

DISTINCT STRATEGIES BY BLOOM-FORMING DIATOMS TO FREQUENTLY-ENCOUNTERED
SHIFTS IN THEIR ENVIRONMENT

Robert H. Lampe

A thesis submitted to the faculty at the University of North Carolina at Chapel Hill in partial fulfillment of
the requirements for the degree of Master of Science in the Department of Marine Sciences in the
College of Arts and Sciences.

Chapel Hill
2018

Approved by:

Adrian Marchetti

Scott M. Gifford

Harvey Seim

© 2018
Robert H. Lampe
ALL RIGHTS RESERVED

ABSTRACT

Robert H. Lampe: Distinct strategies by bloom-forming diatoms to frequently-encountered shifts in their environment
(Under the direction of Adrian Marchetti)

In the world's oceans, diatoms perform approximately 40% of the total primary production and are the most prominent group of eukaryotic phytoplankton. Recent advances in diatom genomics are revealing numerous insights into diatoms' unique evolution and metabolic adaptations that contribute to their ecological success. Here I present two studies that explore some of these distinct strategies within bloom-forming diatoms in their natural environment by combining environmental transcriptomics with additional measures of diatom physiology. First, the response to upward vertical transport during coastal upwelling events was examined. Diatoms display a distinct transcriptional response that includes frontloading nitrogen-related genes in order to outcompete other groups. Laboratory-based simulations of upwelling show that this diatom response occurs over relatively short time scales. Secondly, iron storage mechanisms were investigated with iron addition and removal incubations from varying iron environments. We show that a specific storage mechanism, the protein ferritin, may provide a competitive advantage for ferritin-utilizing diatoms in areas of the world's oceans that undergo prolonged iron limitation with pulsed iron inputs. Together, these studies provide insight into the fundamental ecological question of why diatoms are highly successful in response to frequently-encountered abiotic changes.

ACKNOWLEDGEMENTS

I wish to thank my advisor, Dr. Adrian Marchetti, and committee members, Dr. Scott Gifford and Dr. Harvey Seim, for their guidance and support in the completion of this thesis. I would also like to thank my undergraduate research advisor from North Carolina State University, Dr. Astrid Schnetzer, for providing me with an introduction to the wonders of diatom research.

The research in this thesis was supported by many collaborators across several institutions. In particular, Liz Mann (Bigelow), Ben Twining (Bigelow), and Natalie Cohen (UNC) made significant contributions to the work presented in Chapters 1 and 2. Gus Hernandez (UNC) was instrumental in completing Appendix C. Extremely valuable contributions were also made by Ken Bruland (UCSC, Ch. 1 & 2), Mark Brzezinski (UCSB, Ch. 1 & 2), Kelsey Ellis (UNC, Ch. 1), Maite Maldonado (UBC, Ch. 1), Claire Till (HSU, Ch. 1 & 2), and Kim Thamatrakoln (Rutgers, Ch. 1 & 2).

This thesis was also directly supported by a fellowship from the UNC Graduate School and the National Science Foundation via the Graduate Research Fellowship Program (DGE-1650116) and OCE-1334935 (to A.M.). Data presented within this thesis was also funded by National Science Foundation grants OCE-1259776 (to K.B.), OCE-1334387 (to M.B.), OCE-1341479 (to A.M.), OCE-1333929 (to K.T.), OCE-1334632 (to B.T.), and Natural Sciences and Engineering Research Council of Canada (NSERC) grant 261521-13 (to M.M.)

Lastly, I would like to thank Molly Matty for always encouraging me to be a better scientist.

TABLE OF CONTENTS

LIST OF TABLES.....	viii
LIST OF FIGURES.....	ix
LIST OF ABBREVIATIONS.....	xii
CHAPTER 1: DIVERGENT GENE EXPRESSION AMONG PHYTOPLANKTON TAXA IN RESPONSE TO UPWELLING	1
Introduction	1
Materials and Methods	2
Sample Collection	2
Chlorophyll	3
Biogenic Silica.....	3
Particulate Carbon, Particulate Nitrogen, and Nitrate Uptake	3
Dissolved Inorganic Carbon Uptake	4
RNA Extraction and Sequencing	4
Metatranscriptome Assembly, Annotation, and Read Quantification	5
Differential Expression Analysis.....	5
Data Deposition.....	6
Results and Discussion	7
Experimental Overview and Physiological Observations.....	7
Taxonomic Composition	9
Comparative Gene Expression of Phytoplankton Groups	10
Molecular Characterization of the Nitrogen Assimilation Response	13
<i>Chaetoceros</i> and <i>Pseudo-nitzschia</i> Expressed Genes	16
Conclusions.....	18
References.....	19

CHAPTER 2: DIFFERENT IRON STORAGE STRATEGIES AMONG BLOOM-FORMING DIATOMS	23
Introduction	23
Materials and Methods	25
Experimental Design	25
RNA-Seq Overview	26
Synchrotron X-ray Fluorescence (SXRF) Preparation and Analysis	28
Data Deposition.....	28
Results and Discussion	28
Iron States within the Initial Phytoplankton Communities and Incubations.....	28
Taxonomic Distributions.....	31
Iron Quotas among Diatom Genera.....	33
Ferritin Expression in <i>Pseudo-nitzschia</i> and Other Diatoms	35
Vacuolar-associated Gene Expression.....	39
Biogeochemical and Ecological Implications	41
References.....	44
APPENDIX A: CHAPTER 1 SUPPLEMENTARY INFORMATION	48
Materials and Methods.....	48
Text	53
Figures	56
Tables.....	69
References.....	73
APPENDIX B: CHAPTER 2 SUPPLEMENTARY INFORMATION	75
Materials and Methods.....	75
Text	76
Figures	79
Tables.....	95
References.....	99
APPENDIX C: CHANGES IN GROWTH AND ELEMENTAL COMPOSITION OF A DIATOM AND HAPTOPHYTE THROUGHOUT A SIMULATED UPWELLING CONVEYER BELT CYCLE	100

Introduction	100
Materials and Methods	101
Results and Discussion	104
References.....	108

LIST OF TABLES

Table 2.1 – Incubation site identifiers, location, and initial iron conditions	26
Table A.1 – Summary of environmental RNA sequencing with Illumina HiSeq 2000.....	69
Table A.2 – Summary of transcriptome sequencing, assembly, and gene prediction for isolates from the California Upwelling Zone.....	70
Table A.3 – Summary of functional and taxonomic annotation for the environmental RNA assembly	71
Table A.4 – Cell abundances from microscopic counts	72
Table B.1 – Time points and initial collection depths for each incubation site.....	95
Table B.2 – Significance test results for differences in site parameters	96
Table B.3 – Significance in differential expression of <i>FLDA1</i> and <i>ISIP3</i> expression for <i>Thalassiosira</i> and <i>Chaetoceros</i> determined by DESeq2 and shown in Figure B.3	97
Table C.1 – Description of sampling time points throughout the UCBC simulations.....	103

LIST OF FIGURES

Figure 1.1 - Measurements from the initial upwelled water and incubations at 72 and 120 hours	7
Figure 1.2 – Average taxonomic distribution by mapped reads from each time point.....	9
Figure 1.3 – KEGG Ortholog (KO) gene expression comparison among the four main detected phytoplankton groups.....	11
Figure 1.4 – Diatom gene expression	15
Figure 2.1 – Map of field incubation locations with interpolated statistical means of nitrate concentrations ($\mu\text{mol L}^{-1}$) on a 1° grid of all decades from World Ocean Atlas 2013	26
Figure 2.2 - Macronutrient ($\mu\text{mol L}^{-1}$), size-fractionated chlorophyll a ($\mu\text{g L}^{-1}$), and biogenic silica ($\mu\text{mol L}^{-1}$) concentrations and maximum photochemical yields of photosystem II ($F_v:F_m$) from the initial collected seawater ($t = 0$) and incubations at various time points	30
Figure 2.3 - Differences in relative proportions of taxonomically annotated transcripts between the iron addition and control (Fe-C), control and iron removal (C-DFB), and iron addition and iron removal (Fe-DFB) at each site	32
Figure 2.4 – (a) Cellular iron quotas ($\mu\text{mol Fe:mol C}$) in <i>Chaetoceros</i> and <i>Pseudo-nitzschia</i> from incubations. (b) Time course for site C-Low1	33
Figure 2.5 – <i>Pseudo-nitzschia</i> normalized ferritin (<i>FTN</i>) expression for all initial conditions and incubations.....	36
Figure 2.6 - Heatmap of \log_2 fold change values for ferritin (<i>FTN</i>), natural resistance-associated macrophage protein (<i>NRAMP</i>), and <i>ZIP1</i> expression in <i>Pseudo-nitzschia</i> , <i>Chaetoceros</i> , and <i>Thalassiosira</i>	38
Figure A.1 – Satellite-derived sea surface temperature prior to the incubations.....	56
Figure A.2 – Wind speed (m s^{-1}) and direction from the R/V Melville for the 48 hours prior to starting the incubations	57
Figure A.3 - CTD (conductivity-temperature-depth) measurements	58
Figure A.4 - Temperature and on-deck photosynthetically active radiation (PAR) during the incubations	59
Figure A.5 - Biological coefficient of variation (square root of dispersion) for diatom genes.....	60
Figure A.6 - Additional measurements from the simulated upwelling incubation experiment	61
Figure A.7 - Histograms of KEGG Ortholog (KO) gene expression between 72 and 120 hours for the four main phytoplankton groups detected in the study	62
Figure A.8 - Heatmaps for expressed KOs at 72 and 0 hours	63

Figure A.9 - Counts and proportions of significantly overrepresented KEGG orthologs within KEGG module class level two groups at 0 and 72 hours	64
Figure A.10 - Heatmap of nitrogen related genes for <i>Chaetoceros</i> , <i>Pseudo-nitzschia</i> , other diatoms, dinoflagellates, chlorophytes, and haptophytes at 0 and 72 hours	65
Figure A.11 - Phylogenetic tree of guanylate cyclase soluble subunit beta (GUCY1B)	66
Figure A.12 - Phylogenetic tree of nitrate reductase (NR) in diatoms	67
Figure A.13 - Quantitative Metabolic Fingerprint (QMF) depicting relative expression of KEGG modules for each of the four major phytoplankton groups at each time point.	68
Figure B.1 – Bathymetry of the California Upwelling Zone	79
Figure B.2 – Molecular indicators of iron stress or limitation in diatoms.....	80
Figure B.2 – Average taxonomic distribution by mapped reads for the whole community (phylum-based) and diatom genera within all diatom assigned reads.....	81
Figure B.4 – Light and synchrotron X-ray fluorescence micrographs of <i>Chaetoceros</i> and <i>Pseudo-nitzschia</i> cells from the control and iron (+Fe) incubations at 48 hours at C-Low1	82
Figure B.5 – Normalized transcript abundances for ferritin (<i>FTN</i>), natural resistance-associated macrophage protein (<i>NRAMP</i>), and <i>ZIP1</i> in <i>Pseudo-nitzschia</i> , <i>Chaetoceros</i> , and <i>Thalassiosira</i>	83
Figure B.6 – Phylogenetic tree of <i>Pseudo-nitzschia</i> internal transcribed spacer 1 (<i>ITS1</i>) sequences with environmental sequences from the metatranscriptome.....	84
Figure B.7 – Phylogenetic tree of <i>Pseudo-nitzschia</i> ferritin sequences with environmental sequences from the metatranscriptome.....	85
Figure B.8 – Fold changes in <i>Pseudo-nitzschia</i> <i>FTN</i> expression versus fold changes in iron quotas for the Fe or DFB treatments compared to the control treatment	86
Figure B.9 – Midpoint-rooted phylogenetic tree of diatom ferritins	87
Figure B.10 – Midpoint-rooted phylogenetic tree of diatom ferritins	88
Figure B.11 – Normalized transcript abundances for ferritin (<i>FTN</i>) and natural resistance-associated macrophage protein (<i>NRAMP</i>) in <i>Fragilaropsis</i>	89
Figure B.12 – Phylogenetic tree of natural resistance-associated macrophage protein (<i>NRAMP</i>) homologous genes	90
Figure B.13 – Normalized transcript abundances for <i>ZIP7</i> in <i>Pseudo-nitzschia</i> , <i>Chaetoceros</i> , and <i>Thalassiosira</i>	91
Figure B.14 – Midpoint-rooted phylogenetic tree of <i>ZIP1</i> homologous genes.....	92
Figure B.15 – Domoic acid (DA) concentrations from C-Low2	93
Figure B.16 – Three day average of satellite derived sea surface temperature from 03 June 2015 to 05 June 2015, the day of sampling for C-High.....	94

Figure C.1 – Conceptual model of the upwelling conveyer belt cycle	103
Figure C.2 – $F_v:F_m$ and C:N ratios for <i>C. decipiens</i> and <i>E. huxleyi</i> throughout the UCBC simulations under iron-replete and iron-stressed conditions	105
Figure C.3 – Raw fluorescence units (RFUs) of each sample of <i>C. decipiens</i> and <i>E. huxleyi</i> throughout the UCBC experiments under high and low iron.....	106

LIST OF ABBREVIATIONS

CUZ	California Upwelling Zone
DFB	Desferoxamine B
KEGG	Kyoto Encyclopedia of Genes and Genomes
KO	KEGG Ortholog
RFU	Raw fluorescence unit
SXRF	Synchrotron X-ray Fluorescence
TMC	Trace metal clean
UCBC	Upwelling conveyer belt cycle

CHAPTER 1: DIVERGENT GENE EXPRESSION AMONG PHYTOPLANKTON TAXA IN RESPONSE TO UPWELLING

Introduction

Wind-driven coastal upwelling associated with eastern boundary currents delivers rich supplies of nutrients to illuminated surface waters. This phenomenon provides ideal conditions for blooms of phytoplankton that render coastal upwelling regimes centers of new production even though their relative ocean area is small (Capone and Hutchins 2013). Typically dominated by large chain-forming diatoms, phytoplankton blooms in upwelling zones rapidly sequester carbon dioxide and are the base of short, efficient food chains that comprise a significant percentage of the global fish catch (Estrada and Blasco 1985, Lachkar and Gruber 2013, Lassiter et al 2006, Ryther 1969).

