abundance between treatments using spectral counting. Future efforts for increased depth of protein detection is also possible using two dimensional chromatographic approaches if depth of coverage is the primary goal. The global proteomic approach is most useful in comparing relative protein abundance between treatments and caution should be taken when comparing absolute abundances across proteins. Tryptic peptide biases are introduced during protein digestion that may result in large proteins with many tryptic peptides having higher counts associated with them as compared to similar abundances of small proteins with fewer tryptic peptides. To more accurately compare concentrations of different proteins within a treatment, a targeted approach is recommended which relies upon isotopically-labelled peptide standards (Saito *et al.*, 2014; Mackey *et al.*, 2015).

## Acknowledgements

J. Strauss (EMBL) provided growth data corresponding to *F. cylindrus* culture experiments from Mock *et al.*, (2017). E. Bertrand (DAL) provided the *Pseudo-nitzschia heimii* CBA1 reference sequence. We thank W. Sunda, B. McGregor, M. Alperin, H. Paerl, K. Ellis and C. Moreno (UNC) for their feedback on earlier versions of the manuscript. Two anonymous reviewers offered valuable and constructive comments. All RNA-Seq data was processed using UNC's Research Computing clusters. This research was funded through National Science Foundation (NSF) grant OCE1334935 to A.M. Protein analyses were supported by the Gordon Betty Moore Foundation grant 3782 to M.A.S. N.R.C. was partially supported by the UNC Graduate School and by a grant from the Simons Foundation/SFARI (544236, N.R.C) during the writing of the manuscript.

## Author contributions

NRC and AM designed the study. NRC conducted the *P. granii* culture experiments, performed RNA extractions, processed bioinformatic data and analysed the results. WG assisted with the transcriptomic pipeline. DMM performed protein extractions and MRM conducted the LC/MS/MS proteomic analysis under the guidance of MAS. NRC and AM wrote the manuscript with input from all coauthors.

## References

- Alexander, H., Rouco, M., Haley, S. T., Wilson, S. T., Karl, D. M., and Dyhrman, S. T. (2015) Functional group-specific traits drive phytoplankton dynamics in the oligotrophic ocean. *Proc Natl Acad Sci U S A* **112**: E5972–E5979.

