Different iron storage strategies among bloom-forming diatoms

Supplementary Information

Robert H. Lampe[a], Elizabeth L. Mann[b], Natalie R. Cohen[a,c], Claire P. Till[d], Kimberlee Thamatrakoln[e], Mark A. Brzezinski[f,g], Kenneth W. Bruland[h], Benjamin S. Twining[b], Adrian Marchetti[a,1]

[a]Department of Marine Sciences, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599
[b]Bigelow Laboratory for Ocean Sciences, East Boothbay, ME 04544
[c]Marine Chemistry and Geochemistry Department, Woods Hole Oceanographic Institution, Woods Hole, MA, USA 02543
[d]Department of Chemistry, Humboldt State University, Arcata, CA 95521
[e]Department of Marine and Coastal Sciences, Rutgers, the State University of New Jersey, New Brunswick, NJ 08901
[f]The Marine Science Institute, University of California, Santa Barbara, CA 93106
[g]Department of Ecology Evolution and Marine Biology, University of California, Santa Barbara, CA 93106
[h]Department of Ocean Sciences, University of California, Santa Cruz, CA 95064

[1]To whom correspondence should be addressed:
Dr. Adrian Marchetti
Department of Marine Sciences
University of North Carolina at Chapel Hill
3202 Venable Murray Hall CB 3300
Chapel Hill, NC 27599-3300, USA
Email: amarchetti@unc.edu

This PDF file includes:

Supplementary text
Figs. S1 to S16
Tables S1 to S3
References for SI citations

Supporting Information Text

Evaluation of iron status

Iron status of diatoms was assessed based on oceanographic context, differences in the chemical and biological properties of the seawater, and a combination of gene expression-based molecular indicators that evaluate iron stress or limitation for distinct diatom genera. For molecular analysis in *Pseudo-nitzschia*, the *Pseudo-nitzschia* Iron Limitation Index was used (Fig. S2)(1). In *Thalassiosira*, flavodoxin (*FLDA1*) and iron-starvation induced protein 3 (*ISIP3*) expression were quantified, as these genes have been shown to be indicators of iron limitation in *T. oceanica* (2). The concept of using *ISIP3* expression to assess iron limitation was extended to members of the genus *Chaetoceros* as they are also centric diatoms and possess the gene (Fig. S2).

Sites C-Low1, C-Low2, and P-Low displayed low iron concentrations and Fe:NO_3 ratios, which were anticipated from their locations (Table 1). Although C-Low1 was near a moderate shelf-width area (Fig. S1), relatively low dissolved iron and high micronutrients were still observed (Table 1, Fig. 2). This site is offshore the 200 m isobath, and therefore, it is beyond the observed natural boundary for iron-replete high productivity waters (3). C-Low2 is located over a narrow shelf region (Fig. S1) explaining the reduced iron. Both of these sites had lower initial dissolved Fe concentrations and Fe:NO_3 ratios well below the theoretical 8 nM:20 μM ratio for near-complete drawdown of nitrate (4).

At these low iron sites, significant differences among treatments in many of the bulk community measurements were observed, but these differences were not universal except for F_v:F_m (Fig. 2, Table S1). When comparing the iron and control incubations, significant increases in the iron treatment were observed in chlorophyll for all sites although this was only in the large size fraction at C-Low2 and the small size fraction at P-Low. C-Low1 had strongly significant differences in biogenic silica and nitrate drawdown not observed at the other sites. It is likely given the availability of nutrients and delayed increase in phytoplankton abundance at C-Low2 that more significant differences would have been observed had there been additional samples later in time. The result at P-Low can be attributed to the dominance of small phytoplankton initially present at this oceanic site. Biomass still remained quite low, but significant increases in biogenic silica were observed in the Fe incubations over time and relative to the DFB treatment suggesting enhanced diatom growth when Fe was added.

The *Pseudo-nitzschia* Iron Limitation Index transitioned to near positive values in the control incubations at C-Low1 and C-Low2 implying that these communities have the potential to naturally become iron limited (Fig. S2). It was strongly positive at P-Low indicating substantial iron stress. Indicators for the centric diatoms were also significantly different between the iron and control incubations (Fig. S2, Table S2). Overall, these pronounced differences in the chemical parameters and molecular indicators between the iron and control incubations substantiate that a change in iron status was induced by the addition of iron.

C-High already showed high *in situ* biomass and dissolved iron from being located in a wide-shelf area with strong upwelling (Figs. S1 and S16). As a result, no significant differences in nitrate, chlorophyll, or biogenic silica between the Fe addition and control incubations were observed supporting that iron availability was not limiting. The molecular indicators also display similar values between the iron and control treatments supporting the lack of an iron affect (Fig. S2, Table S2).

P-High showed significant differences between the Fe and control incubations for nitrate drawdown and F_v:F_m similar to that of the low iron sites. Additionally, the molecular indicators suggest that there may have been iron stress in the control incubations, and dissolved iron concentrations were initially less than those of the CUZ sites. These results are indicative of a low iron scenario; however, productivity at this station is known to be dependent on the delivery of macronutrients from seasonal upwelling instead of just iron (5). The initial conditions are consistent with this dependence as macronutrients were fairly depleted and iron still remained (Fig. 2, Table 1). After the addition of nitrate to these incubations (Materials and Methods), significant increases in chlorophyll ($p < 0.0001$) were observed indicating that nitrate availability was primarily limiting. Although significant differences in nitrate drawdown between the iron and the control incubations were observed by the second time point, chlorophyll concentrations were not different, pointing to the potentially depressed iron effect at this site. It is likely that other macronutrients (phosphate and silicate) were limiting preventing greater contrast (Fig. 2). Altogether, these results suggest iron had an effect at this site, although macronutrient availability was most likely driving productivity leading to classification as a high iron site, especially relative to its NE Pacific counterpart, P-Low.

Materials and Methods

Experimental Design

Incubation experiments were conducted at five sites within the California Upwelling Zone (CUZ, July 2014) onboard the *R/V Melville* or along the Line P transect in the subarctic NE Pacific (June 2015) onboard the *CCGS John P. Tully* (Fig. 1, Table 1). Experimental procedures are described in Cohen*, et al.* (6) for all sites except C-Low2, which is described by Lampe*, et al.* (7). In summary, near-surface water (Table S1) was collected using trace-metal clean techniques and distributed into acid-cleaned 10 L Cubitainers® (Hedwin Corporation, Newark, DE, USA) except C-Low2 for which water was collected from 96 m. Cubitainers were incubated on-deck at near-ambient surface water temperature and screened at approximately one-third surface irradiance. Treatments included an unamended control (C), addition of iron with 5 nmol L^{-1} $FeCl_3$ (Fe), removal of iron with 200 nmol L^{-1} of the fungal siderophore desferroxamine B (DFB), and iron addition at t = 0 followed by iron removal at the first time point (FeDFB) to mimic a short-lived iron pulse. At site P-Low, DFB was added to a control and Fe treatment at 48 hours for the DFB and FeDFB treatment sampled at 96 hours. At site P-High, 10 µmol L^{-1} of NO_3 was added to all incubations to support growth since initial NO_3 concentrations were low (1.49 µmol L^{-1}). Timing of incubation sampling for each experiment is described in Table S1. At dawn of each time point, triplicate cubitainers of each treatment were harvested and subsamples were collected for the following measurements: dissolved inorganic nutrients, chlorophyll *a,* biogenic silica, F_v:F_m, RNA, and SXRF. Methods for analyzing chlorophyll *a,* biogenic silica, and F_v:F_m are described in Cohen*, et al.* (6).

RNA-Seq Assembly, Annotation, and Analysis

The initial RNA extraction, sequencing, assemblies, and annotations are described in Cohen*, et al.* (6) and Lampe*, et al.* (7). Briefly, seawater was filtered onto 0.8 µm filters then immediately flash frozen. RNA was extracted using the ToTALLY RNA Total RNA Isolation Kit and treated with DNase 1 (Life Technologies, Grand Island, NY, USA). Library preparation was conducted with the Illumina TruSeq Stranded mRNA Library Preparation Kit and HiSeq v4 reagents. Sequencing of barcoded samples was performed on an Illumina HiSeq 2000 (125bp, paired-end). All samples were sequenced in triplicate except samples for treatments with low yields (all t = 0 and Line-P samples) where triplicate extractions were pooled into one sample. Reads were trimmed for quality

and removal of adapters using Trimmomatic v0.32 (8). Trimmed paired reads that overlap were merged into single reads with BBMerge v8.0. Merged pairs and non-overlapping paired-end reads were used to assemble contigs using ABySS v1.5.2 with varied k-mer sizes (9) then the assemblies were merged using Trans-ABySS v1.5.3 to remove redundant contigs (10). Contigs shorter than 120 bp were discarded.

Annotation was assigned by best homology (lowest E-value) to protein databases using BLASTX v2.2.31 (E-value $\leq 10^{-3}$). For taxonomic identification, MarineRefII, a custom reference database (Moran Lab, University of Georgia) was used. MarineRefII contains predicted protein sequences of marine microbial eukaryotes and prokaroytes including all sequenced transcriptomes from the Marine Microbial Eukaryote Transcriptome Sequencing Project (11). To assign gene function to contigs, the same methodology with the Kyoto Encyclopedia of Genes and Genomes (KEGG; Release 75) was used (12). The best hit with a KEGG Ortholog (KO) number from the top 10 hits was chosen.

For this analysis, assemblies from all sites were then merged again with Trans-ABySS and duplicate contig removal verified with GenomeTools v1.5.1 (13). Read counts were estimated from this combined assembly using the quasi-mapping method implemented in Salmon v0.73 (14). The *Pseudo-nitzschia* Iron Limitation Index (ILI) was calculated using our annotations and un-normalized counts for ferritin and iron starved induced protein 2A (*ISIP2A*) according to the formula presented in Marchetti*, et al.* (1). In all other instances, normalization, gene expression, and differential expression were assessed within each taxonomic group using DESeq2 v1.12.4 (15). Significance was determined by genes with Benjamini & Hochberg adjusted P-values ≤ 0.05 (16). As diatom iron starvation induced protein (*ISIP*) genes are not included in the KEGG Orthology database, they were manually annotated based on top BLAST hits to genes described in Morrissey*, et al.* (17) with verification from the KEGG SSDB database (12).