The phytoplankton community in upwelling zones is postulated to undergo a ‘conveyer belt cycle’ in which viable cells are upwelled into sunlit waters to seed a surface bloom. The community is then advected away from the upwelled source, and some cells eventually sink out of the photic zone. Surviving cells at depth and positioned in future upwelled waters are able to act as seed stock once winds are favorable for upwelling (Wilkerson and Dugdale 1987, Wilkerson and Dugdale 2008). This continuity between a subsurface population and surface bloom during an upwelling event has been observed through a combination of glider and remote sensing techniques (Seegers et al 2015).

The physiological response of phytoplankton to being vertically transported into a higher light and nutrient-rich environment, commonly referred to as shift-up, includes rapid growth rates and strong increases in nitrate uptake and assimilation (MacIsaac et al 1985, Wilkerson and Dugdale 1987). This nitrate-related activity has been repeatedly observed in simulated upwelling mesocosm experiments (Dugdale and Wilkerson 1989, Fawcett and Ward 2011) and in a laboratory experiment on the diatom *Skeletonema costatum* (Smith et al 1992). Shift-up, as expressed through rapid nitrate assimilation, is hypothesized to be linked to the success of diatoms in upwelling regions; it is believed that diatoms respond quickest to available nitrate once conditions are optimal (Fawcett and Ward 2011). Characterization of this

physiological response at a molecular level, however, is lacking. Only upregulation of the nitrogen assimilation gene, nitrate reductase, has been observed in *Skeletonema costatum* under lab-simulated upwelling conditions, indicating that there is a molecular basis for the shift-up response (Smith et al 1992).

Metatranscriptomics is increasingly being applied to eukaryotic phytoplankton communities to provide a deeper understanding of molecular responses among resident phytoplankton groups (Caron et al 2017). With the growing availability of reference transcriptomes and genomes of eukaryotic phytoplankton, unprecedented levels and confidence in gene annotation are being obtained from environmental sequences (Alexander et al 2015, Keeling et al 2014). Here we apply comparative metatranscriptomics to a simulated upwelling event in a shipboard incubation experiment to characterize the phytoplankton community's response and investigate the molecular basis for shift-up. Our results indicate that phytoplankton functional groups exhibit highly distinct transcriptional responses to being upwelled in which diatoms constitutively express genes involved in nitrogen assimilation. This strategy possibly allows diatoms to outcompete other groups for available nitrogen once physical conditions are optimal for growth.

Materials and Methods

Sample Collection

On 17 July 2014, upwelling conditions were not present at a site within the California Upwelling Zone ($35^{\circ} 56.071' N$, $121^{\circ} 44.022' W$; Figures A.1, A.2). At 05:00 PDT (12:00 GMT) at the same location, viable phytoplankton cells were detected via imaging flow cytometry (FlowCAM, Fluid Imaging Technologies Inc., Scarborough, ME, USA) at 96 m which corresponded to the $10^{\circ}C$ isotherm (Figure A.3). Seawater from this depth was processed immediately for the initial time point. To simulate upwelling, additional seawater from the same depth was filled into a large acid-rinsed HDPE barrel for homogenization, dispensed to triplicate 10 L Cubitainer® (Hedwin Corporation, Newark, DE, USA), and incubated in an on-deck plexiglass incubator with flow-through seawater at 33% incident irradiance. Seawater collection and incubations followed trace metal clean techniques as they were conducted as part of a larger study to examine diatom responses to iron addition or removal (Chapter 2) although for the purposes of this study, only the unamended control incubations are considered. Temperature and on deck irradiance values

throughout the incubation are provided in Figure A.4.

Based on macronutrient drawdown, triplicate cubitainer were harvested following 72 hours and 120 hours of incubation. Subsamples from each cubitainer were preserved or measured for chlorophyll *a*, species composition (by microscopy and FlowCAM), biogenic silica, F_v:F_m, domoic acid, nutrients, particulate carbon and nitrogen, carbon and nitrogen uptake, and RNA. Chlorophyll *a*, particulate carbon and nitrogen, and nitrate uptake rates were size fractionated using a series filter cascade. Carbon uptake rates were size fractionated using a mesh spacer. Additional methods are described in Appendix A.

Chlorophyll

Four hundred mL of seawater was gravity-filtered through a 5 µm polycarbonate filter (47 mm) followed by a GF/F filter (25 mm) under gentle vacuum pressure (<100 mm Hg). Filters were rinsed with 0.45 µm filtered seawater and immediately frozen at -80°C until analysis. Chlorophyll *a* extraction was performed using 90% acetone at -20°C for 24 h and measured via *in vitro* fluorometry on a 10-AU fluorometer (Turner Designs, San Jose, CA, USA) using the acidification method (Parsons et al 1984).

Biogenic Silica

Biogenic silica was determined via filtration of 335 mL onto 1.2 µm polycarbonate filters (45 mm). Concentrations were measuring using a NaOH digestion in teflon tubes (Krause et al 2009) and a colorimetric ammonium molybdate method (Brzezinski and Nelson 1995).

Particulate Carbon, Particulate Nitrogen, and Nitrate Uptake

Particulate nitrogen (PN), carbon (PC) and nitrate (NO₃⁻) uptake was obtained by spiking 618 mL of seawater with ¹⁵N-NaNO₃ at no more than 10% of ambient nitrate concentration and incubating for 8 hours in the flow-through plexiglass incubator. Following incubation, seawater filtration commenced immediately and was performed by gravity through a 5 µm polycarbonate filter (47 mm), and with an in-line vacuum (<100 mm Hg) onto a precombusted (450°C for 5 h) GF/F filter (25 mm). Cells on the 5 µm polycarbonate filter were then rinsed onto an additional precombusted GF/F filter (25 mm) using an artificial saline solution. Filters were then stored at -20°C. Prior to analysis, filters were dried at 50°C for 24 hours,

wrapped in tin capsules, and pelletized in preparation for analysis of the atom % ^{15}N , PN, and PC using an elemental analyzer paired with an isotope ratio mass spectrometer (EA-IRMS). Absolute uptake rates (ρ , NO_3 taken up per unit time) were calculated using a constant transport model (Eq. (3) from Dugdale and Wilkerson (1986)). Biomass-specific NO_3 uptake rates (V , NO_3 taken up per unit PN per unit time) were also calculated according to the constant specific uptake model (Eq. (6) Dugdale and Wilkerson (1986)). $^{15}\text{NO}_3$ uptake rates were not corrected for possible losses of ^{15}N in the form of dissolved organic nitrogen (Bronk et al 1994); therefore, the reported values are considered conservative estimates or net uptake.

Dissolved Inorganic Carbon Uptake

Sixty mL samples from each cubitainer were distributed into light and dark bottles cleaned with 1.2 mol L $^{-1}$ HCl. For each bottle, 1.2 μCi of $\text{NaH}^{14}\text{CO}_3$ was added and mixed. A 1 mL subsample was taken and added to vials containing 6 mol L $^{-1}$ NaOH to trap and validate the initial inorganic H^{14}CO_3 quantities. The light and dark bottles were incubated on-deck for 6.5-8 hours. Samples were filtered onto stacked polycarbonate filters (5 μm and 1 μm) separated with a mesh spacer. Blank control bottles also containing 1.2 μCi of $\text{NaH}^{14}\text{CO}_3$ were filtered onto a GF/F filter after 5 minutes and had counts similar to dark bottles. Filters were vacuumed dried, placed in scintillation vials with 0.5 mL of 6 mol L $^{-1}$ HCl, permitted to degas for 24 hours, and counted using a Beckman Coulter LS 6500 scintillation counter. Reported values are light bottles minus dark bottles. Biomass-specific dissolved inorganic carbon (DIC) uptake rates (V_{DIC}) were calculated by normalizing DIC uptake to PC.

RNA Extraction and Sequencing

Seawater was filtered onto 0.8 μm Pall Supor® filters (142 mm) using a peristaltic pump then immediately flash frozen in liquid nitrogen. RNA was extracted using the ToTALLY RNA Total RNA Isolation Kit and treated with DNase 1 (Life Technologies, Grand Island, NY, USA). The extraction procedure was followed according to the manufacturer's instructions with additional first step of glass bead addition to assist with organic matter disruption. RNA quantity and purity was assessed prior to sequencing on an Agilent Bioanalyzer 2100. Total RNA from the triplicate samples for the initial time point (T0) and the first time point (T72) were pooled into one sample due to low RNA yields. Triplicate samples were maintained

for the second time point (T120). Library prep 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).

Metatranscriptome Assembly, Annotation, and Read Quantification

Reads were trimmed for quality and adapter removal 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 paired reads that overlap were merged into single reads with BBMerge v8.0. Merged pairs and non-overlapping paired-end reads were then used to assemble contigs using AbySS v1.5.2 with varied k-mer sizes (Birol et al 2009). Assemblies were merged using Trans-AbySS v1.5.3 to remove redundant contigs (Robertson et al 2010), and those shorter than 125 bp were discarded. Read counts were obtained by mapping raw reads to contigs with Bowtie2 v2.2.6 (Langmead and Salzberg 2012) and filtered by mapping quality with SAMtools v1.2 (Li et al 2009). Mapping percentages are provided in Table A.1.

Annotation was assigned by best homology (lowest E-value) to protein databases using BLASTX v2.2.31 (E-value $\leq 10^{-5}$). For taxonomic identification, MarineRefII, a custom reference database was used. MarineRefII contains predicted protein sequences of marine microbial eukaryotes and prokaryotes including all sequenced transcriptomes from the Marine Microbial Eukaryote Transcriptome Sequencing Project (Keeling et al 2014). MarineRefII was supplemented with transcriptomes of isolated phytoplankton from these incubations adding increased confidence in the taxonomic annotation of some contigs (Table A.2). To assign gene function to contigs, the same methodology with the Kyoto Encyclopedia of Genes and Genomes (KEGG; Release 75) was used (Kanehisa et al 2017). The best hit with a KEGG Ortholog (KO) number from the top 10 hits was chosen. Similarly, analysis of module annotations (MO) was conducted by selecting the top BLASTX hit with a KEGG MO number from the top 10 hits. A summary of annotation results is provided in Table A.3.

Differential Expression Analysis

Differential expression was assessed by summing read counts of contigs within a taxonomic group (phylum-based or genus for only the diatoms, *Chaetoceros* and *Pseudo-nitzschia*) by KEGG Gene

Definition or KEGG Orthology (KO) annotation. EdgeR v3.12.0 was used to calculate normalized fold change and counts-per-million (CPM) from pairwise comparisons within each taxonomic group using the exactTest function (Robinson and Smyth 2008, Robinson et al 2010). Significance (P -value < 0.05) was calculated by using edgeR's estimate of tagwise dispersions utilizing the available replication of all treatments from the second time point (T120; Figure A.5; Chen et al 2014). ExactTest output in combination with the taxonomic distributions per gene were plotted using a custom plotting function available at <https://github.com/marchettilab/mantaPlot>.

Shared expression of gene was considered when a gene was detected in at least one of the libraries under comparison for each taxonomic group. For binning of genes displayed in heatmaps, a positive or negative fold change, variance greater than the number of taxonomic groups, and fold change greater than or less than all other groups were used. Genes with a \log_2 fold change greater than 2 or less than -2 but had a variance less than the number of taxonomic groups were considered similarly overrepresented by all groups. Otherwise, the expression level was considered similar on the basis of fold change. These data were visualized with pheatmap v1.0.8.

Data Deposition

The data reported in this paper have been deposited in the National Center for Biotechnology (NCBI) sequence read archive under the accession no. SRP074302 (BioProject accession no. PRJNA320398). Assembled contigs, read counts, and annotations are available at <https://marchettilab.web.unc.edu/data/>. Isolate 18S sequences, transcriptome raw reads, assemblies, and predicted peptide sequences are deposited in Cyverse (<http://www.cyverse.org>) under the project name unc_phyto_isolates (Appendix A). Isolate 18S sequences are also deposited in Genbank (accession nos. KX229684-KX229691).