- Allen, A. E., LaRoche, J., Maheswari, U., Lommer, M., Schauer, N., Lopez, P. J., et al. (2008) Whole-cell response of the pennate diatom *Phaeodactylum tricornutum* to iron starvation. *Proc Natl Acad Sci U S A* **105**: 10438–10443.
- Allen, A. E., Dupont, C. L., Oborník, M., Horák, A., Nunes-Nesi, A., McCrow, J. P., et al. (2011) Evolution and metabolic significance of the urea cycle in photosynthetic diatoms. *Nature* **473**: 203.
- Allen, A. E., Moustafa, A., Montsant, A., Eckert, A., Kroth, P. G., and Bowler, C. (2012) Evolution and functional diversification of fructose bisphosphate aldolase genes in photosynthetic marine diatoms. *Mol Biol Evol* **29**: 367–379.
- Altschul, S. F., Gish, W., Miller, W., Myers, E. W., and Lipman, D. J. (1990) Basic local alignment search tool. *J Mol Biol* **215**: 403–410.
- Arrigo, K. R., Mills, M. M., Kropuenske, L. R., van Dijken, G. L., Alderkamp, A.-C., and Robinson, D. H. (2010) Photo-physiology in Two Major Southern Ocean Phytoplankton Taxa: Photosynthesis and Growth of *Phaeocystis antarctica* and *Fragilariopsis cylindrus* under Different Irradiance Levels. *Integr Comp Biol* **50**: 950–966.
- Asada, K. (2006) Production and scavenging of reactive oxygen species in chloroplasts and their functions. *Plant Physiol* **141**: 391–396.
- de Baar, H. J. W., Boyd, P. W., Coale, K. H., Landry, M. R., Tsuda, A., Assmy, P., et al. (2005) Synthesis of iron fertilization experiments: from the iron age in the age of enlightenment. *J Geophys Res C Ocean* **110**: 1–24.
- Bertini, I., Luchinat, C., Provenzano, A., Rosato, A., and Vasos, P. R. (2002) Browsing gene banks for Fe<sub>2</sub>S<sub>2</sub> ferredoxins and structural modeling of 88 plant-type sequences: An analysis of fold and function. *Proteins* **46**: 110–127.
- Beja, O., Aravind, L., Koonin, E. V., Suzuki, M. T., Hadd, A., Nguyen, L. P., et al. (2000) Bacterial rhodopsin: evidence for a new type of phototrophy in the sea. *Science* **289**: 1902–1906.
- Bender, S., Durkin, C., Berthiaume, C., Morales, R., and Armbrust, E. V. (2014) Transcriptional responses of three model diatoms to nitrate limitation of growth. *Front Mar Sci* **1**: 3.
- Bender, S. J., Moran, D. M., McIlvin, M. R., Zheng, H., McCrow, J. P., Badger, J., et al. (2018) Iron triggers colony formation in *Phaeocystis antarctica*: connecting molecular mechanisms with iron biogeochemistry. *Biogeosciences* **2018**: 1–52.
- Bertrand, E. M., Allen, A. E., Dupont, C. L., Norden-krichmar, T. M., Bai, J., and Valas, R. E. (2012) Influence of cobalamin scarcity on diatom molecular physiology and identification of a cobalamin acquisition protein. *Proc Natl Acad Sci U S A* **109**: 1762–1771.
- Bertrand, E. M., Moran, D. M., McIlvin, M. R., Hoffman, J. M., Allen, A. E., and Saito, M. A. (2013) Methionine synthase interreplacement in diatom cultures and communities: implications for the persistence of B12 use by eukaryotic phytoplankton. *Limnol Oceanogr* **58**: 1431–1450.