Synchrotron X-ray Fluorescence (SXRF) Preparation and Analysis

Samples were collected and analyzed following Twining, *et al*. (18, 19). Briefly, cells from 500 mL of sample were gently preconcentrated approximately ten-fold over 2µm pore-size polycarbonate membranes via gravity filtration. The remaining 40 mL was transferred to a centrifuge tube, preserved with 0.25% trace-metal clean electron-microscopy grade buffered glutaraldehyde, and centrifuged onto C/formvar-coated Au TEM grids. After centrifugation, grids

were briefly rinsed with Milli-Q water and dried in a darkened Class-100 hood. Grids were then mounted onto custom-machined Al holders for storage and analysis. Light and chlorophyll fluorescence images were collected for target cells using a shipboard microscope. SXRF analyses were performed at the 2-ID-E microprobe beamline at the Advanced Photon Source (Argonne National Laboratory). The incident beam energy was tuned to 10 keV to allow for the stimulation of K-line emissions for all elements ranging in atomic number from Si through Zn. Each target cell was scanned in a 2-D raster fashion by the focused X-ray beam and the entire X-ray fluorescence spectrum recorded at each pixel. The spectra from the pixels covering the cell were summed to generate a single spectrum corrected with a background region. Element concentrations were calculated by comparison to certified reference standards (19, 20), and cellular C was calculated from cell volume (21).

Macronutrient Analyses

For the California Upwelling Zone (CUZ) sites, dissolved nitrate + nitrite (NO_3^- + NO_2^-), phosphate (PO_4^{3-}), and silicic acid (H_4SiO_4) concentrations were measured using a Lachat Quick Chem 8000 Flow Injection Analysis system (22) whereas for Line P sites, an Astoria Analyzer was used (23). Particles were removed by filtering the sample through a GF/F filter using a 60 mL syringe. CUZ samples were analyzed onboard the ship whereas Line P samples were stored in acid-rinsed polypropylene tubes and frozen at -20°C prior to analysis onshore. Reference materials for nutrients in seawater were run alongside samples for quality control. For examining historical data (Fig. 1), interpolated statistical means of nitrate data on a 1° grid for all decades from World Ocean Atlas 2013 (24) was plotted using matplotlib (25) for Python v2.7.

Dissolved Iron

For the CUZ incubations, seawater was acidified at sea with 4 mL 6 N quartz-distilled HCl per L of seawater (pH ~1.7) and stored in acid-cleaned LDPE bottles for at least two months prior to analysis. Iron was preconcentrated from buffered (pH 6.0) seawater on Nobias-chelate PA1 resin and eluted with 1 N quartz-distilled HNO_3 following Biller and Bruland (26) and Parker, Brown and Bruland (27). The eluent was analyzed with a Thermo-Element XR™ ICP-MS in counting mode. Line-P dissolved Fe samples were stored in acid-cleaned LDPE bottles, acidified post-cruise with 1 mL 12 N Optima-grade HCl per L of seawater, and stored for at least three months prior to

analysis. These samples were also preconcentrated on resin and measured with a ICP-MS following Milne, Landing, Bizimis and Morton (28).

Statistical Analyses

One- and two-way ANOVAs followed by Tukey's multiple comparison test were performed on the biological and chemical properties of the seawater in Graphpad PRISM v7.03. Statistical methods for gene expression data are described under *RNA-Seq Assembly, Annotation, and Analysis*.

Phylogenetic Analysis

Internal Transcribed Spacer 1 (ITS1) sequences for known *Pseudo-nitzschia* species (29) were compiled from Hubbard, *et al*. (30, 31), and Marchetti*, et al.* (32). Reference ITS1 sequences were checked for duplicates with GenomeTools v.1.5.1 (13) then compared with contigs using BLAST v2.5.0 with a 98% similarity cutoff. Contigs were then merged based on 98% similarity favoring longer contigs. Reference ferritin (FTN) sequences were derived from Lampe*, et al.* (7), Cohen*, et al.* (33), and Moreno, *et al.* (34), and contigs were selected by their annotation described under *RNA-Seq Assembly, Annotation, and Analysis*. *NRAMP* and *ZIP1* sequences were retrieved from the JGI Genome Portal for diatoms and UniProt for all other organisms. For all trees, sequences were aligned with MUSCLE (35) in Geneious Pro v10.2.2. Maximum likelihood phylogenetic trees of the reference sequences were created with RAxML v8.2.9 (*ITS1*, GTRGAMMA model; all others, PROTGAMMALG model) with 100 bootstraps (36). Contigs were placed on the tree with pplacer v1.1.alpha18-2-gcb55169 (37), and the final trees were visualized with Archaeopteryx v0.9916 (38).

Figure S1. Bathymetry of the California Upwelling Zone. California-based incubation sites are labeled with the 200 m isobath denoted in red. Data are derived from the General Bathymetric Chart of the Oceans (GEBCO) 2014 grid model (www.gebco.net) and were viewed in Ocean Data View 4.

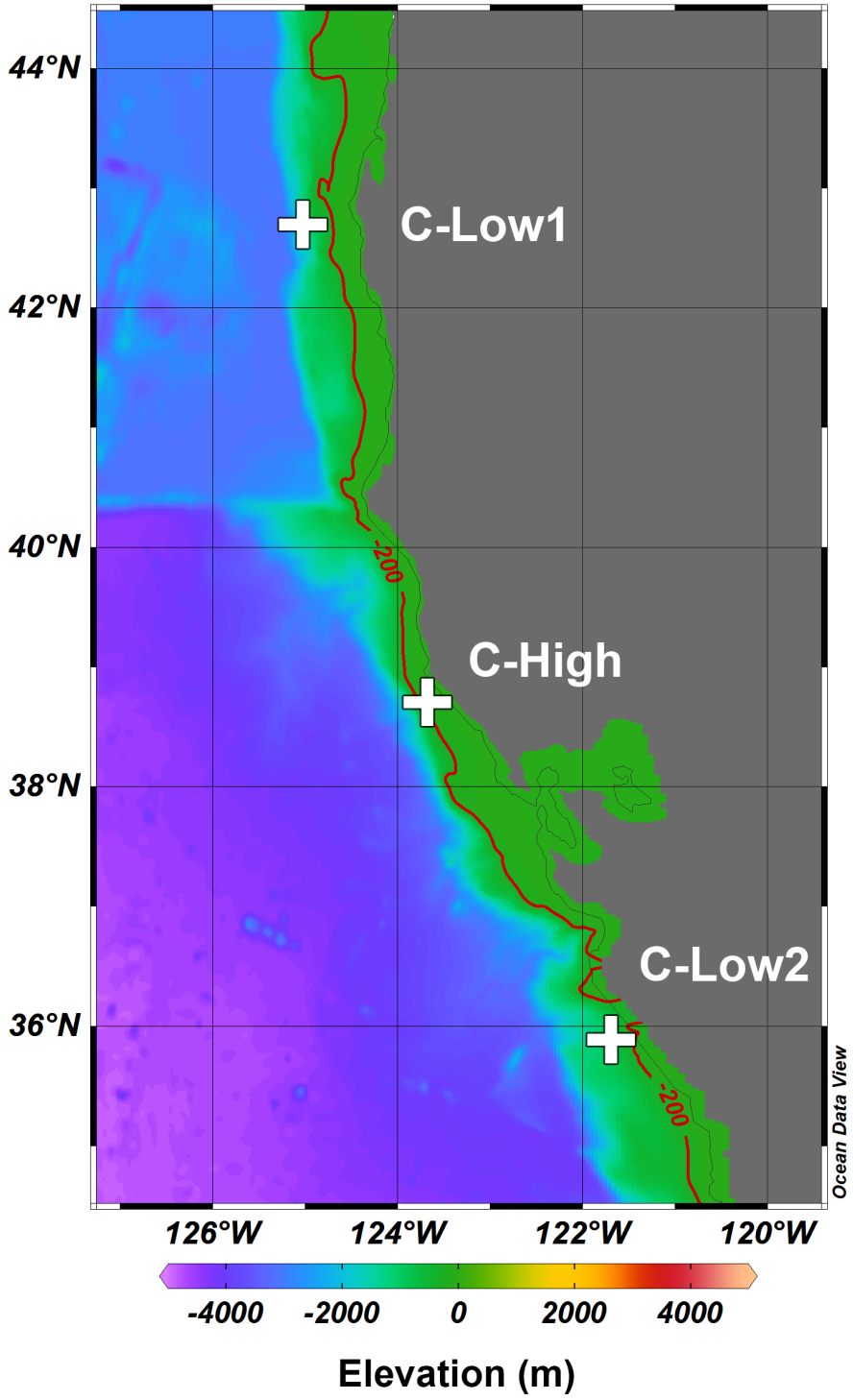

Figure S2. Molecular indicators of iron stress or limitation in diatoms as described by Marchetti*, et al.* (1) and Chappell*, et al.* (2). From left to right, the *Pseudo-nitzschia* Iron Limitation Index (*Ps-n* ILI) and flavadoxin (*FLDA1*) and iron starvation induced protein 3 (ISIP3) normalized counts for *Thalassiosira* (*Thal.)* and *Chaetoceros* (*Chae.*). A positive *Ps-n* ILI value indicates iron stress or limitation. Incubations are labelled as follows: control (C), iron addition (Fe), iron removal (DFB), iron addition then removal (FeDFB) and denoted as the first or second time point (T_1 or T_2) where applicable. An 'X' indicates that the value was unable to be calculated. White asterisks (*) denote that the value extends beyond the axis limits.