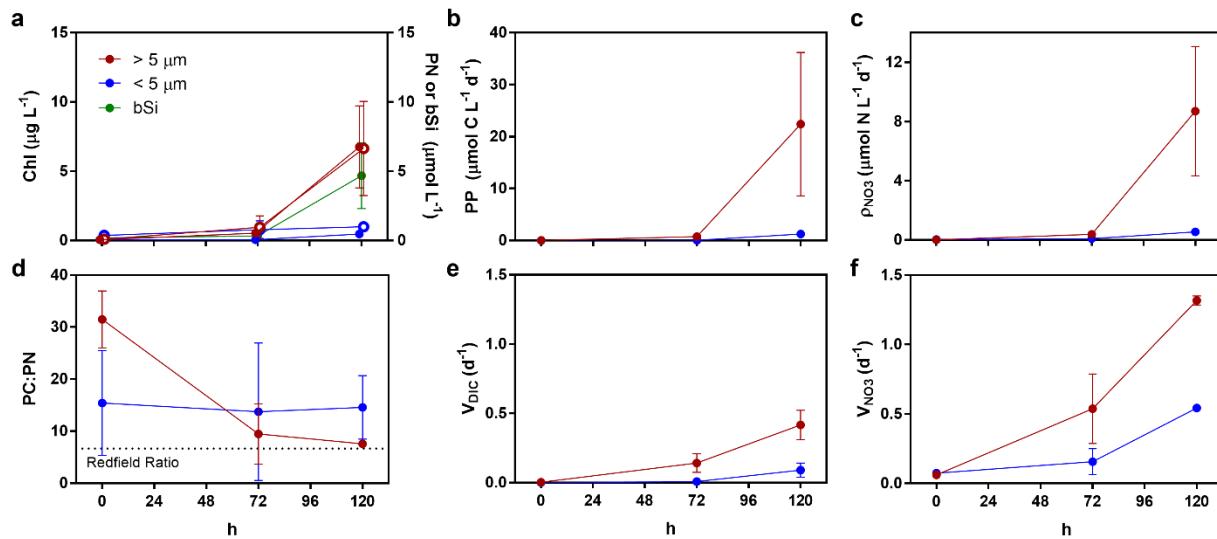


Figure 1.1 - Measurements from the initial upwelled water and incubations at 72 and 120 hours: >5 µm (red) and <5 µm (blue). (A) Chlorophyll a (closed circles), particulate nitrogen (open circles), and biogenic silica (green). (B) Primary productivity expressed as inorganic carbon uptake ($\mu\text{mol C L}^{-1} \text{d}^{-1}$). (C) Absolute nitrate (NO_3^-) uptake rates (ρ, NO_3^- taken up per unit time). (D) Ratios of particulate carbon to particulate nitrogen. (E) Biomass-specific dissolved inorganic carbon uptake rates (V_{DIC}), i.e. inorganic carbon uptake normalized to biomass as particulate carbon ($\mu\text{mol C L}^{-1} \text{d}^{-1} / \mu\text{mol C L}^{-1} \text{ or } \text{d}^{-1}$) (F) Biomass-specific nitrate uptake rates ($V_{\text{NO}_3^-}$), i.e. nitrate uptake rates normalized to biomass as particulate nitrogen ($\mu\text{mol N L}^{-1} \text{d}^{-1} / \mu\text{mol N L}^{-1} \text{ or } \text{d}^{-1}$). Error bars indicate standard deviation of the mean ($n = 3$).

Results and Discussion

Experimental Overview and Physiological Observations

Results from the simulated upwelling experiment indicate that a bloom of large phytoplankton (>5 µm) was induced with observations of shift-up in terms of growth, primary production, and nitrate uptake within these large cells. Macronutrient concentrations in the upwelled waters remained high throughout the incubations; however, significant growth in the large (>5 µm) phytoplankton community was observed (Figure 1.1a and Figure A.6a). The initial dissolved iron concentration was approximately 1.28 nmol L⁻¹ which is marginally higher than the normal values (<1 nmol L⁻¹) observed in the region. For complete drawdown of nitrate, an iron to nitrate ratio of 8 nmol L⁻¹:20 µmol L⁻¹ is typically required (Bruland et al 2001). The initial ratio of 1.28 nmol L⁻¹:21.86 µmol L⁻¹ in the incubations therefore indicates that iron had the potential to be a limiting nutrient which resulted in 15 µmol L⁻¹ of unused nitrate after 120 hours.

In addition to chlorophyll a, the upwelling simulation produced notable increases in biogenic silica suggesting that the phytoplankton growth may be attributed to diatoms (Figure 1.1a). The presence of

domoic acid also suggests a considerable presence of the diatom genus *Pseudo-nitzschia* (Lelong et al 2012). Concentrations were below the detection limit until 120 hours when the average concentration was $1.34 \mu\text{g L}^{-1}$ (Figure A.6b). Although this concentration is lower than peak concentrations observed during blooms in California coastal waters (Schnetzer et al 2013), it nevertheless indicates that *Pseudo-nitzschia* spp. were abundant within the incubations by 120 hours.

This success of large phytoplankton is consistent with previous studies showing phytoplankton from large size fractions as the significant contributors to growth and new production during upwelling. Large phytoplankton consistently have greater increases in biomass and outcompete small cells for nutrients during a bloom (Fawcett and Ward 2011, Wilkerson et al 2000). These large phytoplankton are commonly characterized as chain-forming colonial diatoms such as *Chaetoceros* spp. but include *Pseudo-nitzschia* spp. (Estrada and Blasco 1985, Lassiter et al 2006).

The large phytoplankton community also exhibited clear physiological responses to being upwelled. Maximum photochemical yields ($F_v:F_m$) of the whole community increased from 0.25 to 0.51 within the first 72 hours (Figure A.6c). Dissolved inorganic carbon and nitrate uptake in the large cells increased throughout the experiment and was significantly higher than the small cells (Figures 1.1b, 1.1c). The particulate carbon-to-nitrogen ratio (C:N) was initially 31.5:1 in the $>5 \mu\text{m}$ size fraction but decreased to approach the expected elemental composition of 6.6:1 (Redfield et al 1963), while C:N values remained fairly constant and above the Redfield ratio for the small size fraction (Figure 1.1d). This return to Redfield stoichiometry for the larger community was coupled with increasing biomass-specific NO_3^- uptake rates ($V_{\text{NO}_3^-}$) that were approximately double that of biomass-specific carbon uptake rates (V_{DIC}) (Figures 1.1e, 1.1f).

These data are also a clear indication of a positive response from light limitation, or potentially a resting stage, to high growth for the larger phytoplankton. The initial low $F_v:F_m$ signifies that the community was stressed but quickly able to return to higher photosynthetic efficiencies. A high initial C:N ratio that approaches the Redfield-predicted value has also been observed in similar mesocosm experiments (Fawcett and Ward 2011, Kudela and Dugdale 2000). These studies suggest that the initial high C:N ratio indicates severe N limitation which likely occurred as the phytoplankton cells in aged upwelling water began to sink to depth. Once released from light limitation, the community is able to stabilize with large

phytoplankton controlling the total C:N as time progresses. It is also possible that there was C-rich detrital material elevating the initial measurement (Fawcett and Ward 2011), but acceleration of nitrate uptake, especially in relation to carbon uptake, to drive the community toward the Redfield-predicated ratio is clear. The larger cells are able to take advantage of nitrate as conditions become optimal and dominate the community since they uptake nitrate at higher rates than the smaller cells.

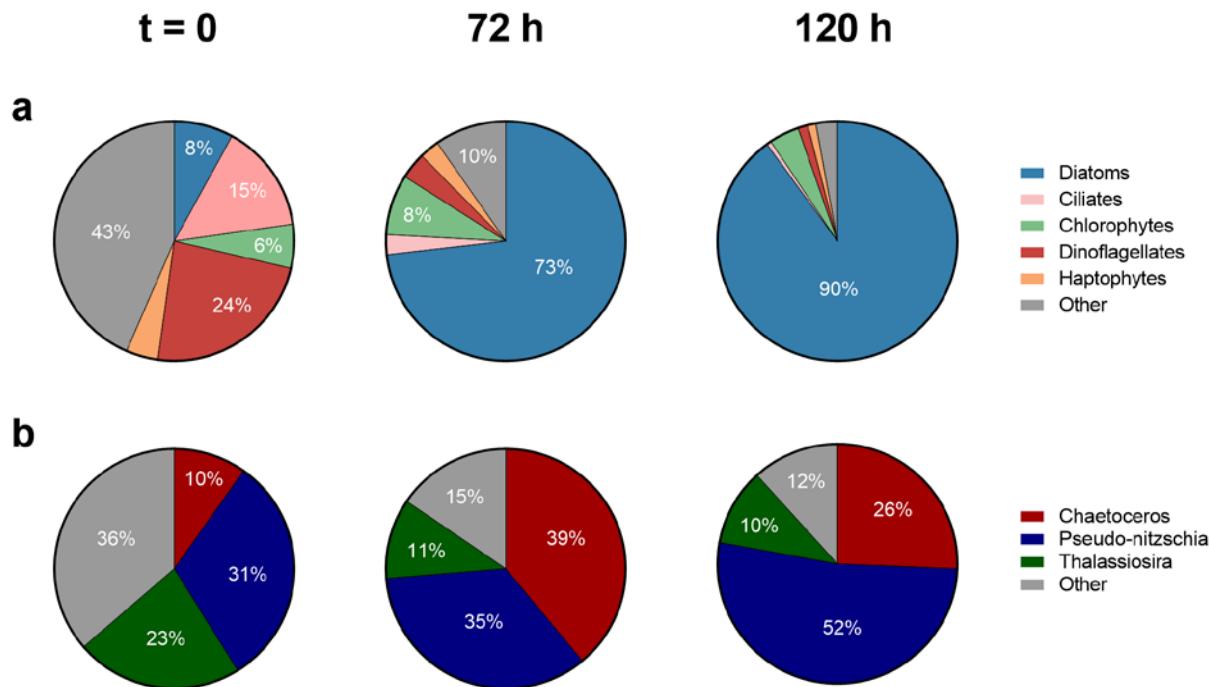


Figure 1.2 - Average taxonomic distribution by mapped reads from each time point. (A) Percentage of reads from the whole community. (B) Percentage of reads for diatom genera within all reads assigned as diatoms.

Taxonomic Composition

Metatranscriptome assembly resulted in 3.1 million contigs with levels of annotation similar to previous studies utilizing KEGG and reference transcriptomes from the Marine Microbial Eukaryote Transcriptome Sequencing Project (MMETSP; Table A.3; Alexander et al 2015, Cohen et al 2017, Keeling et al 2014). Obtaining taxonomically-annotated mRNA read counts allows for inference of relative taxonomic composition which is supported by microscopic cell counts and observations from an imaging flow cytometer (FlowCAM; Figure 1.2a and Table A.4). The initial community biomass and cell abundance were relatively low but quite diverse (Figure 1.1a and Table A.4). Relative abundances of transcripts

suggests that dinoflagellates were the dominant phytoplankton group although microscopic cell counts suggest chlorophytes may have been more abundant. Satellite-derived sea surface temperature and ship-board wind data indicate that upwelling-favorable conditions were not present for 13 days before the incubations (Figures A.1, A.2). Although multiple factors contribute to the residence time of cells at depth, these data suggest an upper limit of 13 days prior to sampling. By 72 and 120 hours following incubation, there was an overwhelming increase in the abundance of mRNA reads attributable to diatoms (Figure 1.2a) consistent with the bulk measurements, FlowCAM, and previous studies (Estrada and Blasco 1985): diatoms were unequivocally the dominant group within the simulated upwelling event.

The taxonomic composition of diatoms followed a similar trend as the whole community with an initially more diverse diatom community that transitioned into one dominated by just two genera: *Chaetoceros* and *Pseudo-nitzschia* (Figure 1.2b). *Chaetoceros* appeared to make rapid early gains but the community became mostly *Pseudo-nitzschia* by 120 hours. These two genera were also dominant within a previous mesocosm experiment examining shift-up at a nearby coastal California site (Kudela and Dugdale 2000). *Chaetoceros* spp. were found as resting spores and may quickly germinate following upwelling to make early gains in cell abundance (Pitcher 1990). Although a resting stage for *Pseudo-nitzschia* spp. is not known (Lelong et al 2012), they are significant members of the phytoplankton community throughout the upwelling cycle and dominated after 120 hours consistent with a peak in the *Pseudo-nitzschia* produced toxin, domoic acid (Figure A.6b). The presence of *Pseudo-nitzschia* is unsurprising considering the reports that members of this genus often dominate subsurface chlorophyll maxima (Ryan et al 2005), thin layers (McManus et al 2008, Rines et al 2002), and upwelled communities (Seegers et al 2015) that often result in harmful algal blooms in this region.

Comparative Gene Expression of Phytoplankton Groups

Examining shifts in the total transcript pool provides a broad depiction of the responsiveness of different groups. By comparing expression levels at 0 and 72 hours, these shifts reveal the initial whole transcriptome responses to simulated upwelling by the main detected groups of phytoplankton (Figure 1.3a). Diatoms had a high proportion of overrepresented genes after upwelling compared to other groups, over 950 (20%) of which were significantly overrepresented (P -value < 0.05). Dinoflagellates showed an

opposite pattern with gene expression skewed towards overrepresentation in the pre-upwelled condition while haptophytes had an even distribution of overrepresented genes in both phases. Interestingly, chlorophytes also had a higher number of significantly overrepresented genes post-upwelling, and they were able to maintain their relative proportion of the overall transcript pool unlike the dinoflagellates and haptophytes.