- Bertrand, E. M., McCrow, J. P., Moustafa, A., Zheng, H., McQuaid, J. B., Delmont, T. O., et al. (2015) Phytoplankton-bacterial interactions mediate micronutrient colimitation at the coastal Antarctic Sea ice edge. *Proc Natl Acad Sci U S A* **112**: 9938–9943.
- Bertrand, E. M., Saito, M. A., Lee, P. A., Dunbar, R. B., Sedwick, P. N., and Ditullio, G. R. (2011) Iron limitation of a springtime bacterial and phytoplankton community in the Ross Sea: implications for vitamin B12 nutrition. *Front Microbiol* **2**: 1–12.
- Bolger, A. M., Lohse, M., and Usadel, B. (2014) Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**: 2114–2120.
- Boyd, P., and Harrison, P. J. (1999) Phytoplankton dynamics in the NE subarctic Pacific. *Deep Res Part II Top Stud Oceanogr* **46**: 2405–2432.
- Boyd, P. W., Jickells, T., Law, C. S., Blain, S., Boyle, E. A., Buesseler, K. O., et al. (2007) Mesoscale iron enrichment experiments 1993–2005: synthesis and future directions. *Science* **315**: 612–618.
- Brand, L. E., Guillard, R. R. L., and Murphy, L. S. (1981) A method for the rapid and precise determination of acclimated phytoplankton reproduction rates. *J Plankt Res* **3**: 193–201.
- Browning, T. J., Achterberg, E. P., Rapp, I., Engel, A., Bertrand, E. M., Tagliabue, A., et al. (2017) Nutrient co-limitation at the boundary of an oceanic gyre. *Nature* **551**: 242.
- Carradec, Q., Pelletier, E., Da Silva, C., Alberti, A., Seeleuthner, Y., Blanc-Mathieu, R., et al. (2018) A global ocean atlas of eukaryotic genes. *Nat Commun* **9**: 373.
- Chappell, P. D., Whitney, L. P., Wallace, J. R., Darer, A. I., Jean-Charles, S., and Jenkins, B. D. (2015) Genetic indicators of iron limitation in wild populations of *Thalassiosira oceanica* from the northeast Pacific Ocean. *ISME J* **9**: 592–602.
- Chen, H., and Boutros, P. C. (2011) VennDiagram: a package for the generation of highly-customizable Venn and Euler diagrams in R. *BMC Bioinformatics* **12**: 35.
- Cohen, N. R., Ellis, K. A., Burns, W. G., Lampe, R. H., Schuback, N., Johnson, Z., et al. (2017a) Iron and vitamin interactions in marine diatom isolates and natural assemblages of the Northeast Pacific Ocean. *Limnol Oceanogr* **62**: 2076–2096.
- Cohen, N. R., Ellis, K. A., Lampe, R. H., McNair, H., Twining, B. S., Maldonado, M. T., et al. (2017b) Diatom transcriptional and physiological responses to changes in iron bioavailability across ocean provinces. *Front Mar Sci* **4**: 360.
- Cohen, N. R., Elizabeth, M., Brooke, S., Moreno, C. M., Rauschenberg, S., Jacquot, J. E., et al. (2018) Iron storage capacities and associated ferritin gene expression among marine diatoms. *Limnol Oceanogr* **63**: 1677–1691.
- Doucette, G., Erdner, D., Peleato, M., Hartman, J., and DM, A. (1996) Quantitative analysis of iron-stress related proteins in *Thalassiosira weissflogii*: measurement of flavodoxin and ferredoxin using HPLC. *Mar Ecol Prog Ser* **130**: 269–276.
- Dyhrman, S. T., Jenkins, B. D., Rynearson, T. A., Saito, M. A., Mercier, M. L., Alexander, H., et al. (2012) The transcriptome and proteome of the diatom *Thalassiosira pseudonana* reveal a diverse phosphorus stress response. *PLoS One* **7**: 1–10.