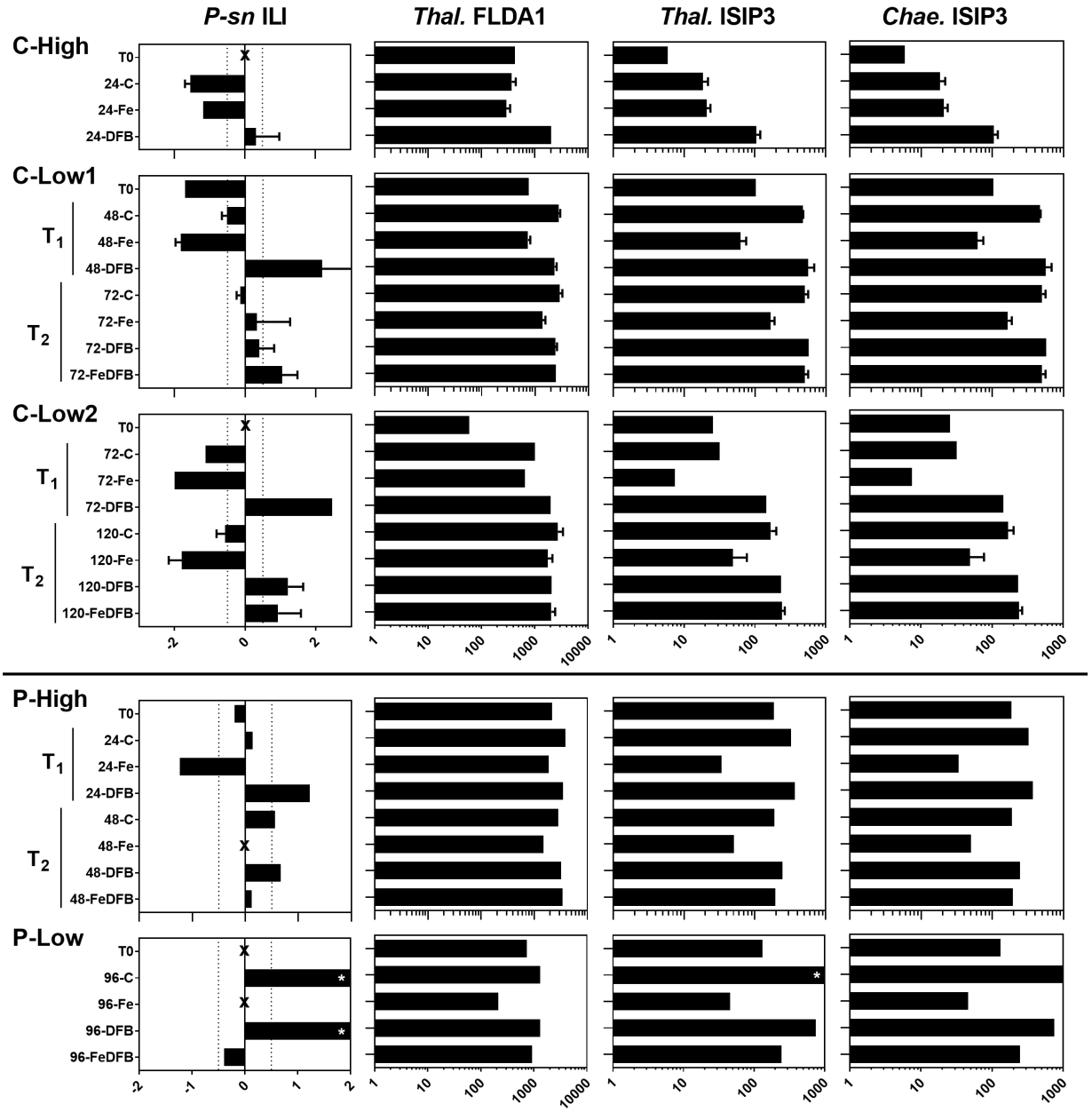

Figure S3. Average taxonomic distribution by mapped reads for the whole community (phylum-based) and diatom genera within all diatom assigned reads. Incubations are labelled as follows: control (C), iron addition (Fe), iron removal (DFB), iron addition then removal (FeDFB) and denoted as the first or second time point (T_1 or T_2) where applicable.

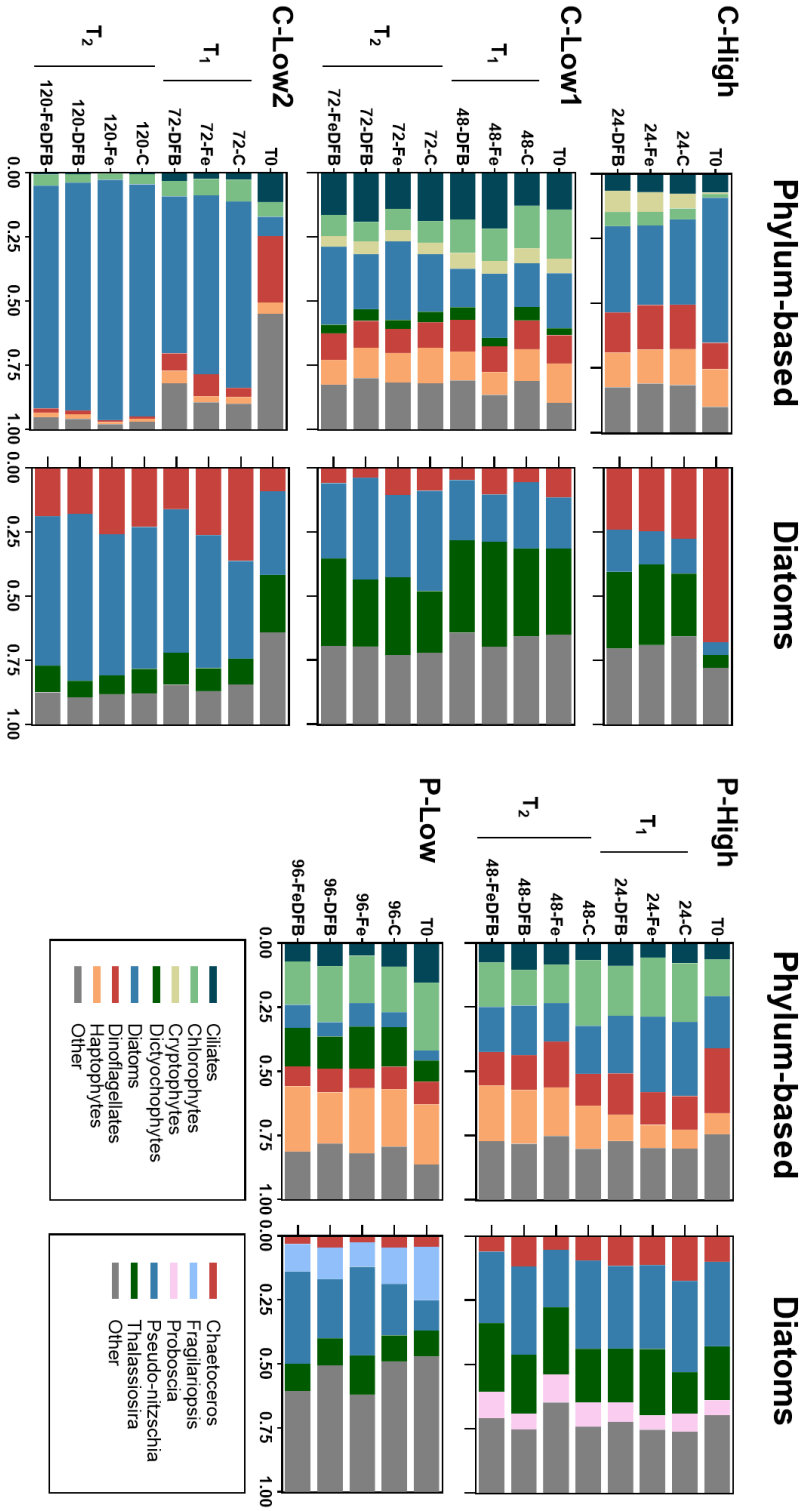

Figure S4. Light and synchrotron X-ray fluorescence micrographs of *Chaetoceros* and *Pseudo-nitzschia* cells from the control and iron (+Fe) incubations at 48 hours at C-Low1. The red box outlines the region of the SXRF scan for these particular cells.

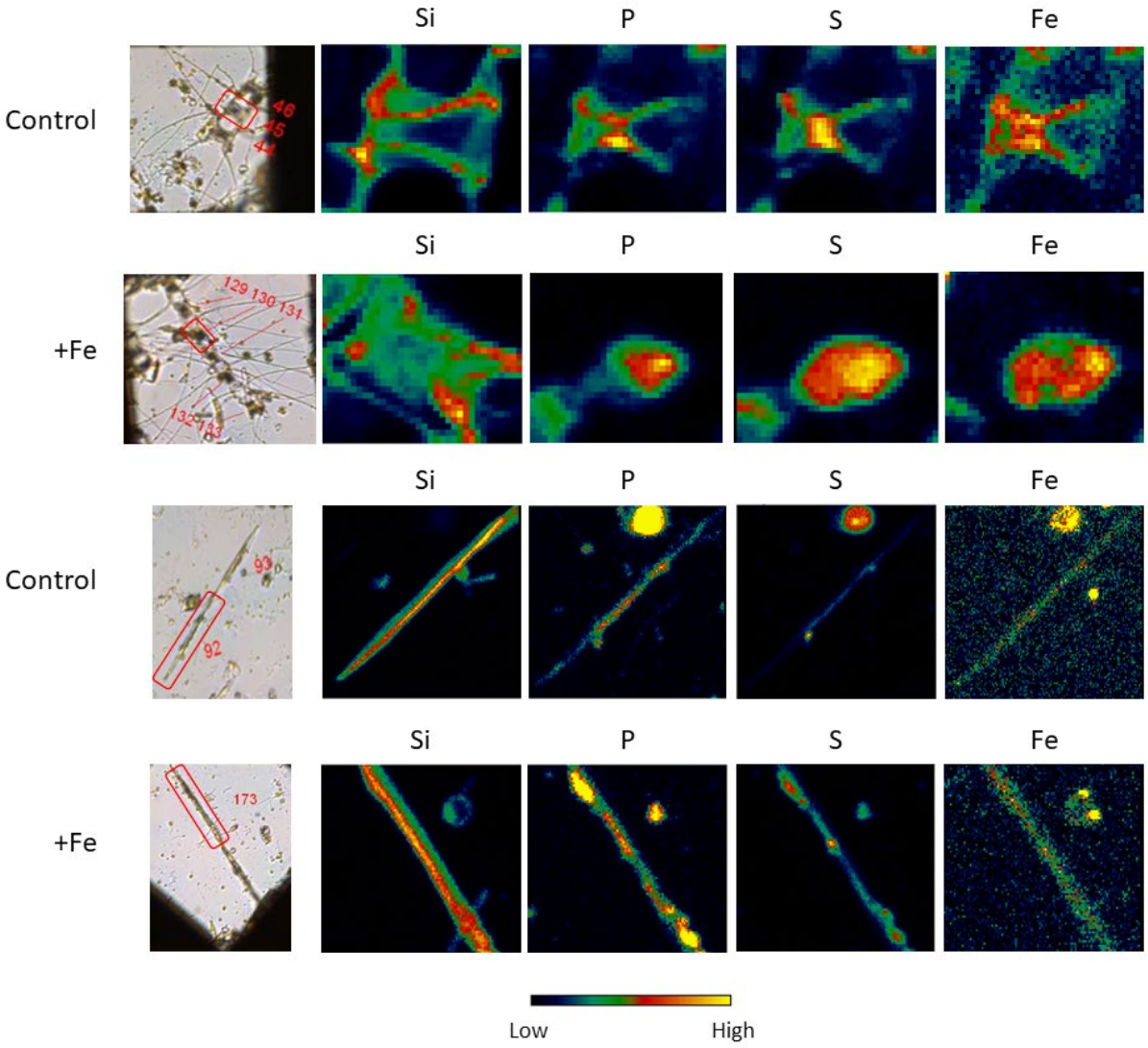

Figure S5. Normalized transcript abundances for ferritin (*FTN),* natural resistance-associated macrophage protein (*NRAMP*), and *ZIP1* in *Pseudo-nitzschia* (P), *Chaetoceros* (C), and *Thalassiosira* (T). Dark gray indicates that the gene was not detected.