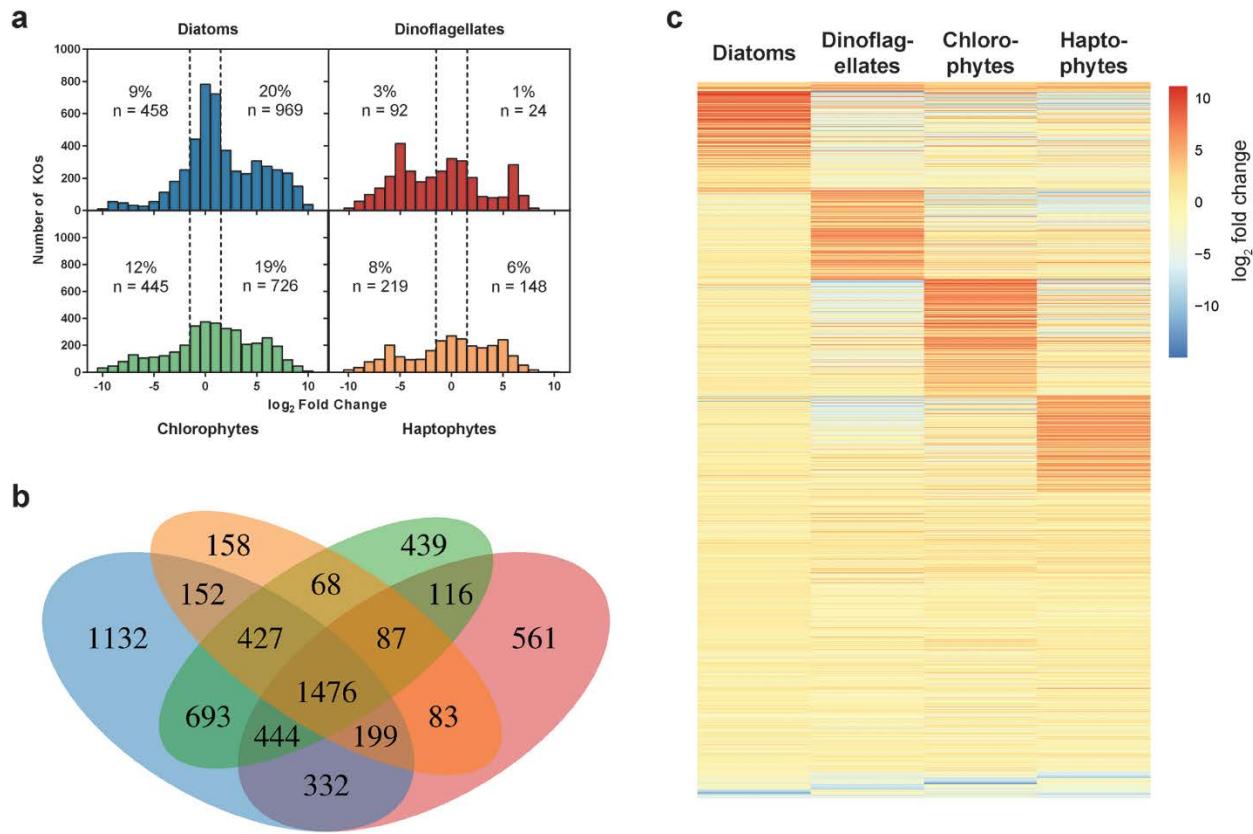


Figure 1.3 - KEGG Ortholog (KO) gene expression comparison among the four main detected phytoplankton groups: diatoms (blue), dinoflagellates (red), chlorophytes (green), and haptophytes (orange). (A) Histograms of KO counts binned by \log_2 fold change intervals of 1 for 0 and 72 hours. Dashed vertical lines indicate a \log_2 fold change of -1 or 1. The number and percentage of only the significantly ($P\text{-value} < 0.05$) overrepresented genes at 0 hours (pre-upwelling; left) and 72 hours (post-upwelling; right) are annotated on each plot. (B) Venn diagram of expressed KOs at 72 and 0 hours for each group. (C) Heatmap for the 1,476 commonly expressed KOs at 0 and 72 hours. Each row indicates an expressed KO with darker red (positive fold change) indicating overrepresentation at 72 hours and darker blue (negative fold change) indicating overrepresentation at 0 hours.

Smaller changes across all four groups were observed when examining shifts from 72 to 120 hours (Figure A.7). Relatively minor changes in the whole transcript pool and a less pronounced taxonomic shift from 72 to 120 hours indicates that most of the activity in relation to diatom dominance likely occurred in

the first 72 hours. This timing and slowing of response also corresponds to field observations that predict a 5-7 day window for cells to achieve balanced growth and transition from shift-up to a low nutrient shift-down (Dugdale and Wilkerson 1989, Wilkerson et al 2006). It has been speculated that these shifts, or variable transcript allocation, are a reflection of r- and K-type growth strategies (Alexander et al 2015). Our observations appear to follow this paradigm with diatoms exhibiting r-type growth and the highest transcript reallocation in terms of gene count.

Analysis of the expression of genes with shared KEGG Orthology (KO) annotation allows for direct comparisons between taxonomic groups as orthologs normally retain the same function throughout evolutionary history. Similar or different expression of a gene among groups may signify correspondingly similar or different investments in cellular processes at given time points. We detected 1,476 orthologous genes expressed by all four taxonomic groups at 0 or 72 hours (Figures 1.3b, 1.3c). Only 18 genes were binned as highly overrepresented at 72 hours by all four phytoplankton groups, of which many were related to chlorophyll synthesis. Over 550 genes had low absolute fold change values, many of them positive, across all four groups. These included more photosynthesis-related genes such as photosystem II constituents, photosynthesis electron transport proteins, light-harvesting chlorophyll protein complex proteins, and most of the genes associated with the Calvin cycle. The shared expression of these genes across groups is unsurprising considering the community is transitioning from a deep and dark environment to a sunlit environment, and would benefit from investing in photosynthetic machinery. Other genes that were highly expressed but showed little change in expression across all four groups were associated with other predictable cellular functions such as ribosomal proteins, translation initiation factors, and all constituents of the citric acid cycle.

Of particular interest is the clear overrepresentation at 72 hours of approximately 200 genes per taxonomic group that show little or negative fold change in the other three groups (Figure 1.3c). It is important to note that although differences in shifts in the total transcript pool were observed (Figure 1.3a), all groups are still responding and highly increasing their expression of a distinct set of genes compared to the other groups. This pattern continues to hold when examining the genes that were shared between diatoms and just one or two of the other groups (Figure A.8). The genes highly expressed by each group appear to be of diverse function as they do not cluster into certain categories or modules but can be broadly

interpreted as investments in different metabolic processes (Figure A.9). These unique responses may reflect fundamental differences in life strategies and ecological traits among functional groups.

To further explain the dominance of diatoms in these systems, expression of diatom annotated genes was investigated. 1,132 KOs were found solely in diatoms, likely due to the abundance of diatoms in our samples resulting in an improved metatranscriptome assembly for that group (Figure 1.3b). However, only 173 of these KOs were significantly overrepresented at either 0 or 72 h. It is difficult to determine the importance of the remaining genes that were expressed in low abundances.

Diatom taxa, however, were not found to respond equivalently to being upwelled; clear differences were noted between *Chaetoceros*, *Pseudo-nitzschia*, and other diatoms (Figure 1.4a). Expression of 2,807 orthologs was detected in the genera *Chaetoceros*, *Pseudo-nitzschia*, and all other diatom genera combined mostly consisting of *Thalassiosira*. Similar to what was observed for major taxonomic groups, there was large overrepresentation of distinct sets of genes, particularly in *Chaetoceros* spp., also potentially reflecting transcriptional investments in different processes at different times.

Eighty-five genes showed significantly opposite fold-changes in *Chaetoceros* and *Pseudo-nitzschia* when compared to other diatoms. This highlights that gene expression may not be as accurately assessed by combining genes at high level taxonomic groupings as done in previous metatranscriptomics studies (Alexander et al 2015, Bertrand et al 2015). The high expression of a gene at one time point or treatment by one group may be cancelled out by another group with opposing expression leading to the incorrect conclusion for the group as a whole. Additionally, one genus could be driving expression of many genes rather than being distributed across the entire group.

Molecular Characterization of the Nitrogen Assimilation Response

Gene expression was assessed among specific diatom genera and other phytoplankton groups to investigate nitrogen assimilation and utilization. Querying nitrogen-related genes for these groups and partitioning by k-means clustering revealed differences in gene expression for the diatoms compared to other phytoplankton (Figures 1.4b, A.10). The genes that clustered as highly expressed both pre- and post-simulated upwelling, referred to here as frontloaded, were almost all from diatoms and related to nitrogen assimilation: nitrate transporter, nitrate reductase, nitrite reductase, and ammonium transporter (Figure

1.4c). The only constitutively and highly expressed nitrogen-related gene by other groups was glutamine synthetase within dinoflagellates and haptophytes. Within diatoms, the nitrate assimilation genes all had a positive fold change suggesting slightly greater abundance of these genes post-upwelling when compared to pre-upwelling matching our observations of increased nitrate uptake at 72 and 120 hours (Figures 1.1c, 1.1f). The change in expression in nitrate reductase was very low which contrasts a simulated upwelling experiment with a *Skeletonema* species (Smith et al 1992). *Skeletonema*, however, was not found to be an abundant genus within this study, and this variation further highlights potential genera-specific differences in the upwelling response.

The urea cycle is believed to facilitate recovery from prolonged nitrogen limitation for diatoms (Allen et al 2011), but may also be important for the shift-up response. The urea cycle genes carbamoyl-phosphate synthetase and argininosuccinate synthase were also frontloaded by diatoms (Figures 1.4b, 1.4c, and A.10). Several others were significantly overrepresented post-upwelling including ornithine carbamoyltransferase, arginosuccinate lyase, and urease. The exception in diatoms was arginase, the final enzyme in the urea cycle which was significantly overrepresented pre-upwelling by diatoms not including *Chaetoceros* (Figure A.10). Low expression of arginase post-upwelling is similar to the diatom response to iron enrichment (Marchetti et al 2012) and may suggest that in both of these scenarios, there are alternative fates for urea cycle intermediates such as nitrogen storage or silica precipitation (Kröger et al 2001, Llácer et al 2008).

High relative expression of almost all of these nitrogen-related genes in diatoms compared to most of the other phytoplankton groups at both 0 and 72 hours suggests that investing in nitrogen assimilation and utilization is a priority even when conditions are not optimal for growth. Constitutively high expression of the primary genes for nitrate assimilation such as nitrate transporters and nitrate reductase pre-upwelling may contribute to the rapid response of diatoms as part of their shift-up process. By maintaining elevated pools of nitrogen-related gene transcripts or expressed proteins, upwelled cells are set up to rapidly assimilate available nitrogen whereas other phytoplankton groups appear to wait to upregulate these genes once upwelled into the euphotic zone. These results further support the hypothesis that one reason diatoms dominate upwelling regions is because they have the ability to take up and assimilate nitrate more quickly than other phytoplankton groups (Fawcett and Ward 2011).