- Edgar, R. C. (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res* **32**: 1792–1797.
- Ellis, K. A., Cohen, N. R., Moreno, C., and Marchetti, A. (2017) Cobalamin-independent methionine synthase distribution and influence on vitamin B12 growth requirements in marine diatoms. *Protist* **168**: 32–47.
- Erdner, D. L., and Anderson, D. M. (1999) Ferredoxin and flavodoxin as biochemical indicators of iron limitation during open-ocean iron enrichment. *Limnol Oceanogr* **44**: 1609–1615.
- Godhe, A., and Rynearson, T. (2017) The role of intraspecific variation in the ecological and evolutionary success of diatoms in changing environments. *Philos Trans R Soc B Biol Sci* **372**: 20160399.
- Gorbunov, M. Y., and Falkowski, P. (2004) Fluorescence Induction and Relaxation (FIRe) technique and instrumentation for monitoring photosynthetic processes and primary production in aquatic ecosystems. In *13th International Congress of Photosynthesis*, Vol. 2, van der Est, A., and Bruce, D. (eds). Montreal, QC: Allen Press, pp. 1029–1031.
- Grabherr, M. G., Haas, B. J., Yassour, M., Levin, J. Z., Thompson, D. A., Amit, I., et al. (2011) Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nat Biotech* **29**: 644–652.
- Groussman, R. D., Parker, M. S., and Armbrust, E. V. (2015) Diversity and evolutionary history of iron metabolism genes in diatoms. *PLoS One* **10**: e0129081.
- Gutowska, M. A., Shome, B., Sudek, S., McRose, D. L., Hamilton, M., Giovannoni, S. J., et al. (2017) Globally important haptophyte algae use exogenous pyrimidine compounds more efficiently than thiamin. *MBio* **8**: e01459–e01417.
- Hanke, G., and Mulo, P. (2012) Plant type ferredoxins and ferredoxin-dependent metabolism. *Plant Cell Environ* **36**: 1071–1084.
- Hasle, G. R. (2002) Are most of the domoic acid-producing species of the diatom genus *Pseudo-nitzschia* cosmopolites? *Harmful Algae* **1**: 137–146.
- Heal, K. R., Qin, W., Ribalet, F., Bertagnolli, A. D., Coyote-Maestas, W., Hmelo, L. R., et al. (2017) Two distinct pools of B12 analogs reveal community interdependencies in the ocean. *Proc Natl Acad Sci U S A* **114**: 364–369.
- Helliwell, K. E., Lawrence, A. D., Holzer, A., Kudahl, U. J., Sasso, S., Kräutler, B., et al. (2016) Cyanobacteria and eukaryotic algae use different chemical variants of vitamin B12. *Curr Biol* **26**: 999–1008.
- Hippmann, A. A., Schuback, N., Moon, K. M., McCrow, J. P., Allen, A. E., Foster, L. J., et al. (2017) Contrasting effects of copper limitation on the photosynthetic apparatus in two strains of the open ocean diatom *Thalassiosira oceanica*. *PLOS ONE* **12**: e0181753.
- Hockin, N. L., Mock, T., Mulholland, F., Kopriva, S., and Malin, G. (2012) The response of diatom central carbon metabolism to nitrogen starvation is different from that of green algae and higher plants. *Plant Physiol* **158**: 299–312.
- Hostetter, H. P., and Hoshaw, R. W. (1972) Asexual developmental patterns of the diatom *Stauroneis anceps* in culture. *J Phycol* **8**: 289–296.
- Horecker, B. L., Tsolas, O., and Lai, C. Y. (1972) 6 aldolases. In *The Enzymes*, Boyer, P. D. (ed). New York: Academic Press, pp. 213–258.
- Hunter, T., Bonetta, R., Sacco, A., Marita, V., Sultana, P., Trinh, C. H., et al. (2017) A single mutation is sufficient to modify the metal selectivity and specificity of a eukaryotic manganese superoxide dismutase to encompass iron. *Chem: A Eur J* **24**: 5303–5308.
- Hutchins, D. A., and Fu, F. (2017) Microorganisms and ocean global change. *Nat Microbiol* **2**: 17058.
- Jeanjean, R., Latifi, A., Matthijs, H. C. P., and Havaux, M. (2008) The PsaE subunit of photosystem I prevents light-induced formation of reduced oxygen species in the cyanobacterium *Synechocystis* sp. PCC 6803. *Biochim Biophys Acta - Bioenerg* **1777**: 308–316.
- Katoh, H., Hagino, N., Grossman, A. R., and Ogawa, T. (2001) Genes essential to iron transport in the cyanobacterium *Synechocystis* sp. strain PCC 6803. *J Bacteriol* **183**: 2779–2784.
- Kazamia, E., Sutak, R., Paz-Yepes, J., Dorrell, R. G., Vieira, F. R. J., Mach, J., et al. (2018) Endocytosis-mediated siderophore uptake as a strategy for Fe acquisition in diatoms. *Sci Adv* **4**: eaar4536.
- Kelley, L. A., Mezulis, S., Yates, C. M., Wass, M. N., and Sternberg, M. J. E. (2015) The Phyre2 web portal for protein modeling, prediction and analysis. *Nat Protoc* **10**: 845.
- Lam, P. J., and Bishop, J. K. B. (2008) The continental margin is a key source of iron to the HNLC North Pacific Ocean. *Geophys Res Lett* **35**: 1–5.
- Lam, P. J., Bishop, J. K. B., Henning, C. C., Marcus, M. A., Waychunas, G. A., and Fung, I. Y. (2006) Wintertime phytoplankton bloom in the subarctic Pacific supported by continental margin iron. *Global Biogeochem Cycles* **20**: 1–12.
- Langmead, B., and Salzberg, S. L. (2012) Fast gapped-read alignment with bowtie 2. *Nat Methods* **9**: 357–359.
- LaRoche, J., Murray, H., Orellana, M., and Newton, J. (1995) Flavodoxin expression as an indicator of iron limitation in marine diatoms. *J Phycol* **31**: 520–530.
- Lelong, A., Hégaret, H., Soudant, P., and Bates, S. S. (2012) *Pseudo-nitzschia* (Bacillariophyceae) species, domoic acid and amnesic shellfish poisoning: revisiting previous paradigms. *Phycologia* **51**: 168–216.
- Lin, Y.-H., Pan, K.-Y., Hung, C.-H., Huang, H.-E., Chen, C.-L., Feng, T.-Y., et al. (2013) Overexpression of Ferredoxin, PETF, enhances tolerance to heat stress in *Chlamydomonas reinhardtii*. *Int J Mol Sci* **14**: 20913–20929.
- Lommer, M., Specht, M., Roy, A.-S., Kraemer, L., Andreson, R., Gutowska, M. A., et al. (2012) Genome and low-iron response of an oceanic diatom adapted to chronic iron limitation. *Genome Biol* **13**: R66.
- Luo, C.-S., Liang, J.-R., Lin, Q., Li, C., Bowler, C., Anderson, D. M., et al. (2014) Cellular responses associated with ROS production and cell fate decision in early stress response to iron limitation in the diatom *Thalassiosira pseudonana*. *J Proteome Res* **13**: 5510–5523.
- Mackey, K. R. M., Post, A. F., McIlvin, M. R., Cutter, G. A., John, S. G., and Saito, M. A. (2015) Divergent responses of Atlantic coastal and oceanic *Synechococcus* to iron limitation. *Proc Natl Acad Sci U S A* **112**: 9944–9949.