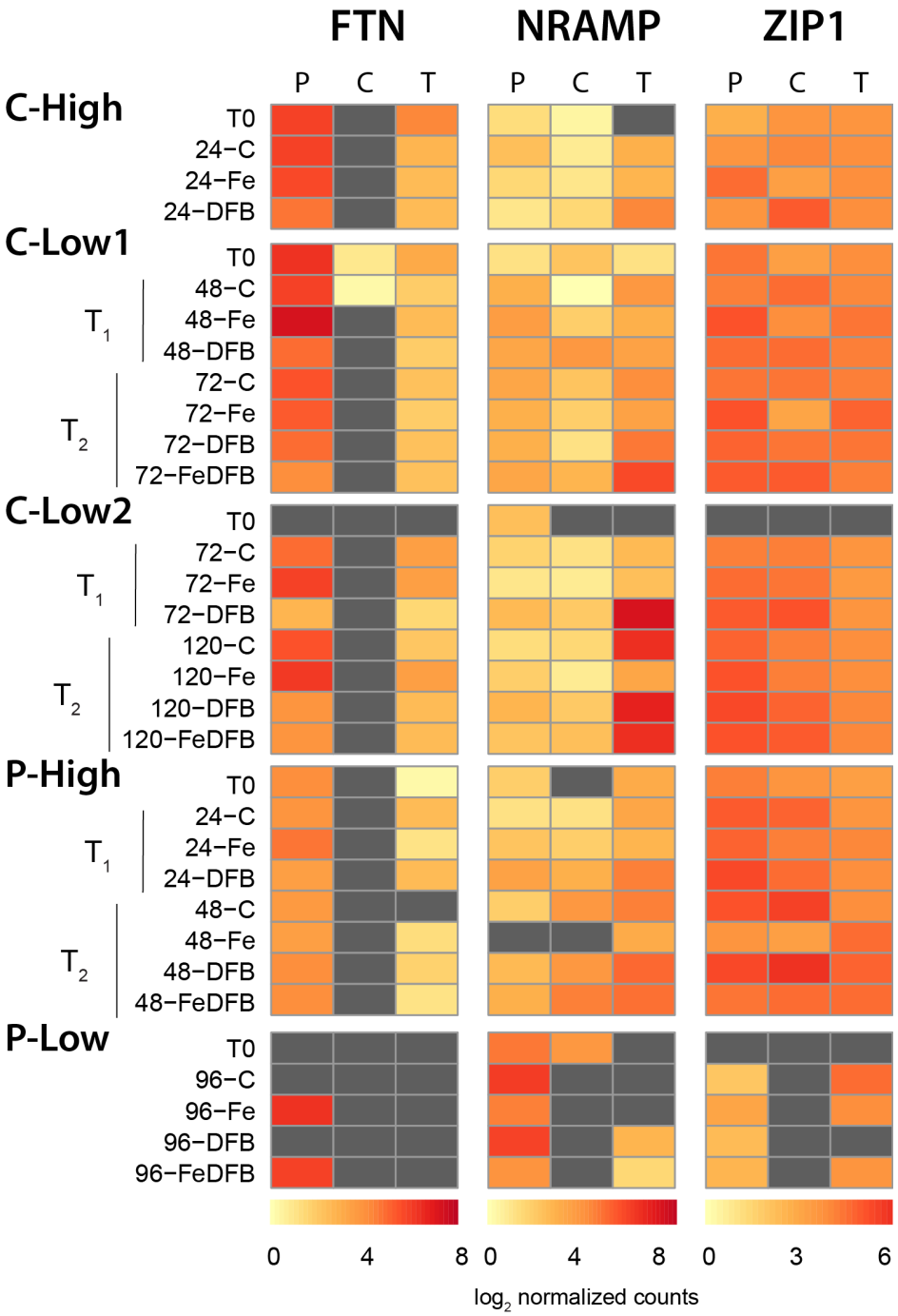

Figure S6. Phylogenetic tree of *Pseudo-nitzschia* internal transcribed spacer 1 (*ITS1*) sequences with environmental sequences from the metatranscriptome in blue. Bootstrap values ≥ 50 from the reference sequences are shown. Detection of reads mapping for each site is indicated in blue in the table.

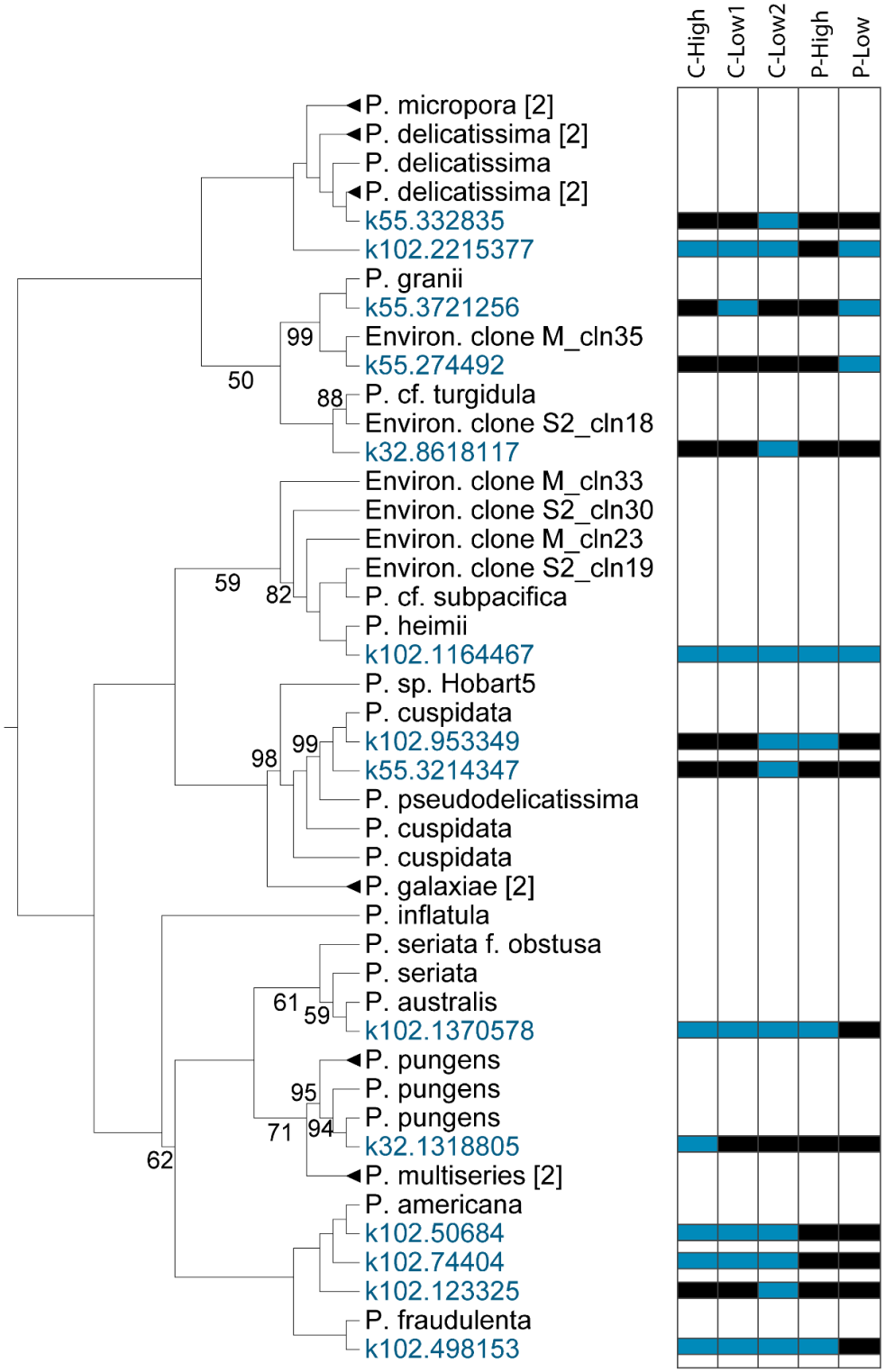

Figure S7. Phylogenetic tree of *Pseudo-nitzschia* ferritin sequences with environmental sequences from the metatranscriptome in blue. Bootstrap values ≥ 50 from the reference tree are shown. Detection of reads mapping for each site is indicated in blue in the table.

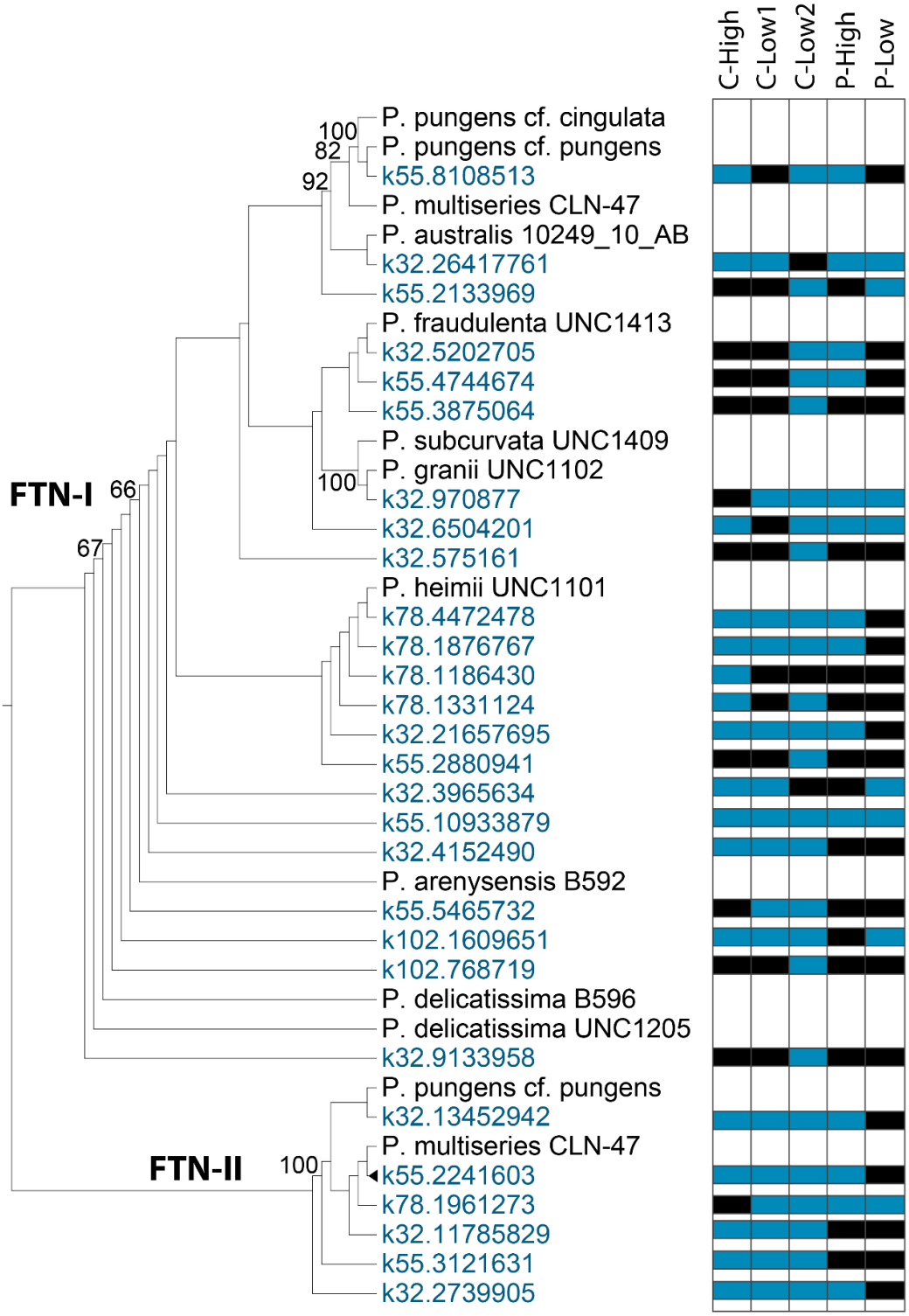

Figure S8. Fold changes in *Pseudo-nitzschia FTN* expression (x-axis) versus fold changes in iron quotas (µmol Fe:mol C; y-axis) for the Fe or DFB treatments compared to the control treatment. The P-High DFB vs C comparison (red point) was excluded as an outlier as silicate depletion likely uncoupled the relationship between *FTN* and iron quotas. A Pearson correlation indicates a significantly positive relationship ($\rho = 0.7551$, $P = 0.0186$).