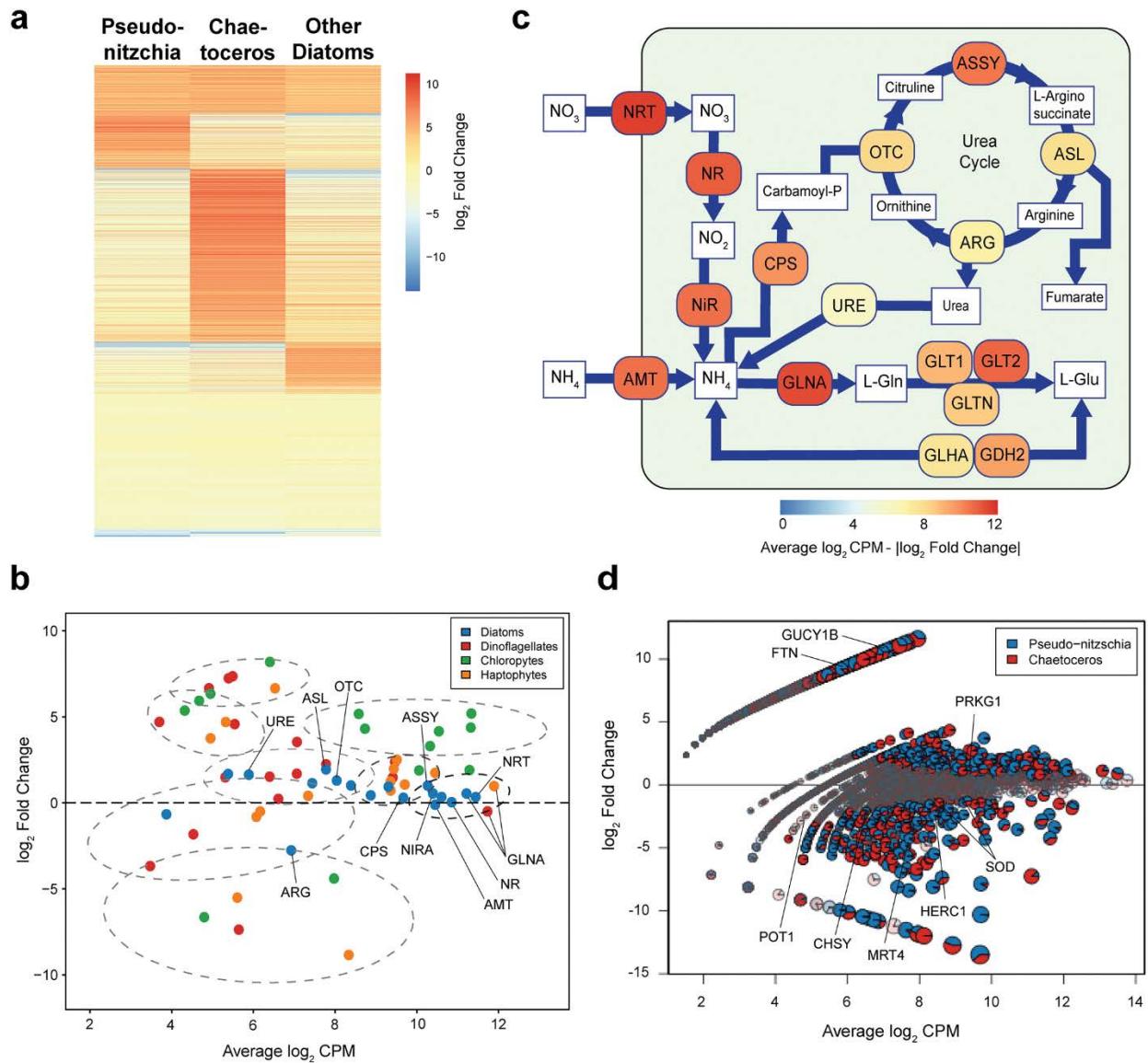


Figure 1.4 - Diatom gene expression. (A) Heatmap for the 2,807 commonly expressed KOs at 0 and 72 hours for *Chaetoceros*, *Pseudo-nitzchia*, and all other diatoms. Each row indicates an expressed KO with darker red (positive fold change) indicating overrepresentation at 72 hours and darker blue (negative fold change) indicating overrepresentation at 0 hours. (B) MA plot of nitrogen-related genes for the four main detected phytoplankton groups: diatoms (blue), dinoflagellates (red), chlorophytes (green), and haptophytes (orange). Genes are grouped into clusters using k-means clustering with confidence ellipses at the 90% level. Select genes are labelled within clusters that represent the most highly frontloaded genes (black ellipse) and additional frontloaded genes (medium gray ellipse). These are labelled as follows: NRT, nitrate transporter; GLNA, glutamine synthetase; NR, nitrate reductase; AMT, ammonium transporter; NIRA, nitrite reductase (ferredoxin); ASSY, arginosuccinate synthase; CPS, carbamoyl-phosphate synthetase; OTC, ornithine carbamoyltransferase; ASL, argininosuccinate lysase; URE, urease, ARG, arginase. (C) Cell schematic depicting frontloading of genes associated with nitrogen assimilation and utilization for diatoms. The model is based on Alexander et al. (2015a) and utilizes the same KO numbers. Color indicates the average abundance of the genes (\log_2 CPM) minus the absolute value of the \log_2 fold change to highlight the most abundant, lowest change (i.e., frontloaded) genes. Labels are as indicated in Figure 1.4b and as follows: GLT1, glutamate synthase (NADP/NADH); GLT2, glutamate synthase (ferredoxin); GLTD, glutamate synthase; small chain. (D) Differential transcript

abundance between 0 and 72 hours for *Chaetoceros* (red) and *Pseudo-nitzschia* (blue) for expressed KEGG Orthologs (KO). Each pie represents a KO and increases in size with absolute values of its coordinates to optimize visibility. Gene circles that are shaded and have grey borders are not significantly represented in either library (P -value ≥ 0.05). Select gene names discussed in the text are labeled as follows: CHSY, chondroitin sulfate synthase; FTN, ferritin; GUCY1B, guanylate cyclase soluble subunit beta; HERC1, ubiquitin-protein ligase; MRT4; mRNA turnover protein, POT1, protection of telomeres protein 1; PRKG1, cGMP-dependent protein kinase 1; SOD, superoxide dismutases.

This transcriptionally proactive approach to abiotic changes, termed ‘frontloading’, has been characterized with environmental stress response genes in coral and yeast (Barshis et al 2013, Berry and Gasch 2008). It is also similar to what has been observed in diatoms within a previous metatranscriptomics study in relation to iron stress. Iron-enrichment experiments in the northeastern Pacific Ocean demonstrated oceanic diatoms continued expressing genes encoding for iron-free photosynthetic proteins rather than substituting genes encoding for iron-containing functionally equivalent proteins which was in contrast to other phytoplankton groups (Cohen et al 2017, Marchetti et al 2012). This strategy is speculated to provide oceanic diatoms with the ability to rapidly acclimate to the inevitable return to iron-limited conditions just as our observations show a strategy that provides certain diatoms with the ability to rapidly take up nitrogen following upwelling. Constitutive frontloading is suggested to provide organisms with resilience to such stressors (Barshis et al 2013). Along similar lines, bloom-forming diatoms such as *Chaetoceros* and *Pseudo-nitzschia* may have evolved to frontload transcripts of particular genes depending on frequently encountered environmental fluctuations such as conditions associated with the upwelling conveyor belt cycle rather than simply reactively responding to these changes.

Chaetoceros and *Pseudo-nitzschia* Expressed Genes

Analyzing genes assigned to two of the most dominant diatom genera, *Chaetoceros* and *Pseudo-nitzschia*, provides further insight into the molecular mechanisms these genera use at depth and as part of their shift-up response. From KOs with module annotations, it is evident that the significantly overrepresented genes at both time points fall into a diverse set of functional categories even at a high-level grouping (Figure A.9). To obtain finer resolution, gene expression among all KOs for these genera was examined.

Many genes overrepresented in the pre-upwelling condition, such as those involved in proteolysis, stabilization of DNA or RNA, defense against reactive oxygen species (superoxide dismutases)(Fridovich

1998), and TEP production, suggest that diatoms are responding to stress (Figure 1.4d; Appendix A). *Pseudo-nitzschia* in particular expressed a set of distinctive genes as part of its shift-up response. Ferritin was highly expressed post-upwelling, possibly providing a method of storing the essential micronutrient iron. As iron availability in the California upwelling regime can be sporadic and potentially growth limiting, ferritin may provide an advantage to *Pseudo-nitzschia* by concentrating iron for longer-term storage (Bruland et al 2001, Marchetti et al 2009) although it may also be used for iron homeostasis (Pfaffen et al 2015).

At 72 hours, *Pseudo-nitzschia* also highly expressed a subunit of soluble guanylate cyclase (GUCY1B, sGC, Figure 1.4d). sGC is the only proven receptor of nitric oxide (Denninger and Marletta 1999) and synthesizes cyclic guanosine monophosphate (cGMP), a second messenger related to many physiological responses (Delledonne 2005). cGMP activates protein kinase G (PRKG1) which was also significantly expressed (Figure 1.4d). Although nitric oxide has been hypothesized to be an infochemical for intercellular signaling and monitoring of stress in diatoms (Amin et al 2012, Vardi 2008), *Pseudo-nitzschia* are generally not believed to have a nitric oxide synthase gene as a putative sequence was detected in only one species, *P. multistriata* (Di Dato et al 2015). *Pseudo-nitzschia* may be using sGC to monitor exhibition of stress from other genera which could allow them to rapidly adapt to changing conditions or respond to sexual cues (Basu et al 2017). Nitric oxide is also produced by activation of nitrate reductase (Sakihama et al 2002). As increased nitrate reductase activity occurs as part of the shift-up response, sGC may be used to monitor the continuation of that response and promote certain cellular functions such as gliding of pennates or binary fission (Thompson et al 2008). Inhibition of sGC prevents the germination of *Leptocyclindrus danicus* resting spores suggesting that it may be involved in transitioning from a resting stage in certain diatoms (Shikata et al 2011). Examination of this gene in our reference database reveals that it is highly conserved among *Pseudo-nitzschia* spp. but not ubiquitously present among diatoms (Figure A.11). The high expression of sGC as part of the upwelling response, evolutionary conservation of this gene, and potential to act as an important signaling device among *Pseudo-nitzschia* spp. may play a role in their success.

Conclusions

Our simulated upwelling experiment in the California Upwelling Zone is consistent with previous physiological observations of the shift-up response in upwelled phytoplankton: growth of large chain-forming diatoms and increased nitrate assimilation rates. The application of metatranscriptomics to the entire phytoplankton community highlights the divergent transcriptional response of major phytoplankton groups and diatom genera, potentially reflecting variations in their life history strategies. By frontloading, diatoms exhibit the potential for abundant nitrogen-related genes in their enzyme pool allowing them to respond to available nitrate more rapidly than other phytoplankton. This trait is not unlike the response of oceanic diatoms to iron enrichment and may indicate that diatoms have evolved to frontload transcripts in response to frequently encountered changes in their environment. Although the response to upwelling has largely been focused on nitrogen-related pathways, it is likely that other uncharacterized genes and pathways are also important to diatom success.

REFERENCES

- Alexander H, Rouco M, Haley ST, Wilson ST, Karl DM, Dyhrman ST (2015). Functional group-specific traits drive phytoplankton dynamics in the oligotrophic ocean. *P Natl Acad Sci USA* **112**: E5972-E5979.
- Allen AE, Dupont CL, Obornik M, Horak A, Nunes-Nesi A, McCrow JP et al (2011). Evolution and metabolic significance of the urea cycle in photosynthetic diatoms. *Nature* **473**: 203-207.
- Amin SA, Parker MS, Armbrust EV (2012). Interactions between Diatoms and Bacteria. *Microbiol Mol Biol R* **76**: 667-684.
- Barshis DJ, Ladner JT, Oliver TA, Seneca FO, Traylor-Knowles N, Palumbi SR (2013). Genomic basis for coral resilience to climate change. *P Natl Acad Sci USA* **110**: 1387-1392.
- Basu S, Patil S, Mapleson D, Russo MT, Vitale L, Fevola C et al (2017). Finding a partner in the ocean: molecular and evolutionary bases of the response to sexual cues in a planktonic diatom. *New Phytologist* **215**: 140-156.
- Berry DB, Gasch AP (2008). Stress-activated Genomic Expression Changes Serve a Preparative Role for Impending Stress in Yeast. *Mol Biol Cell* **19**: 4580-4587.
- Bertrand EM, McCrow JP, Moustafa A, Zheng H, McQuaid JB, Delmont TO et al (2015). Phytoplankton–bacterial interactions mediate micronutrient colimitation at the coastal Antarctic sea ice edge. *P Natl Acad Sci USA* **112**: 9938-9943.
- Birol I, Jackman SD, Nielsen CB, Qian JQ, Varhol R, Stazyk G et al (2009). De novo transcriptome assembly with AbySS. *Bioinformatics* **25**: 2872-2877.
- Bolger AM, Lohse M, Usadel B (2014). Trimmomatic: A flexible trimmer for Illumina Sequence Data. *Bioinformatics* **30**: 2114-2120.
- Bronk DA, Glibert PM, Ward BB (1994). Nitrogen Uptake, Dissolved Organic Nitrogen Release, and New Production. *Science* **265**: 1843-1846.
- Bruland KW, Rue EL, Smith GJ (2001). Iron and macronutrients in California coastal upwelling regimes: Implications for diatom blooms. *Limnol Oceanogr* **46**: 1661-1674.
- Brzezinski MA, Nelson DM (1995). The annual silica cycle in the Sargasso Sea near Bermuda. *Deep-Sea Res Pt I* **42**: 1215-1237.
- Capone DG, Hutchins DA (2013). Microbial biogeochemistry of coastal upwelling regimes in a changing ocean. *Nature Geosci* **6**: 711-717.
- Caron DA, Alexander H, Allen AE, Archibald JM, Armbrust EV, Bachy C et al (2017). Probing the evolution, ecology and physiology of marine protists using transcriptomics. *Nat Rev Micro* **15**: 6-20.
- Chen Y, Lun A, Smyth G (2014). Differential expression analysis of complex RNA-seq experiments using edgeR. In: Datta S, Nettleton DS (eds). *Statistical Analysis of Next Generation Sequence Data*. Springer: New York.
- Cohen NR, Ellis KA, Lampe RH, McNair HM, Twining BS, Brzezinski MA et al (2017). Variations in diatom transcriptional responses to changes in iron availability across ocean provinces. *Front Mar Sci* **4**: 360.
- Delledonne M (2005). NO news is good news for plants. *Curr Opin Plant Biol* **8**: 390-396.