- Marchetti, A., Maldonado, M. T., Lane, E. S., and Harrison, P. J. (2006) Iron requirements of the pennate diatom pseudo-nitzschia: comparison of oceanic (high-nitrate, low-chlorophyll waters) and coastal species. *Limnol Oceanogr* **51**: 2092–2101.
- Marchetti, A., Parker, M. S., Moccia, L. P., Lin, E. O., Arrieta, A. L., Ribalet, F., et al. (2009) Ferritin is used for iron storage in bloom-forming marine pennate diatoms. *Nature* **457**: 467–470.
- Marchetti, A., Schrueth, D. M., Durkin, C. A., Parker, M. S., Kodner, R. B., Berthiaume, C. T., et al. (2012) Comparative metatranscriptomics identifies molecular bases for the physiological responses of phytoplankton to varying iron availability. *Proc Natl Acad Sci U S A* **109**: E317–E325.
- Marchetti, A., Catlett, D., Hopkinson, B. M., Ellis, K., and Cassar, N. (2015) Marine diatom proteorhodopsins and their potential role in coping with low iron availability. *ISME J* **9**: 2745–2748.
- Marchetti, A., Moreno, C. M., Cohen, N. R., Oleinikov, I., deLong, K., Twining, B. S., et al. (2017) Development of a molecular-based index for assessing iron status in bloom-forming pennate diatoms. *J Phyco* **53**: 820–832.
- McKay, R. M. L., La Roche, J., Yakunin, A. F., Durnford, D. G., and Geider, R. J. (1999) Accumulation of ferredoxin and flavodoxin in a marine diatom in response to Fe. *J Phycol* **35**: 510–519.
- McQuaid, J. B., Kustka, A. B., Oborník, M., Horák, A., McCrow, J. P., Karas, B. J., et al. (2018) Carbonate-sensitive phytoferritin controls high-affinity iron uptake in diatoms. *Nature* **555**: 534.
- McRose, D., Guo, J., Monier, A., Sudek, S., Wilken, S., Yan, S., et al. (2014) Alternatives to vitamin B1 uptake revealed with discovery of riboswitches in multiple marine eukaryotic lineages. *ISME J* **8**: 2517–2529.
- Mihara, H., and Esaki, N. (2002) Bacterial cysteine desulfurases: their function and mechanisms. *Appl Microbiol Biotechnol* **60**: 12–23.
- Mock, T., Otilar, R. P., Strauss, J., McMullan, M., Paajanen, P., Schmutz, J., et al. (2017) Evolutionary genomics of the cold-adapted diatom *Fragilariopsis cylindrus*. *Nature* **541**: 536–540.
- Moore, J. K., Doney, S. C., Glover, D. M., and Fung, I. Y. (2002) Iron cycling and nutrient-limitation patterns in surface waters of the World Ocean. *Deep Res II* **49**: 463–507.
- Moore, C. M., Mills, M. M., MILNE, A., Langlois, R., Achterberg, E. P., Lochte, K., et al. (2006) Iron limits primary productivity during spring bloom development in the Central North Atlantic. *Glob Chang Biol* **12**: 626–634.
- Morrissey, J., Satak, R., Paz-Yepes, J., Tanaka, A., Moustafa, A., Veluchamy, A., et al. (2015) A novel protein, ubiquitous in marine phytoplankton, concentrates iron at the cell surface and facilitates uptake. *Curr Biol* **25**: 364–371.
- Neuwirth, E. (2014) RColorBrewer: ColorBrewer palettes. *R Packag version 1.1–2*. <https://cran.R-project.org/package=RColorBrewer>
- Nunn, B. L., Faux, J. F., Hippmann, A. A., Maldonado, M. T., Harvey, H. R., Goodlett, D. R., et al. (2013) Diatom proteomics reveals unique acclimation strategies to mitigate Fe limitation. *PLoS One* **8**: e75653.