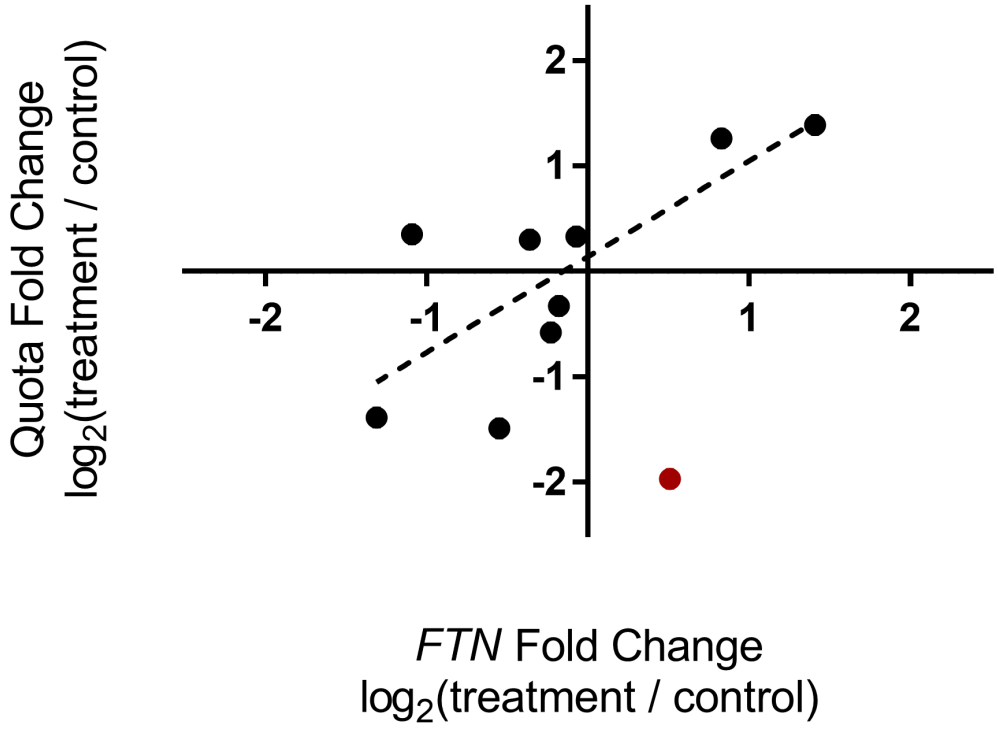

Figure S9. Midpoint-rooted phylogenetic tree of diatom ferritins. Branches that do not contain *Chaetoceros* sequences were collapsed with the number of sequences shown. *Chaetoceros*-assigned contigs are highlighted in red. Bootstrap values ≥ 50 are shown.

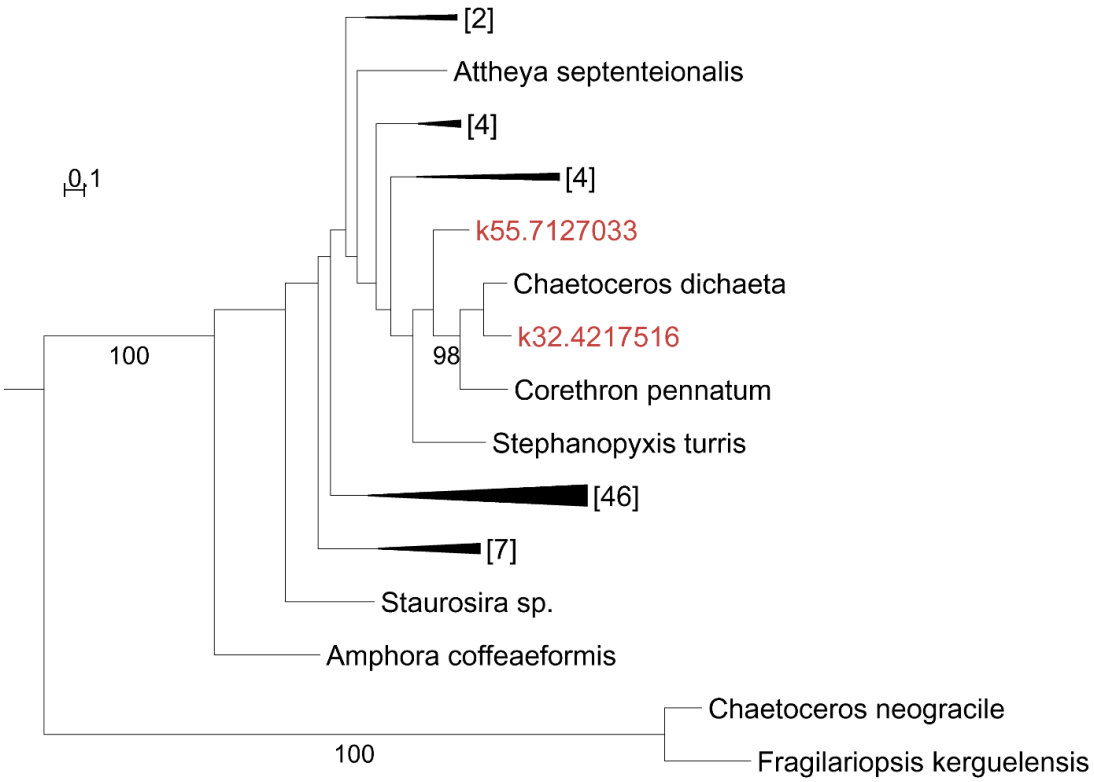

Figure S10. Midpoint-rooted phylogenetic tree of diatom ferritins. Branches that do not contain *Thalassiosira* sequences were collapsed with the number of sequence shown. *Thalassiosira*-assigned contigs are highlighted in red. Bootstrap values ≥ 50 are shown.

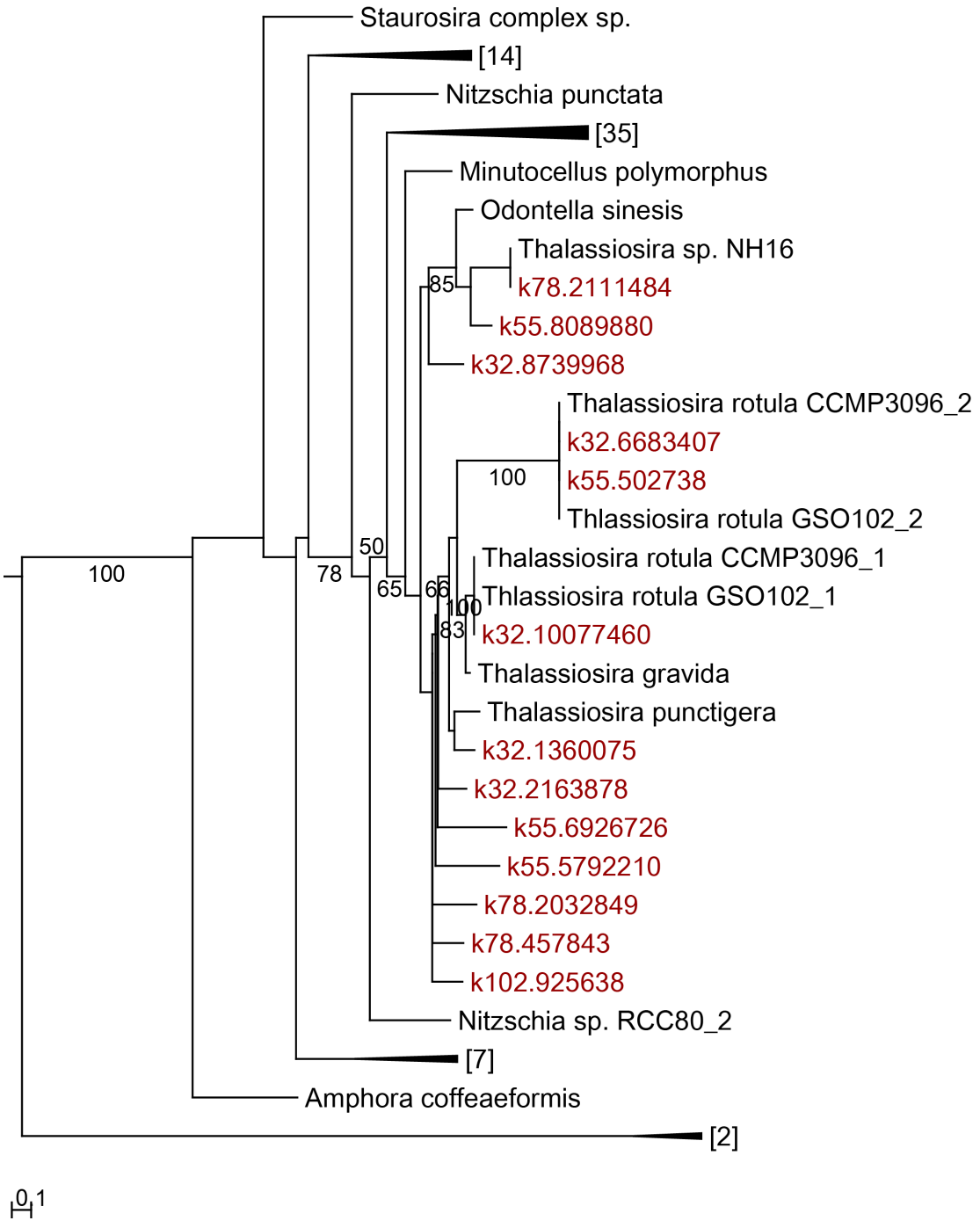

Figure S11. Normalized transcript abundances for ferritin (*FTN)* and natural resistance-associated macrophage protein (*NRAMP*) in *Fragilariopsis*. Dark gray indicates that the gene was not detected.