- Denninger JW, Marletta MA (1999). Guanylate cyclase and the ·NO/cGMP signaling pathway. *BBA-Bioenergetics* **1411**: 334-350.
- Di Dato V, Musacchia F, Petrosino G, Patil S, Montresor M, Sanges R *et al* (2015). Transcriptome sequencing of three *Pseudo-nitzschia* species reveals comparable gene sets and the presence of Nitric Oxide Synthase genes in diatoms. *Sci Rep* **5**: 12329.
- Dugdale RC, Wilkerson FP (1986). The use of ¹⁵N to measure nitrogen uptake in eutrophic oceans; experimental considerations. *Limnol Oceanogr* **31**: 673-689.
- Dugdale RC, Wilkerson FP (1989). New production in the upwelling center at Point Conception, California: temporal and spatial patterns. *Deep-Sea Res* **36**: 985-1007.
- Estrada M, Blasco D (1985). Phytoplankton assemblages in coastal upwelling areas. In: Bas C, Margalef R, Rubies P (eds). *International Symposium on the Most Important Upwelling Areas off Western Africa*: Barcelona. Pp 379-402.
- Fawcett S, Ward B (2011). Phytoplankton succession and nitrogen utilization during the development of an upwelling bloom. *Mar Ecol-Prog Ser* **428**: 13-31.
- Fridovich I (1998). Oxygen toxicity: a radical explanation. *J Exp Biol* **201**: 1203-1209.
- Kanehisa M, Furumichi M, Tanabe M, Sato Y, Morishima K (2017). KEGG: new perspectives on genomes, pathways, diseases and drugs. *Nucleic Acids Res* **45**: D353-D361.
- Keeling PJ, Burki F, Wilcox HM, Allam B, Allen EE, Amaral-Zettler LA *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**: e1001889.
- Krause JW, Nelson DM, Lomas MW (2009). Biogeochemical responses to late-winter storms in the Sargasso Sea, II: Increased rates of biogenic silica production and export. *Deep-Sea Res Pt I* **56**: 861-874.
- Kröger N, Deutzmann R, Sumper M (2001). Silica-precipitating Peptides from Diatoms: The Chemical Structure of Silaffin-1A from *Cylindrotheca Fusiformis*. *J Biol Chem* **276**: 26066-26070.
- Kudela RM, Dugdale RC (2000). Nutrient regulation of phytoplankton productivity in Monterey Bay, California. *Deep-Sea Res Pt II* **47**: 1023-1053.
- Lachkar Z, Gruber N (2013). Response of biological production and air-sea CO₂ fluxes to upwelling intensification in the California and Canary Current Systems. *J Marine Syst* **109-110**: 149-160.
- Langmead B, Salzberg SL (2012). Fast gapped-read alignment with Bowtie 2. *Nat Meth* **9**: 357-359.
- Lassiter AM, Wilkerson FP, Dugdale RC, Hogue VE (2006). Phytoplankton assemblages in the CoOP-WEST coastal upwelling area. *Deep-Sea Res Pt II* **53**: 3063-3077.
- Lelong A, Hégaret H, Soudant P, Bates SS (2012). *Pseudo-nitzschia* (Bacillariophyceae) species, domoic acid and amnesic shellfish poisoning: revisiting previous paradigms. *Phycologia* **51**: 168-216.
- Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N *et al* (2009). The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**: 2078-2079.
- Llácer JL, Fita I, Rubio V (2008). Arginine and nitrogen storage. *Curr Opin Struct Biol* **18**: 673-681.

- MacIsaac JJ, Dugdale RC, Barber RT, Blasco D, Packard TT (1985). Primary production cycle in an upwelling center. *Deep-Sea Res* **32**: 503-529.
- Marchetti A, Parker MS, Moccia LP, Lin EO, Arrieta AL, Ribalet F *et al* (2009). Ferritin is used for iron storage in bloom-forming marine pennate diatoms. *Nature* **457**: 467-470.
- Marchetti A, Schruth DM, Durkin CA, Parker MS, Kodner RB, Berthiaume CT *et al* (2012). Comparative metatranscriptomics identifies molecular bases for the physiological responses of phytoplankton to varying iron availability. *P Natl Acad Sci USA* **109**: E317-E325.
- McManus MA, Kudela RM, Silver MW, Steward GF, Donaghay PL, Sullivan JM (2008). Cryptic Blooms: Are Thin Layers the Missing Connection? *Estuar Coast* **31**: 396-401.
- Parsons TR, Maita Y, Lalli CM (1984). *A Manual of Chemical & Biological Methods for Seawater Analysis*. Pergamon: Amsterdam.
- Pfaffen S, Bradley JM, Abdulqadir R, Firme MR, Moore GR, Le Brun NE *et al* (2015). A Diatom Ferritin Optimized for Iron Oxidation but not Iron Storage. *J Biol Chem* **290**: 28416-28427.
- Pitcher GC (1990). Phytoplankton seed populations of the Cape Peninsula upwelling plume, with particular reference to resting spores of Chaetoceros (bacillariophyceae) and their role in seeding upwelling waters. *Estuar Coast Shelf S* **31**: 283-301.
- Redfield AC, Ketchum BH, Richards FA (1963). The influence of organisms on the composition of sea water. In: Hill M (ed). *The Sea*. Interscience: New York. Pp 26-77.
- Rines JEB, Donaghay PL, Dekshenieks MM, Sullivan JM, Twardowski MS (2002). Thin layers and camouflage: hidden Pseudo-nitzschia spp. (Bacillariophyceae) populations in a fjord in the San Juan Islands, Washington, USA. *Mar Ecol-Prog Ser* **225**: 123-137.
- Robertson G, Schein J, Chiu R, Corbett R, Field M, Jackman SD *et al* (2010). De novo assembly and analysis of RNA-seq data. *Nat Meth* **7**: 909-912.
- Robinson MD, Smyth GK (2008). Small-sample estimation of negative binomial dispersion, with applications to SAGE data. *Biostatistics* **9**: 321-332.
- Robinson MD, McCarthy DJ, Smyth GK (2010). EdgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26**: 139-140.
- Ryan JP, Chavez FP, Bellingham JG (2005). Physical-biological coupling in Monterey Bay, California: topographic influences on phytoplankton ecology. *Mar Ecol-Prog Ser* **287**: 23-32.
- Ryther JH (1969). Photosynthesis and Fish Production in the Sea. *Science* **166**: 72-76.
- Sakihama Y, Nakamura S, Yamasaki H (2002). Nitric Oxide Production Mediated by Nitrate Reductase in the Green Alga Chlamydomonas reinhardtii: an Alternative NO Production Pathway in Photosynthetic Organisms. *Plant Cell Physiol* **43**: 290-297.
- Schnetzer A, Jones BH, Schaffner RA, Cetinic I, Fitzpatrick E, Miller PE *et al* (2013). Coastal upwelling linked to toxic Pseudo-nitzschia australis blooms in Los Angeles coastal waters, 2005–2007. *J Plankton Res* **35**: 1080-1092.
- Seegers BN, Birch JM, Marin R, Scholin CA, Caron DA, Seubert EL *et al* (2015). Subsurface seeding of surface harmful algal blooms observed through the integration of autonomous gliders, moored environmental sample processors, and satellite remote sensing in southern California. *Limnol Oceanogr* **60**: 754-764.

Shikata T, Iseki M, Matsunaga S, Higashi S-I, Kamei Y, Watanabe M (2011). Blue and Red Light-Induced Germination of Resting Spores in the Red-Tide Diatom *Leptocylindrus danicus*. *Photochem Photobiol* **87**: 590-597.

Smith GJ, Zimmerman RC, Alberte RS (1992). Molecular and physiological responses of diatoms to variable levels of irradiance and nitrogen availability: growth of *Skeletonema costatum* in simulated upwelling conditions. *Limnol Oceanogr* **37**: 989-1007.

Thompson SEM, Taylor AR, Brownlee C, Callow ME, Callow JA (2008). The Role of Nitric Oxide in Diatom Adhesion in Relation to Substratum Properties. *J Phycol* **44**: 967-976.

Vardi A (2008). Cell signaling in marine diatoms. *Commun Integr Biol* **1**: 134-136.

Wilkerson FP, Dugdale RC (1987). The Use of Large Shipboard Barrels and Drifters to Study the Effects of Coastal Upwelling on Phytoplankton Dynamics. *Limnol Oceanogr* **32**: 368-382.

Wilkerson FP, Dugdale RC, Kudela RM, Chavez FP (2000). Biomass and productivity in Monterey Bay, California: contribution of the large phytoplankton. *Deep-Sea Res Pt II* **47**: 1003-1022.

Wilkerson FP, Lassiter AM, Dugdale RC, Marchi A, Hogue VE (2006). The phytoplankton bloom response to wind events and upwelled nutrients during the CoOP WEST study. *Deep-Sea Res Pt II* **53**: 3023-3048.

Wilkerson FP, Dugdale RC (2008). *Nitrogen in the Marine Environment (2nd Edition)*. Academic Press: Burlington, MA, USA

CHAPTER 2: DIFFERENT IRON STORAGE STRATEGIES AMONG BLOOM-FORMING DIATOMS

Introduction

In wide-ranging areas of the world's oceans, phytoplankton growth is frequently limited by the availability of the micronutrient iron (Martin and Fitzwater 1988, Moore et al 2001). Inputs of iron via atmospheric dust deposition or resuspension of continental shelf sediment during upwelling are spatially and temporally variable, leading to gradients or sporadic episodes of increased iron concentrations (Bruland et al 2001, Mahowald et al 2005). From the tropics to the poles, the introduction of iron to iron-limited surface waters is known to stimulate phytoplankton growth usually dominated by large diatoms (Boyd et al 2007). One cosmopolitan genus of pennate diatoms in particular, *Pseudo-nitzschia*, consistently thrives when iron is added (de Baar et al 2005). Consequently, certain bloom-forming diatoms have evolved to possess a number of mechanisms that allow them to persist under chronically low iron and rapidly divide when it is reintroduced (Armbrust 2009, Marchetti and Maldonado 2016). One such mechanism is the ability to store iron; however, methods for storage are not universal among all diatoms, which in turn may affect iron storage capacities. Two iron storage strategies in diatoms have been previously described: within the protein ferritin and vacuoles (Marchetti and Maldonado 2016).

Diatom ferritins are unlike those from other eukaryotes as the gene was acquired via lateral gene transfer from cyanobacteria, some of which use ferritin for long-term storage (Groussman et al 2015, Keren et al 2004). Although ferritin is present across all four diatom lineages, it appears to be mostly absent in many centric diatoms such as several *Thalassiosira* species (Cohen et al 2018, Groussman et al 2015). Furthermore, evidence suggests that diatom ferritins do not all serve the same functional role. In addition to long-term storage, ferritin may play a role in iron homeostasis. Free intracellular iron is toxic to cells, and some microalgae express ferritin to buffer iron released from degrading proteins (Botebol et al 2015, Liochev and Fridovich 1999, Long et al 2008). In the green algae, *Ostreococcus*, ferritin that serves this buffering role was regulated by the circadian clock rather than iron availability (Botebol et al 2015).

Similarly in some *FTN*-containing diatoms, gene expression changes are comparatively subdued between different iron states suggesting a role other than long-term storage (Cohen et al 2018).

In contrast, *Pseudo-nitzschia* ferritins are highly conserved, and when acclimated to varied iron concentrations, *Pseudo-nitzschia* substantially increase *FTN* expression with dissolved iron concentrations (Marchetti et al 2009, Marchetti et al 2017). Additionally, the oceanic species *Pseudo-nitzschia granii* was also able to perform significantly more cell divisions than a diatom without ferritin, *Thalassiosira oceanica*, when transitioned from high to low iron conditions (Marchetti et al 2009). As a result, ferritin in *Pseudo-nitzschia* provides an explanation for their unusually high iron storage capacities (Cohen et al 2018, Marchetti et al 2009); however, biochemical examination of *Pseudo-nitzschia* ferritin shows that iron remineralization is slow, suggesting a role in buffering (Pfaffen et al 2015). Beyond speculation over its true function, the ecological importance of ferritin also remains in question; it has been suggested that competitive strength for diatoms in iron-limited regions lies primarily in using iron-free proteins rather than iron storage (Mock and Medlin 2012).

In the non-ferritin containing species of the genus *Thalassiosira*, evidence supports an intracellular vacuole mechanism for iron storage. *Thalassiosira pseudonana* possesses a divalent transporter belonging to the natural resistance associated macrophage protein (*NRAMP*) family (Kustka et al 2007), and some *NRAMP* proteins are used in other eukaryotes to transport iron out of vacuoles (Lanquar et al 2005, Portnoy et al 2000). Observed up-regulation of this gene under low dissolved iron concentrations suggests that *NRAMP* could be used to mobilize iron out of a vacuole although a role in cell-surface uptake has not been ruled out (Kustka et al 2007). Anomalously high intracellular regions of iron with stoichiometries consistent with polyphosphate-bound iron in *T. pseudonana* and *Thalassiosira weissflogii* also support a vacuolar storage mechanism for these diatoms (Nuester et al 2012).

Advances in analytical capabilities have improved the ability to study the relationship between iron and phytoplankton in natural communities. Reference genomic data for marine organisms allows for annotation of environmental RNA, or metatranscriptomes, including accurate taxonomic annotation for well-represented genera (Caron et al 2017). Synchrotron X-ray fluorescence (SXRF) enables quantitative elemental analysis, including that of iron, for individual cells (Twining et al 2003). Here we combine metatranscriptomics and SXRF to examine iron storage dynamics in natural eukaryotic phytoplankton

communities. Using shipboard incubations, iron was added to or removed from phytoplankton with varying initial iron states. We show that iron storage ability can be exceptionally high, and ferritin expression is unique in *Pseudo-nitzschia* while other non-ferritin utilizing diatoms likely employ vacuolar storage. This ferritin-linked storage ability may provide an advantage to ferritin-utilizing diatoms such as *Pseudo-nitzschia* under prolonged periods of iron limitation.

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* (Figure 2.1 and Table 2.1). Experimental procedures are described in (Cohen et al 2017b) for all sites except C-Low2, which is described in Chapter 1. In summary, near-surface water (Table B.1) was collected using trace-metal clean techniques and distributed into acid-cleaned 10L 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 FeCl₃ (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 NO₃ was added to all incubations to support growth since initial NO₃ concentrations were low (1.49 µmol L⁻¹). Timing of incubation sampling for each experiment is described in Table B.1. 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 (2017b). Additional methods are described in Appendix B.