- Olgun, N., Duggen, S., Andronico, D., Kutterolf, S., Croot, P. L., Giammanco, S., et al. (2013) Possible impacts of volcanic ash emissions of Mount Etna on the primary productivity in the oligotrophic Mediterranean Sea: Results from nutrient-release experiments in seawater. *Mar Chem* **152**: 32–42.
- Paajanen, P., Strauss, J., Van Oosterhout, C., McMullan, M., Clark, M. D., and Mock, T. (2017) Data descriptor: building a locally diploid genome and transcriptome of the diatom *Fragilariopsis cylindrus*. *Sci Data* **4**: 170149.
- Paerl, R. W., Bouget, F.-Y., Lozano, J.-C., Vergé, V., Schatt, P., Allen, E. E., et al. (2016) Use of plankton-derived vitamin B1 precursors, especially thiazole-related precursor, by key marine picoeukaryotic phytoplankton. *ISME J* **11**: 753.
- Pavelka, N., Pelizzola, M., Vizzardelli, C., Capozzoli, M., Splendiani, A., Granucci, F., and Ricciardi-Castagnoli, P. (2004) A power law global error model for the identification of differentially expressed genes in microarray data. *BMC Bioinformatics* **5**: 203.
- Pavelka, N., Fournier, M. L., Swanson, S. K., Pelizzola, M., Ricciardi-Castagnoli, P., Florens, L., and Washburn, M. P. (2008) Statistical similarities between transcriptomics and quantitative shotgun proteomics data. *Mol Cell Proteomics* **7**: 631–644.
- Pearson, G. A., Lago-Leston, A., Cánovas, F., Cox, C. J., Verret, F., Lasternas, S., et al. (2015) Metatranscriptomes reveal functional variation in diatom communities from the Antarctic peninsula. *ISME J* **9**: 2275–2289.
- Peers, G., and Price, N. M. (2006) Copper-containing plastocyanin used for electron transport by an oceanic diatom. *Nature* **441**: 341–344.
- Picaud, S., Kavanagh, K. L., Yue, W. W., et al. (2011) Structural basis of fumarate hydratase deficiency. *J Inher Metab Dis* **34**: 671–676.
- Pourcel, L., Moulin, M., and Fitzpatrick, T. B. (2013) Examining strategies to facilitate vitamin B(1) biofortification of plants by genetic engineering. *Front Plant Sci* **4**: 160.
- Price, N. M., Harrison, G. I., Hering, J. G., Hudson, R. J., Nirel, P. M. V., Palenik, B., and Morel, F. M. M. (1989) Preparation and chemistry of the artificial algal culture medium Aquil. *Biol Oceanogr* **6**: 443–461.
- R Foundation for Statistical Computing (2016) R: A language and environment for statistical computing *R A Lang Environ Stat Comput* 3.3.1.
- Ribalet, F., Marchetti, A., Hubbard, K. A., Brown, K., Durkin, C. A., Morales, R., et al. (2010) Unveiling a phytoplankton hotspot at a narrow boundary between coastal and offshore waters. *Proc Natl Acad Sci U S A* **107**: 16571–16576.
- Robinson, M. D., and Smyth, G. K. (2008) Small-sample estimation of negative binomial dispersion, with applications to SAGE data. *Biostatistics* **9**: 321–332.
- Robinson, M. D., McCarthy, D. J., and Smyth, G. K. (2010) edgeR: a bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26**: 139–140.
- Rodionov, D. A., Vitreschak, A. G., Mironov, A. A., and Gelfand, M. S. (2003) Comparative genomics of the vitamin B12 metabolism and regulation in prokaryotes. *J Biol Chem* **278**: 41148–41159.
- Rueler, J. G., and Ades, D. R. (1987) The role of iron nutrition in photosynthesis and nitrogen assimilation in