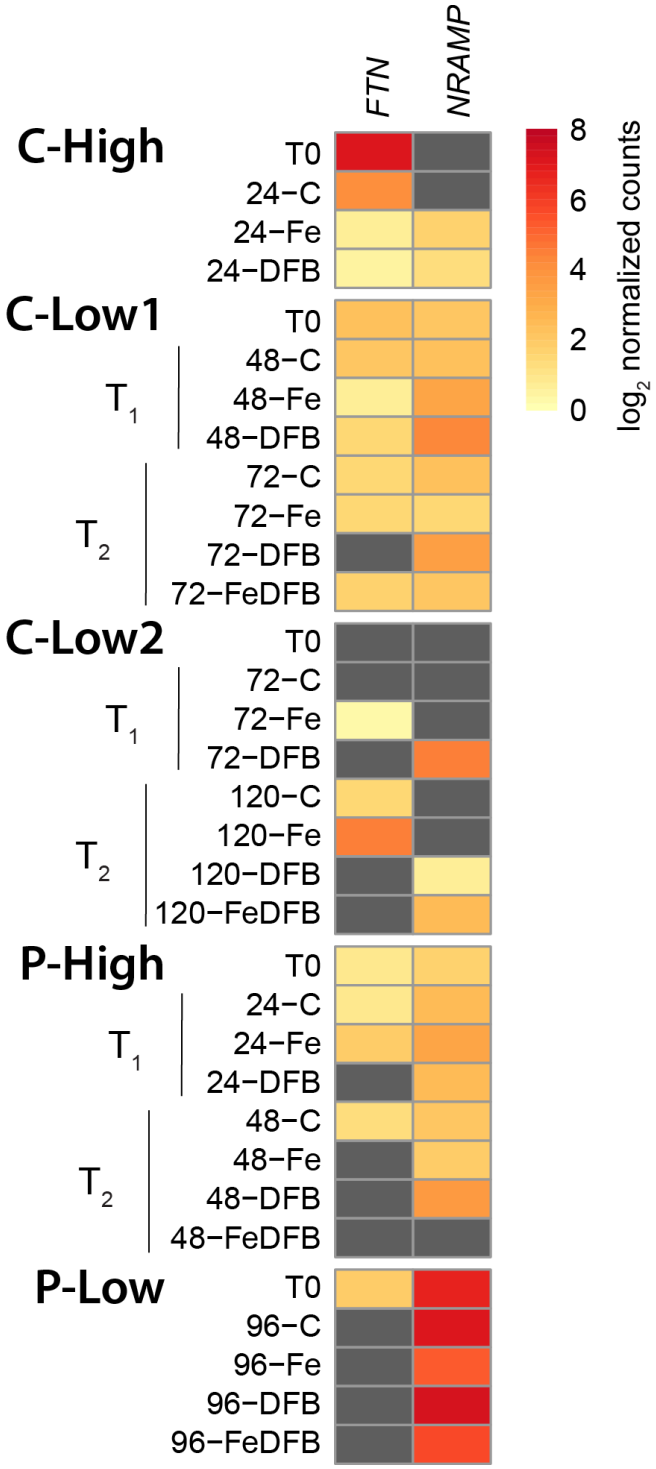

Figure S12. Phylogenetic tree of natural resistance-associated macrophage protein (*NRAMP*) homologous genes. Diatom genes (blue) are labeled as Fracy 172829, Psemu 325037, and Thaps 9840 from *Fragilariopsis cylindrus*, *Pseudo-nitzschia multiseries*, and *Thalassiosira pseudonana* with corresponding JGI gene IDs. The red algae, *Cyanidioschyzon merole* and *Galdieria sulphuraria*, are presented with gene IDs corresponding to entries in Uniprot. Remaining sequences are those presented by Thomine, Wang, Ward, Crawford and Schroeder (39) from *Arabidopsis thaliana* (At), *Oryza sativa* (Os), *Saccharomyces cerevisiae* (SMF), *Mus musculus* (Mm), *Deinococcus radidurans* (Dr), *Escherirchia coli* (Ec), *Pseudomonas aeruginosa* (Pa), and *Salmonella typhimurium* (St).

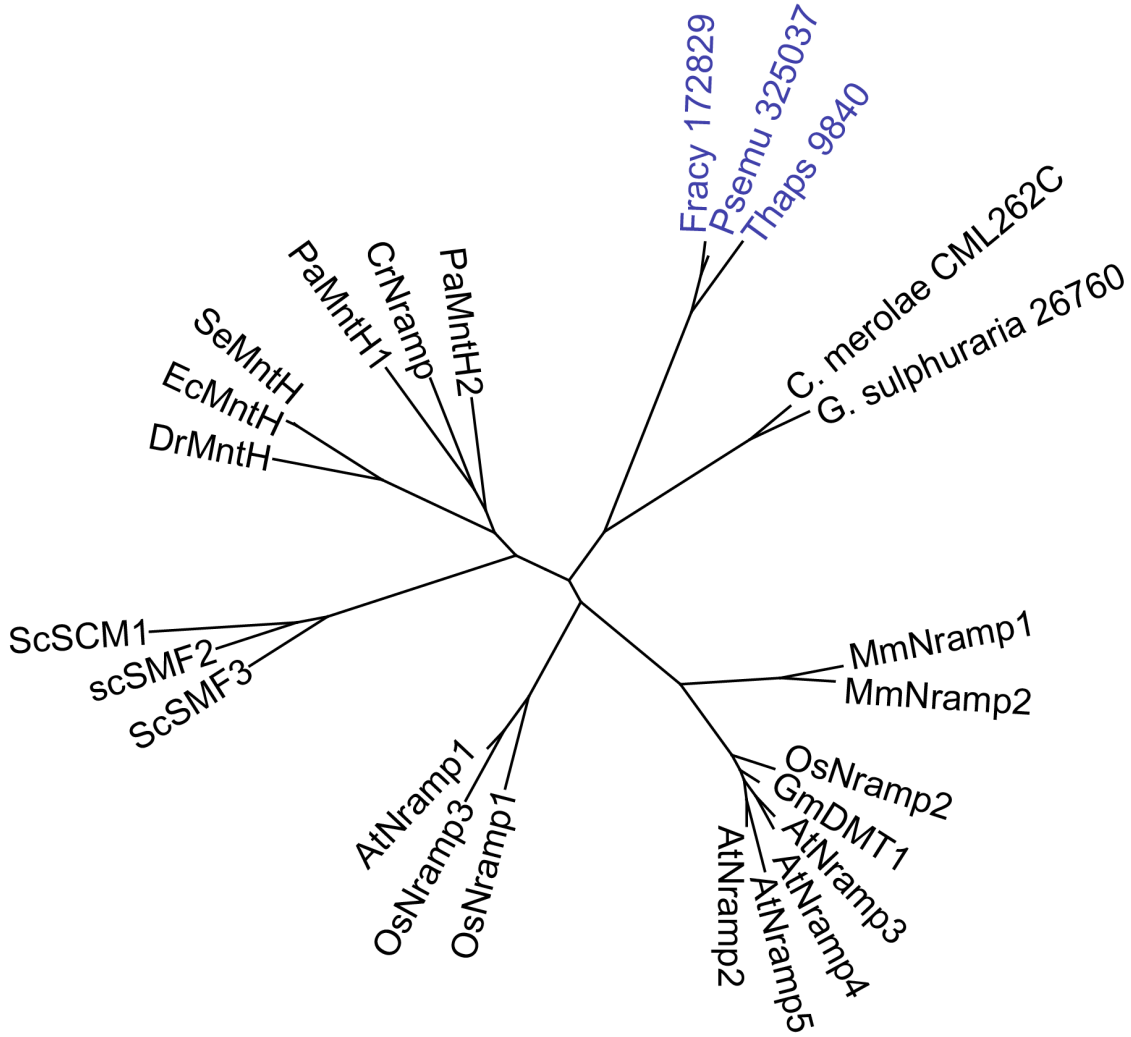

Figure S13. Normalized transcript abundances for *ZIP7* in *Pseudo-nitzschia* (P), *Chaetoceros* (C), and *Thalassiosira* (T). Dark gray indicates that the gene was not detected.

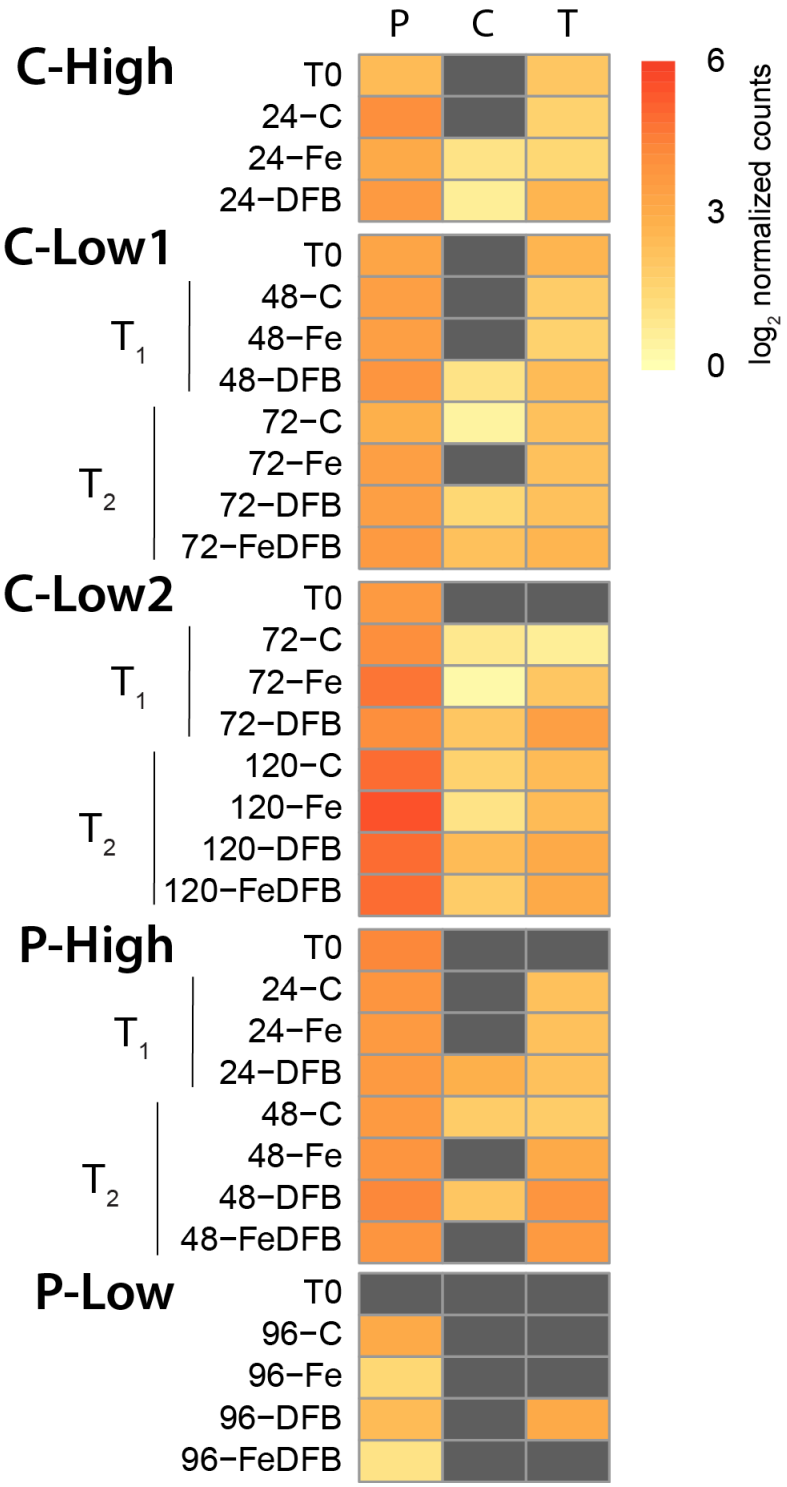

Figure S14. Midpoint-rooted phylogenetic tree of *ZIP1* homologous genes. Bootstrap values ≥ 50 are shown.