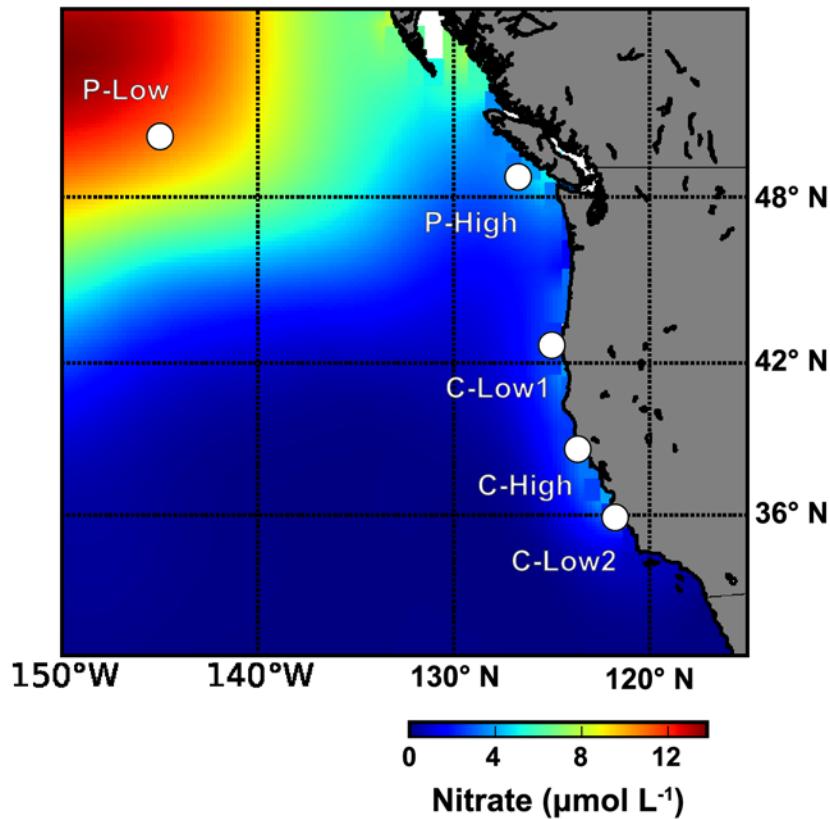


Figure 2.1 - Map of field incubation locations with interpolated statistical means of nitrate concentrations ($\mu\text{mol L}^{-1}$) on a 1° grid of all decades from World Ocean Atlas 2013.

Name	Latitude & Longitude	Fe:NO ₃ (nM:μM)	Initial [Fe] (nM)
C-High	38.7° N, 123.7° W	0.26	3.57
C-Low1	42.7° N, 125.0° W	0.06	1.05
C-Low2	35.9° N, 121.7° W	0.06	1.28
P-High	48.7° N, 126.7° W	0.42	0.64
P-Low	50.0° N, 145.0° W	0.01	0.05

Table 2.1. Incubation site identifiers, location, and initial iron conditions. At site P-High, 10 $\mu\text{mol L}^{-1}$ of NO₃ was added.

RNA-Seq Overview

The initial RNA extraction, sequencing, assemblies, and annotations are described in Cohen et al (2017b) and Chapter 1. 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 (Bolger et al 2014). 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 (Birol et al 2009) then the assemblies were merged using Trans-ABySS v1.5.3 to remove redundant contigs (Robertson et al 2010). 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 prokaryotes including all sequenced transcriptomes from the Marine Microbial Eukaryote Transcriptome Sequencing Project (Keeling et al 2014). To assign gene function to contigs, the same methodology with the Kyoto Encyclopedia of Genes and Genomes (KEGG; Release 75) was used (Kanehisa et al 2017). 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 (Gremme et al 2013). Read counts were estimated from this combined assembly using the quasi-mapping method implemented in Salmon v0.73 (Patro et al 2017). 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 (2017). In all other instances, normalization, gene expression, and differential expression were assessed within each taxonomic group using DESeq2 v1.12.4 (Love et al 2014). Significance was determined by genes with Benjamini & Hochberg adjusted P-values ≤ 0.05 (Benjamini and Hochberg 1995). 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 (2015) with verification from the KEGG SSDB database (Kanehisa et al 2017).

Synchrotron X-ray Fluorescence (SXRF) Preparation and Analysis

Samples were collected and analyzed following Twining et al (2003, 2011). Briefly, cells from 500 mL of sample were gently preconcentrated approximately ten-fold over 2 µm 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 was 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 (Nunez-Milland et al 2010, Twining et al 2003), and cellular C was calculated from cell volume (Twining et al 2004).

Data Deposition

The data reported in this paper have been deposited in the National Center for Biotechnology (NCBI) sequence read archive under the accession nos. SRP074302 (BioProject no. PRJNA320398) and SRP108216 (BioProject no. PRJNA388329).

Results and Discussion

Iron States within the Initial Phytoplankton Communities and Incubations

The iron addition and removal experiments were conducted at four coastal sites and one oceanic site (Figure 2.1 and Table 2.1). Three of the coastal sites are located in the California Upwelling Zone (CUZ): C-High, C-Low1, and C-Low2. In the CUZ, iron delivery is primarily dependent on upwelling-driven resuspension of continental shelf sediments creating a mosaic of iron-limited regions that is largely

dependent on shelf width (Figure B.1)(Bruland et al 2001, Hutchins et al 1998). The other experiments were conducted along the Line-P transect, a well-characterized iron gradient extending into the high-nutrient low-chlorophyll (HNLC) region of the Northeast Pacific Ocean (Harrison 2002). Specifically, the sites correspond to the coastal station, P4 (P-High), and an oceanic site, Ocean Station Papa or P26 (P-Low).

Herein named based on their initial iron states (high or low), these locations provided varying initial phytoplankton biomass, macronutrient, and iron concentrations that were then further manipulated by the addition of iron (Fe) or the removal of bioavailable iron through addition of the strong iron chelator desferrioxamine B (DFB)(Figure 2.2 and Table 2.1). Short-lived pulse additions were also simulated by an initial addition of iron, then followed by addition of DFB at the first time point (FeDFB treatments). Iron status was assessed based on oceanographic context (Appendix B), differences in the chemical and biological properties of the water¹ (Table B.2), and a combination of gene expression-based molecular indicators that evaluate iron stress or limitation for distinct diatom genera (Figure B.2).

Low iron sites displayed comparatively lower Fe:NO₃ ratios (Table B.1). Further substantiating their initial status, significant differences between the Fe and control incubations were always observed in photosynthetic efficiency (F_v:F_m) and commonly found for chlorophyll, biogenic silica, nitrate drawdown, and the molecular indicators of iron stress (Table B.3). C-High phytoplankton dynamics were largely unaffected by the addition of iron compared to controls as shown by a lack of significant differences, and those at P-High were driven more by macronutrient availability rather than iron.

At all sites, significant differences were observed between the Fe and DFB treatments, indicating that low and high iron scenarios were created at each site, regardless of the initial iron status. At C-High, only nitrate was significantly reduced by iron addition out of the macronutrient, chlorophyll, and biogenic silica concentrations; however, the molecular indicators validate that the diatom community's physiology was influenced by the addition of DFB, creating a low iron contrast. The remaining sites show significant differences between the Fe and DFB treatments in nitrate drawdown, chlorophyll, and biogenic silica, although these differences are also not entirely universal, which may be a result of the dissimilar initial

¹nitrate drawdown, chlorophyll a, biogenic silica, and photosynthetic efficiency (F_v:F_m)

conditions. $F_v:F_m$, however, was consistently reduced by DFB addition as were the molecular indicators of iron stress in diatoms at all sites.

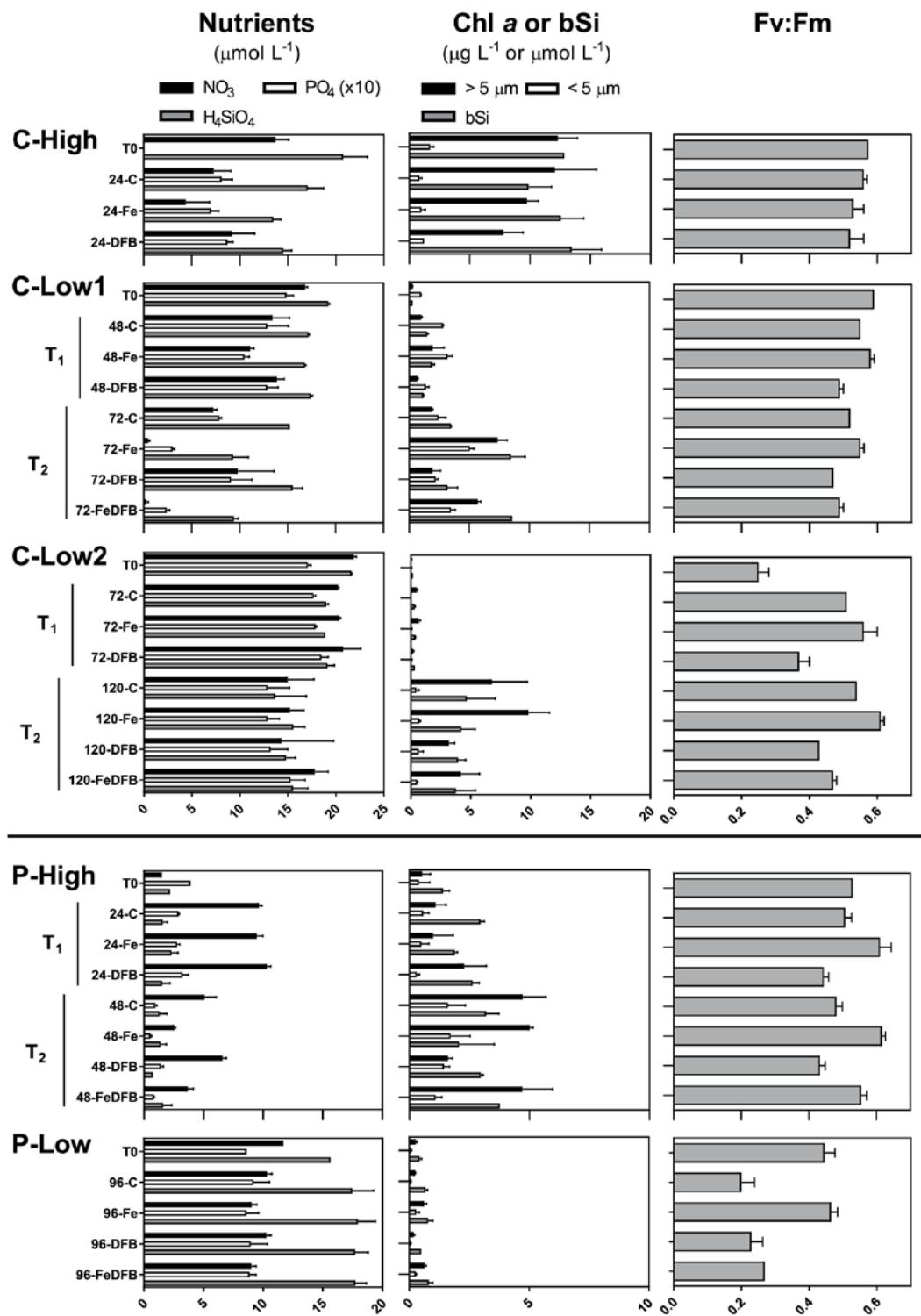


Figure 2.2 - Macronutrient ($\mu\text{mol L}^{-1}$), size-fractionated chlorophyll a ($\mu\text{g L}^{-1}$), and biogenic silica ($\mu\text{mol L}^{-1}$) concentrations and maximum photochemical yields of photosystem II ($F_v:F_m$) from the initial collected seawater ($t = 0$) and incubations at various time points. Incubations are labelled as follows: control (C), iron addition (Fe), iron removal (DFB), iron addition followed by removal (FeDFB) and denoted as the first or second time point (T_1 or T_2) where applicable. California Upwelling Zone (CUZ) sites and Line-P sites are grouped separately as two of the x-axis scales are different.

Taxonomic Distributions

Relative taxonomic distributions based on transcript proportions were assessed within the whole community for phylum-based groupings and among diatom genera (Figure 2.3 and Figure B.3). Although transcript proportions may not always relate to cell abundance, they are indicative of the relative activity among groups and within this study, are consistent with inferred cell abundances. Diatoms consistently comprised higher relative abundances in the Fe treatments compared to the DFB treatments at the low iron sites while the opposite was found by the final time points at both high iron sites (Figure 2.3a). Although these higher relative transcript proportions at the low iron sites appear minor (<10%), they may still have translated to a large absolute abundance of certain diatoms. These higher relative abundances were coupled to increases in chlorophyll (>5 μm) and biogenic silica concentrations indicating that the absolute cellular abundance of diatoms was greater following iron addition rather than simply a shift in proportions (Figure 2.2).