- Scenedesmus quadricauda* (Chlorophyceae) 1. *J Phycol* **23**: 452–457.
- Saito, M. A., McIlvin, M. R., Moran, D. M., Goepfert, T. J., DiTullio, G. R., Post, A. F., and Lamborg, C. H. (2014) Multiple nutrient stresses at intersecting Pacific Ocean biomes detected by protein biomarkers. *Science* **345**: 1173–1177.
- Sañudo-Wilhelmy, S. A., Gómez-Consarnau, L., Suffridge, C., and Webb, E. A. (2014) The role of B vitamins in marine biogeochemistry. *Ann Rev Mar Sci* **6**: 339–367.
- Saunders, J. K. (2017) *Evolutionary response of marine bacteria to the co-occurring pnictogens phosphorus and arsenic (Doctoral dissertation)*: University of Washington. <https://digital.lib.washington.edu/researchworks/handle/1773/38191>.
- Schilling, B., Rardin, M. J., MacLean, B. X., Zawadzka, A. M., Frewen, B. E., Cusack, M. P., et al. (2012) Platform-independent and label-free quantitation of proteomic data using MS1 extracted ion chromatograms in skyline. *Mol Cell Proteomics* **11**: 202–214.
- Shen, W., Le, S., Li, Y., and Hu, F. (2016) SeqKit: a cross-platform and ultrafast toolkit for FASTA/Q file manipulation. *PLoS One* **11**: e0163962.
- Simão, F. A., Waterhouse, R. M., Ioannidis, P., Kriventseva, E. V., and Zdobnov, E. M. (2015) BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics* **31**: 3210–3212.
- Smetacek, V., Klaas, C., Strass, V. H., Assmy, P., Montresor, M., Cisewski, B., et al. (2012) Deep carbon export from a Southern Ocean iron-fertilized diatom bloom. *Nature* **487**: 313–319.
- Smith, D. R. (2013) RNA-Seq data: a goldmine for organelle research. *Brief Funct Genomics* **12**: 454–456.
- Strauss, J. (2012) A genomic analysis using RNA-Seq to investigate the adaptation of the psychrophilic diatom *Fragilariopsis cylindrus* to the polar environment.
- Strzepek, R. F., Hunter, K. A., Frew, R. D., Harrison, P. J., and Boyd, P. W. (2012) Iron – light interactions differ in Southern Ocean phytoplankton. *Limnol Oceanogr* **57**: 1182–1200.
- Tang, Y. Z., Koch, F., and Gobler, C. J. (2010) Most harmful algal bloom species are vitamin B 1 and B 12 auxotrophs. *Proc Natl Acad Sci U S A* **107**: 20756–20761.
- Strzepek, R. F., and Harrison, P. J. (2004) Photosynthetic architecture differs in coastal and oceanic diatoms. *Nature* **431**: 689–692.
- Tang, S., Lomsadze, A., and Borodovsky, M. (2015) Identification of protein coding regions in RNA transcripts. *Nucleic Acids Res* **43**: e78.
- Vizcaíno, J. A., Csordas, A., Del-Toro, N., Dianes, J. A., Griss, J., Lavidas, I., et al. (2016) 2016 update of the PRIDE database and its related tools. *Nucleic Acids Res* **44**: D447–D456.
- Walworth, N. G., Fu, F. X., Webb, E. A., Saito, M. A., Moran, D., McIlvin, M. R., et al. (2016) Mechanisms of increased *Trichodesmium* fitness under iron and phosphorus co-limitation in the present and future ocean. *Nat Commun* **7**: 1208.
- Westerlund, S., and Öhman, P. (1991) Iron in the water column of the Weddell Sea. *Mar Chem* **35**: 199–217.
- Whitney, L. P., Lins, J. J., Hughes, M. P., Wells, M. L., Chappell, P. D., and Jenkins, B. D. (2011) Characterization of putative iron responsive genes as species-specific indicators of iron stress in Thalassiosiroid diatoms. *Front Microbiol* **2**: 234.
- Wickham, H. (2009) Elegant graphics for data analysis. *Media* **35**: 211.
- Wolfe-Simon, F., Starovoytov, V., Reinfelder, J. R., Schofield, O., and Falkowski, P. G. (2006) Localization and role of manganese superoxide dismutase in a marine diatom. *Plant Physiol* **142**: 1701–1709.