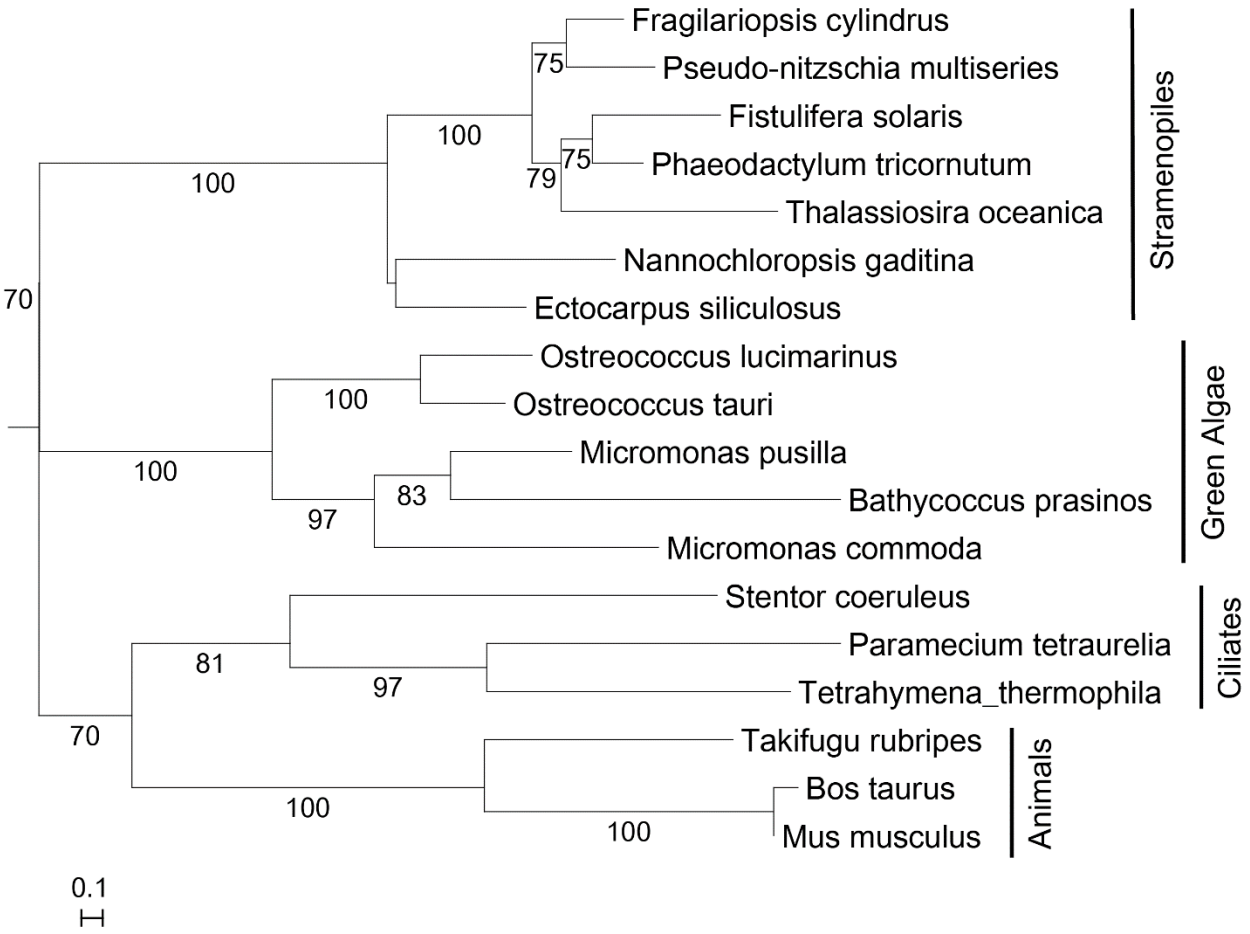

Figure S15. Domoic acid (DA) concentrations from C-Low2. Domoic acid was measured for all California Upwelling Zone (CUZ) incubations with the remaining data reported in Cohen*, et al.* (6).

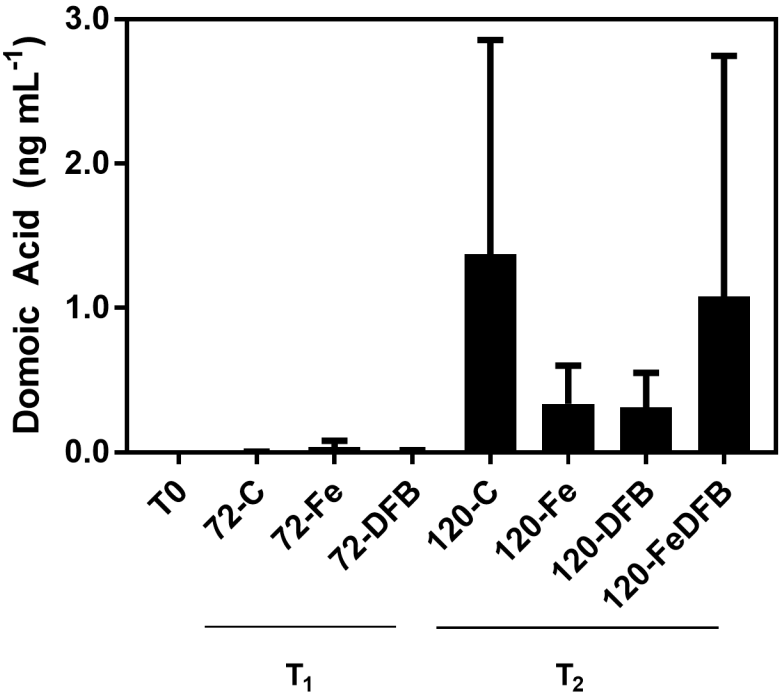

Figure S16. Three day average of satellite-derived sea surface temperature (SST, 0.0125°) from 03 June 2014 to 05 June 2014, the day of sampling for C-High (marked with +). Data is from the NOAA POES AVHRR satellite courtesy of the NOAA / NESDIS Center for Satellite Applications Research and downloaded from the NOAA CoastWatch Browser.

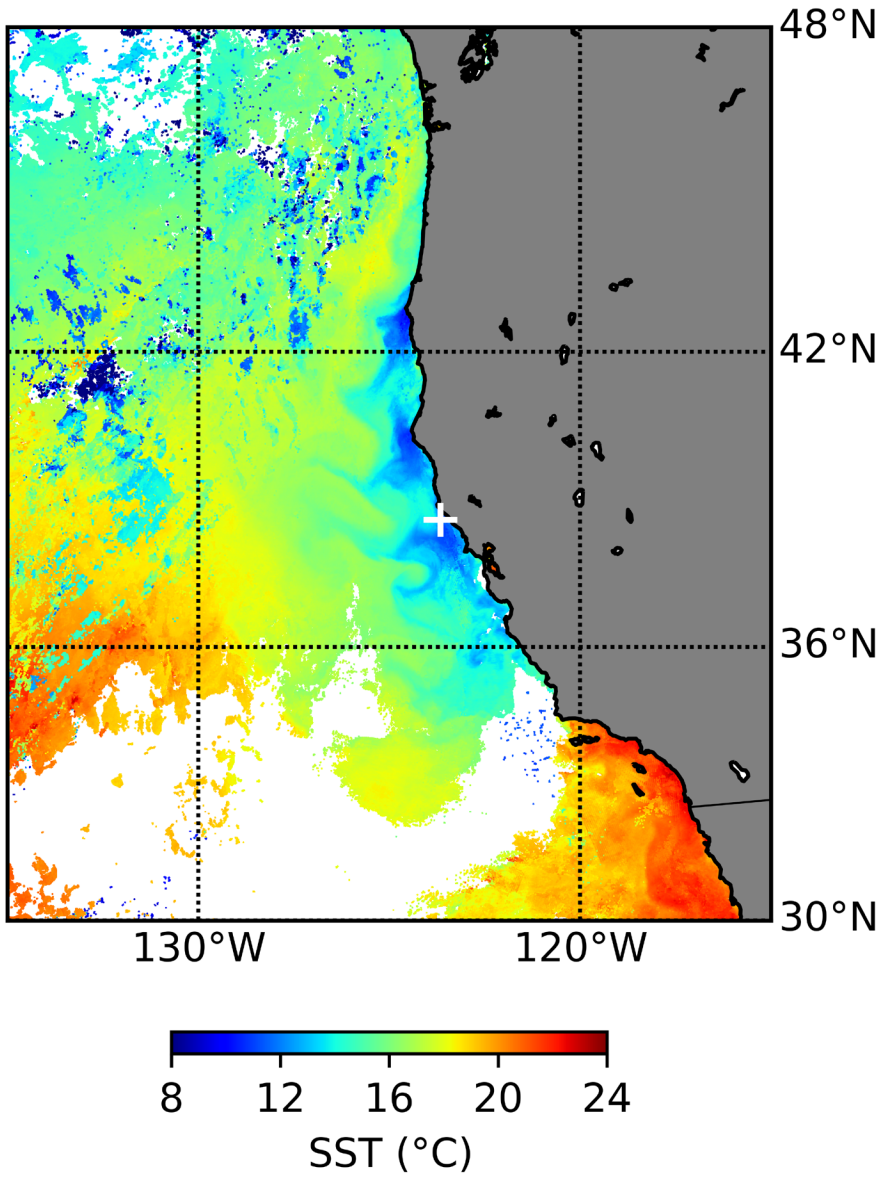

Table S1. Time points and initial seawater collection depths for each incubation site.

Name	Time points (hours)	Collection Depth (m)
C-High	$T_1 = 24$	2
C-Low1	$T_1 = 48$, $T_2 = 72$	15
C-Low2	$T_1 = 72$, $T_2 = 120$	96
P-High	$T_1 = 24$, $T_2 = 48$	3.5
P-Low	$T_1 = 96$	11

Table S2. Significance (Tukey) test results for differences in site parameters. Significance is shown as follows: ns, not significant; displayed numerically; P ≤ 0.1; *, P ≤ 0.05; **, P ≤ 0.01; ***, P ≤ 0.001; ****, P ≤ 0.0001; or (-), not applicable.