## Supporting Information

Additional Supporting Information may be found in the online version of this article at the publisher's web-site:

Fig. S1. Exponential phase growth curves depicted as the natural log of raw fluorescence versus time, specific growth rates ( $\mu$ ) and maximum photochemical yield of PSII ( $F_v:F_m$ ) of iron-replete (black) and iron-limited (white) *P. granii* cultures grown for transcriptomic sequencing (A) or proteomic analyses (B). Circles represent biological replicate 1, triangles replicate 2, and squares replicate 3. Asterisk denotes a significant difference between treatments based on a paired *t*-test ( $p < 0.05$ ).

Fig. S2. Heatmap displaying  $\log_2$  normalized transcript counts obtained via edgeR (trimmed mean of *M* values [TMM]) for genes of interest across iron-replete (+Fe) and iron-limited (–Fe) *P. granii* samples. Row labels are distinguished by biological replicates in the iron-replete treatments ( $n = 3$ ) and iron-limited treatments ( $n = 2$ ). Genes of interest are grouped by iron (red), nitrogen (green) or vitamin-related (purple) metabolism. Genes encoding proteins involved in other metabolic processes are labelled in black. Dendrogram reflects similarity in transcript abundances across treatments and replicates based on Euclidean distance and hierarchical clustering, created with the R packages pheatmap and viridis. Grey boxes indicate transcripts were not detected. See Fig. S4 for full list of gene names.

Fig. S3. Normalized transcript counts (edgeR's trimmed mean of *M* values [TMM]) and normalized spectral abundance factor scores (NSAF) for proteins of interest detected in both the *P. granii* transcriptome (A) and proteome (B) shown on a  $\log_{10}$  scale separated by iron treatment. Asterisks denote a statistically significant difference between treatments (edgeR's likelihood ratio test, FDR < 0.05 or PLGEM, estimated FPR < 0.05). Error bars represent the standard deviation between biological triplicates (Fe-replete) and biological duplicates (Fe-limited) in the transcriptome and biological duplicates for both treatments in the proteome. See Fig. S4 for full list of gene names.

Fig. S4. Alignment of ferredoxin (*PETF*) amino acid sequence identified within the *P. granii* transcriptome compared to *PETF* from other examined diatom species within the MMETSP database (A). Alignment was created using MUSCLE within Geneious version 5.6.4 (Edgar 2004). Residues are colour-coded by blosum62 score matrix similarity (threshold = 5); residues with 100% similarity are represented in green, 80%–100% in gold, 60%–80% in yellow, and less than 60% in white. Iron-binding cysteine residues are outlined in black (Bertini et al. 2002; Groussman et al., 2015). *Pseudo-nitzschia granii* *PETF* secondary structure

and protein model was obtained using Phyre2 (B), with  $\alpha$  helices in pink,  $\beta$  sheets in yellow and turns in blue; 99.8% confidence and 47% identity to the 2Fe-2S ferredoxin-related family (Kelley *et al.*, 2015).

Fig. S5. MA (log ratio vs. mean average) plot displaying module expression (based on transcripts) between acclimated iron-replete (+Fe) and iron-limited (−Fe) treatments. Normalized counts corresponding to individual genes (KOs) were summed by module; each point on the plot corresponds to a unique KEGG module. Modules are coloured based on class 2 and class 3 KEGG groupings.  $\log_2$  sumA represents the average of normalized count abundance between treatments. Points increase in size with higher percentages of module completeness. Some modules corresponding to vitamin synthesis (purple) and metal transport (red) are labelled.

Fig. S6. Example MS1 chromatograms and MS2 fragment ions for putative biomarker peptides of the iron complex transport system substrate-binding protein ABC.FEV.S (A, contig ID: TR27940lc2\_g3\_i1), iron starvation induced protein 1 (B, contig ID: TR12112lc0\_g1\_i1), and flavodoxin (C, contig ID: TR29638lc1\_g1\_i1). MS1 filtering was performed in Skyline (Schilling *et al.*, 2012). The ABC.FEV.S and flavodoxin peptides shown here contained methionine and cysteine modifications respectively. Multiple variants of these peptides were detected with each containing different

modifications, though all increased in peak area abundance under low iron limitation.

Fig. S7. Representation of overall cellular metabolism visualized using iPath software (<http://pathways.embl.de>) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway information for *P. granii* transcripts (A) and proteins (B) with twofold increases in expression under iron-replete (red) or iron-limited (cyan) conditions, as described in Nunn *et al.* (2013).

Supplemental Table S1. Illumina Miseq transcriptomic sequencing statistics.

Supplemental Table S2. *Pseudo-nitzschia granii* differential gene expression results between the acclimated +Fe and −Fe treatments.

Supplemental Table S3. *Pseudo-nitzschia granii* normalized spectral abundance factor (NSAF) scores in the acclimated +Fe and −Fe treatments.

Supplemental Table S4. *Pseudo-nitzschia granii* differential gene expression and protein abundance results between the +Fe and −Fe treatments obtained using the edgeR log ratio test (LRT) (transcriptome) or power law global error model (proteome).

Supplemental Table S5. Putative peptide biomarkers of iron stress in *Pseudo-nitzschia granii*.