		C-High	P-High	C-Low1	C-Low2	P-Low
NO_3						
T_1-Fe	T_1-C	ns	ns	ns	ns	ns
T_1-Fe	T_1-DFB	**	ns	ns	ns	ns
T_2-Fe	T_2-C	-	****	****	ns	-
T_2-Fe	T_2-DFB	-	****	****	ns	-
Chl *a* > 5 µm						
T_1-Fe	T_1-C	ns	ns	ns	ns	0.05
T_1-Fe	T_1-DFB	ns	ns	*	ns	*
T_2-Fe	T_2-C	-	ns	****	**	-
T_2-Fe	T_2-DFB	-	****	****	****	-
T_1-Fe	t = 0	ns	ns	***	ns	0.06
T_2-Fe	t = 0	-	****	****	****	-
Chl *a* < 5 µm						
T_1-Fe	T_1-C	ns	ns	ns	ns	**
T_1-Fe	T_1-DFB	ns	ns	**	ns	****
T_2-Fe	T_2-C	-	ns	****	ns	-
T_2-Fe	T_2-DFB	-	ns	****	ns	-
T_1-Fe	t = 0	ns	ns	****	ns	***
T_2-Fe	t = 0	-	ns	****	ns	-
bSi						
T_1-Fe	T_1-C	ns	ns	ns	ns	ns
T_1-Fe	T_1-DFB	ns	ns	ns	ns	**
T_2-Fe	T_2-C	-	ns	****	ns	-
T_2-Fe	T_2-DFB	-	ns	****	ns	-
T_1-Fe	t = 0	ns	ns	-	-	***
T_2-Fe	t = 0	-	ns	-	-	-
F_v:F_m						
T_1-Fe	T_1-C	ns	***	**	**	****
T_1-Fe	T_1-DFB	ns	****	****	****	****
T_2-Fe	T_2-C	-	****	**	**	-
T_2-Fe	T_2-DFB	-	****	****	****	-

Table S3. Significance in differential expression FLDA1 and ISIP3 expression for *Thalassiosira* and *Chaetoceros* determined by DESeq2 and shown in Figure S3. Significance is shown as follows: ns, not significant; *, $P \leq 0.05$; **, $P \leq 0.01$; ***, $P \leq 0.001$; ****, $P \leq 0.0001$; or (-), not applicable.

		C-High	C-Low1	C-Low2
***Thalassiosira* FLDA1**				
C vs Fe	T_1	ns	****	-
	T_2	-	****	ns
Fe vs DFB	T_1	****	****	-
	T_2	-	**	ns
***Thalassiosira* ISIP3**				
C vs Fe	T_1	ns	****	-
	T_2	-	****	***
Fe vs DFB	T_1	****	****	-
	T_2	-	****	****
***Chaetoceros* ISIP3**				
C vs Fe	T_1	ns	****	-
	T_2	-	****	***
Fe vs DFB	T_1	****	****	-
	T_2	-	****	****

References

1. Marchetti A, et al. (2017) Development of a molecular-based index for assessing iron status in bloom-forming pennate diatoms. *J Phycol*:820-832.
2. Chappell PD, et al. (2015) Genetic indicators of iron limitation in wild populations of *Thalassiosira oceanica* from the northeast Pacific Ocean. *ISME J* 9(3):592-602.
3. Kudela RM, Garfield N, Bruland KW (2006) Bio-optical signatures and biogeochemistry from intense upwelling and relaxation in coastal California. *Deep-Sea Res Pt II* 53(25–26):2999-3022.
4. Bruland KW, Rue EL, Smith GJ (2001) Iron and macronutrients in California coastal upwelling regimes: Implications for diatom blooms. *Limnol Oceanogr* 46(7):1661-1674.
5. Whitney FA, Crawford DW, Yoshimura T (2005) The uptake and export of silicon and nitrogen in HNLC waters of the NE Pacific Ocean. *Deep-Sea Res Pt Ii* 52(7):1055-1067.
6. Cohen NR, et al. (2017) Variations in diatom transcriptional responses to changes in iron availability across ocean provinces. *Front Mar Sci* 4:360.
7. Lampe RH, et al. (2018) Divergent gene expression among phytoplankton taxa in response to upwelling. *Environ Microbiol* 20:3069-3082.
8. Bolger AM, Lohse M, Usadel B (2014) Trimmomatic: A flexible trimmer for Illumina Sequence Data. *Bioinformatics* 30(15):2114-2120.
9. Birol I, et al. (2009) De novo transcriptome assembly with ABySS. *Bioinformatics* 25(21):2872-2877.
10. Robertson G, et al. (2010) De novo assembly and analysis of RNA-seq data. *Nat Meth* 7(11):909-912.
11. Keeling PJ, et al. (2014) The Marine Microbial Eukaryote Transcriptome Sequencing Project (MMETSP): Illuminating the Functional Diversity of Eukaryotic Life in the Oceans through Transcriptome Sequencing. *PLoS Biol* 12(6):e1001889.
12. Kanehisa M, Furumichi M, Tanabe M, Sato Y, Morishima K (2017) KEGG: new perspectives on genomes, pathways, diseases and drugs. *Nucleic Acids Res* 45(D1):D353-D361.
13. Gremme G, Steinbiss S, Kurtz S (2013) GenomeTools: a comprehensive software library for efficient processing of structured genome annotations. *IEEE/ACM Trans Comput Biol Bioinform* 10(3):645-656.
14. Patro R, Duggal G, Love MI, Irizarry RA, Kingsford C (2017) Salmon provides fast and bias-aware quantification of transcript expression. *Nat Meth* 14(4):417-419.
15. Love MI, Huber W, Anders S (2014) Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* 15(12):550.
16. Benjamini Y, Hochberg Y (1995) Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J R Stat Soc Ser B Methodol* 57:289-300.
17. Morrissey J, et al. (2015) A Novel Protein, Ubiquitous in Marine Phytoplankton, Concentrates Iron at the Cell Surface and Facilitates Uptake. *Curr Biol* 25(3):364-371.
18. Twining BS, et al. (2011) Metal quotas of plankton in the equatorial Pacific Ocean. *Deep-Sea Res Pt II* 58(3–4):325-341.
19. Twining BS, et al. (2003) Quantifying Trace Elements in Individual Aquatic Protist Cells with a Synchrotron X-ray Fluorescence Microprobe. *Anal Chem* 75(15):3806-3816.
20. Nunez-Milland DR, Baines SB, Vogt S, Twining BS (2010) Quantification of phosphorus in single cells using synchrotron X-ray fluorescence. *J Synchrotron Radiat* 17(4):560-566.

21. Twining BS, Baines SB, Fisher NS, Landry MR (2004) Cellular iron contents of plankton during the Southern Ocean Iron Experiment (SOFeX). *Deep-Sea Res Pt I* 51(12):1827-1850.
22. Parsons TR, Maita Y, Lalli CM (1984) *A Manual of Chemical & Biological Methods for Seawater Analysis*, (Pergamon, Amsterdam).
23. Barwell-Clarke J, Whitney F (1996) IOS nutrient methods and analysis. *Can Tech Rep Hydrog Ocean Sci* 182:1-43.
24. Garcia HE, R. A. Locarnini, T. P. Boyer, J. I. Antonov, O.K. Baranova, M.M. Zweng, J.R. Reagan, D.R. Johnson (2014) World Ocean Atlas 2013, Volume 4: Dissolved Inorganic Nutrients (phosphate, nitrate, silicate). *NOAA Atlas NESDIS 76*, eds S. Levitus & Mishonov A), p 25.
25. Hunter JD (2007) Matplotlib: A 2D Graphics Environment. *Comput Sci Eng* 9(3):90-95.
26. Biller DV, Bruland KW (2012) Analysis of Mn, Fe, Co, Ni, Cu, Zn, Cd, and Pb in seawater using the Nobias-chelate PA1 resin and magnetic sector inductively coupled plasma mass spectrometry (ICP-MS). *Mar Chem* 130:12-20.
27. Parker CE, Brown MT, Bruland KW (2016) Scandium in the open ocean: A comparison with other group 3 trivalent metals. *Geophys Res Lett* 43(6):2758-2764.
28. Milne A, Landing W, Bizimis M, Morton P (2010) Determination of Mn, Fe, Co, Ni, Cu, Zn, Cd and Pb in seawater using high resolution magnetic sector inductively coupled mass spectrometry (HR-ICP-MS). *Anal Chim Acta* 665(2):200-207.
29. Luisa O, Gabriele P, Diana S, Marina M (2004) Multiple rDNA ITS-types within the diatom Pseudo-nitzschia delicatissima (Bacillariophyceae) and their relative abundances across a spring bloom in the Gulf of Naples. *Mar Ecol-Prog Ser* 271:87-98.
30. Hubbard KA, Rocap G, Armbrust EV (2008) Inter- and intraspecific community structure within the diatom genus *Pseudo-nitzschia* (Bacillariophyceae). *J Phycol* 44(3):637-649.
31. Hubbard KA, Olson CE, Armbrust EV (2014) Molecular characterization of Pseudo-nitzschia community structure and species ecology in a hydrographically complex estuarine system (Puget Sound, Washington, USA). *Mar Ecol-Prog Ser* 507:39-55.
32. Marchetti A, et al. (2008) Identification and assessment of domoic acid production in oceanic *Pseudo-nitzschia* (Bacillariophyceae) from iron-limited waters in the Northeast Subarctic Pacific. *J Phycol* 44(3):650-661.
33. Cohen NR, et al. (2018) Iron storage capacities and associated ferritin gene expression among marine diatoms. *Limnol Oceanogr* 63:1677-1691.
34. Moreno CM, et al. (2017) Examination of gene repertoires and physiological responses to iron and light limitation in Southern Ocean diatoms. *Polar Biol* 41:679-696.
35. Edgar RC (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res* 32(5):1792-1797.
36. Stamatakis A (2014) RAxML Version 8: A tool for Phylogenetic Analysis and Post-Analysis of Large Phylogenies. *Bioinformatics* 30(9):1312-1313.
37. Matsen FA, Kodner RB, Armbrust EV (2010) pplacer: linear time maximum-likelihood and Bayesian phylogenetic placement of sequences onto a fixed reference tree. *BMC Bioinform* 11(1):538.
38. Han MV, Zmasek CM (2009) phyloXML: XML for evolutionary biology and comparative genomics. *BMC Bioinform* 10(1):1-6.

39. Thomine S, Wang R, Ward JM, Crawford NM, Schroeder JI (2000) Cadmium and iron transport by members of a plant metal transporter family in Arabidopsis with homology to Nramp genes. *Proc Nat Acad Sci* 97(9):4991-4996.