This increase in diatom transcript abundance and presumed cellular abundance when iron was added aligns with diatoms' known ability to respond positively when transitioning from low iron to high iron (Boyd et al 2007). Diatoms were also favored when the environment transitioned from high to low iron as shown by higher relative abundance in the DFB treatments compared to the Fe treatments by the last time points at the high iron sites (Figure 2.3a). Regardless of the situation (high iron to low iron or vice versa), diatoms were ultimately able to increase in relative abundance aligning with their responsiveness to iron availability (Cohen et al 2017b, Marchetti et al 2012). In the transition from high to low iron, iron storage may play a role in allowing diatoms to continue to divide as growth rates in other taxa slow.

The dominant diatom genera within these experiments were *Chaetoceros*, *Pseudo-nitzschia*, and *Thalassiosira*, which are also among the most common in the open ocean (Figure B.3)(Malviya et al 2016). When iron was added, *Thalassiosira* generally comprised higher transcript proportions of the diatom community, while inconsistent differences for *Chaetoceros* were observed (Figure 2.3b).

Chaetoceros appeared to be favored by iron addition in the coastal low iron sites while having slightly higher proportions with DFB in the high iron sites. Interestingly, *Pseudo-nitzschia* consistently had higher transcript proportions in the control and DFB treatments compared to when iron was added. As anticipated from *Pseudo-nitzschia* species' high iron storage capacities, their ability to store iron may play a role in these shifts. As iron is removed, *Pseudo-nitzschia* can use stored iron to maintain a higher growth rate while others may not (Marchetti et al 2009).

The exception was the oceanic site, P-Low, where *Pseudo-nitzschia* transcripts greatly increased in the Fe treatment compared to the control and DFB treatments, consistent with previous observations of iron enrichment in HNLC regions (Figure B.3)(de Baar et al 2005). Here, *Chaetoceros* proportions remained relatively low, but similar to the other sites, *Thalassiosira* rose in percentage when iron was added. Also as observed at the other sites in the DFB treatments, *Pseudo-nitzschia* still rose to become the most prevalent genus showing that it persisted compared to the other genera although iron was depleted.

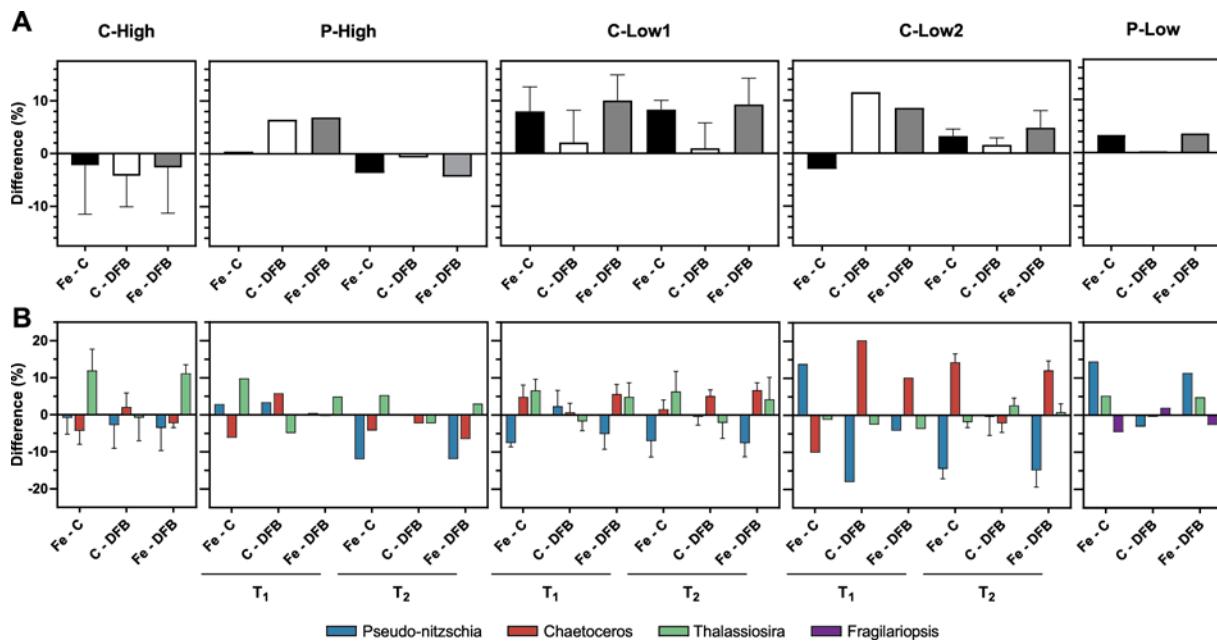


Figure 2.3 - Differences in relative proportions of taxonomically annotated transcripts between the iron addition and control (Fe-C), control and iron removal (C-DFB), and iron addition and iron removal (Fe-DFB) at each site for (a) diatom reads within the total library (Fe-C, black; C-DFB, white; Fe-DFB, gray) and (b) *Pseudo-nitzschia* (blue), *Chaetoceros* (red), and *Thalassiosira* (green) within the diatom reads. *Chaetoceros* were a small percentage of the diatom population at P-Low and *Fragilariopsis* (purple) are shown instead. Plots are grouped by initial iron state (high or low).

Iron Quotas among Diatom Genera

Cellular iron contents, or quotas, were quantified in individual *Chaetoceros* and *Pseudo-nitzschia* cells from within the incubations using synchrotron X-ray fluorescence (SXRF) microscopy (Figure 2.4a and Figure B.4)(Twining et al 2003). Both coastal and oceanic *Pseudo-nitzschia* species are known to have exceptionally high iron storage capacities, (Cohen et al 2018, Marchetti and Maldonado 2016) which was observed in these experiments (Figure 2.4a). At C-High, where *in situ* iron concentrations were highest (Table 2.1), *Pseudo-nitzschia* had significantly higher quotas in the high iron treatments (control and Fe) compared to *Chaetoceros* and at the other sites.

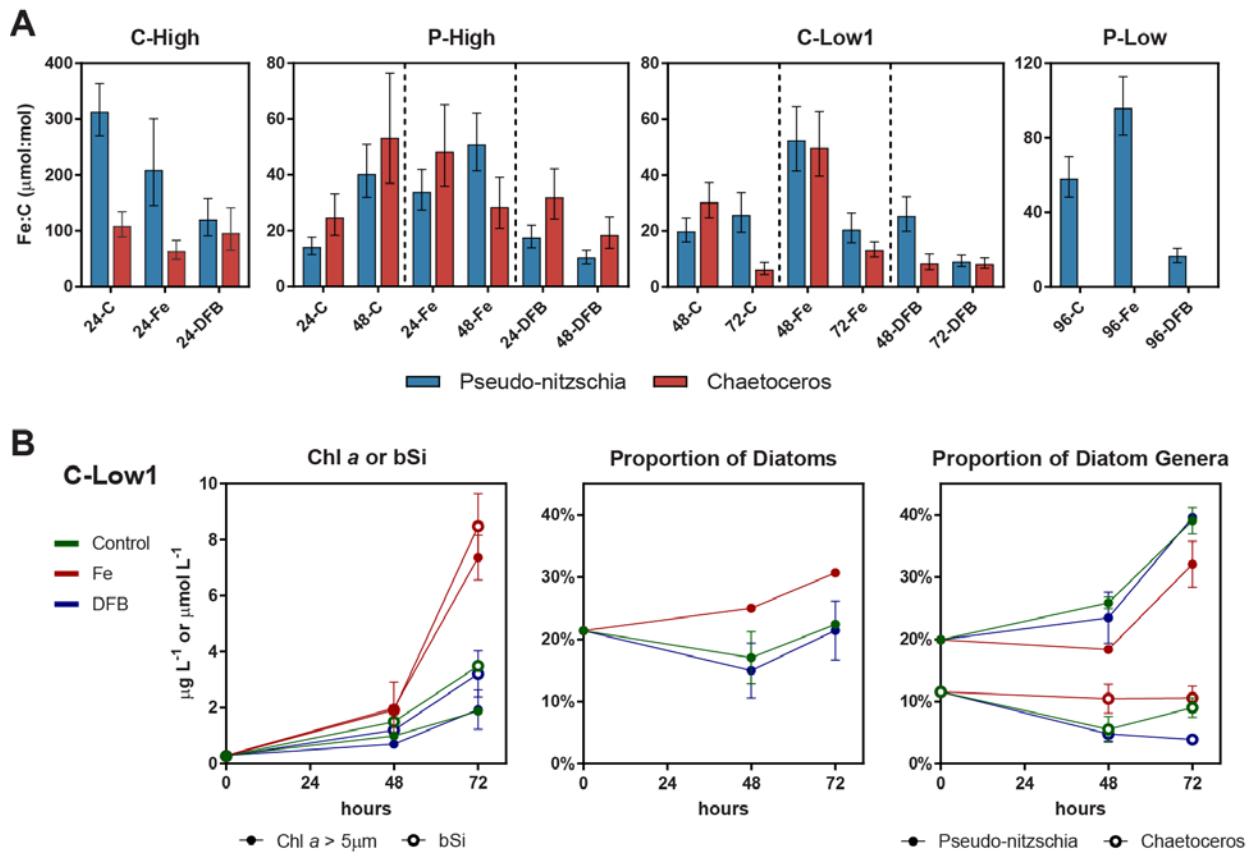


Figure 2.4 - (a) Cellular iron quotas ($\mu\text{mol Fe:mol C}$) in *Chaetoceros* and *Pseudo-nitzschia* from incubations. Bars are geometric means \pm standard error. **(b) Time course for site C-Low1** of chlorophyll a (> 5 μm size-fraction only), biogenic silica, diatom relative transcript proportions, and relative transcript proportions of *Pseudo-nitzschia* and *Chaetoceros* within diatoms. The treatments shown are control (green), iron addition (Fe; red), iron removal (DFB; blue). Plots are grouped by initial iron state (high or low).

In many cases, the quotas in *Pseudo-nitzschia* and *Chaetoceros* were similar. For instance, there are no significant differences in the quotas achieved between the genera at P-High and both increased their quotas when Fe was added at C-Low1 (Figure 2.4a). At P-High, although phytoplankton growth was induced by the addition of nitrate (primarily haptophytes), diatoms declined in relative transcript proportions from 24 to 48 hours as silicate concentrations decreased to $< 1 \mu\text{mol L}^{-1}$, which likely became limiting to diatom growth (Figure 2.2 and Figure B.3). With silicate depletion preventing cell division, cells likely accumulated iron in the Fe and control treatments (30) and transitioned toward their minimum quotas in the DFB treatment, where iron was bound to DFB (Figure 2.4a).

At P-Low, *Pseudo-nitzschia* showed extraordinary ability to maintain a substantial Fe:C quota of 58.02 (+11.95, -9.90) $\mu\text{mol:mol}$ in the control treatment even though the initial dissolved iron concentrations were 0.05 nmol L^{-1} and markers of iron stress were high (Figure B.2). The abundance of *Chaetoceros* cells was too low to quantify iron quotas at this site. Overall, these results align with previous studies showing overlap in iron quotas for diatoms, including *Thalassiosira* under moderate iron conditions, although *Pseudo-nitzschia* species appear to have higher iron maximum quotas and therefore a greater iron storage capacity (Cohen et al 2018, Marchetti and Maldonado 2016).

A closer examination of the results at C-Low1 suggests that maintenance of stored iron between the two diatom genera is not the same. At this site, *Pseudo-nitzschia* appeared to be able to use iron storage to continue to divide and outcompete *Chaetoceros*. As iron concentrations decreased in the control incubations and indicators of iron stress were exhibited (Figure B.2), *Pseudo-nitzschia* were able to maintain iron quotas ($25.7 +8.06, -6.13 \mu\text{mol Fe: mol C}$) while those in *Chaetoceros* declined ($6.21 +2.67, -1.86 \mu\text{mol Fe: mol C}$) from 48 hours to 72 hours. Meanwhile, chlorophyll, biogenic silica, diatom transcript proportions, and transcript proportions of *Chaetoceros* and *Pseudo-nitzschia* all increased between the same two time points (Figure 2.4b). Notably, *Pseudo-nitzschia* increased from 26% to 39% of diatom transcripts while *Chaetoceros* remained lower.

In the DFB treatments, *Chaetoceros* quotas at both time points were similar to the control, suggesting that *Chaetoceros* cells reached their minimum quota quickly as iron was depleted. Although the quotas at 72 hours in the DFB treatment were similar in both *Chaetoceros* and *Pseudo-nitzschia*, the reduction in quota for *Pseudo-nitzschia* can be attributed to their proportions approximately doubling over

the time frame of the incubations to 40% of diatom transcripts while *Chaetoceros* proportions steadily declined to 4% (Figure 2.4b). Meanwhile, diatom transcript proportions increased over time from 15% to 21% as did chlorophyll in the large size fraction from 0.70 to 1.93 µg L⁻¹. *Pseudo-nitzschia* were likely able to use their stored iron at 48 hours to continue to divide and become a larger proportion matching the control treatment by 72 hours despite iron removal by addition of DFB. Taken together, iron storage at this low iron site likely allowed *Pseudo-nitzschia* to maintain a higher growth rate while *Chaetoceros* growth slowed, consistent with what has been previously observed in laboratory experiments between *Pseudo-nitzschia granii* and *Thalassiosira oceanica* (Marchetti et al 2009).

Ferritin Expression in Pseudo-nitzschia and Other Diatoms

Ferritin expression in *Pseudo-nitzschia* was detected at all sites and at comparatively high levels in several incubations, being undetected only in the extreme low iron scenarios at P-Low and the initial community at C-Low2 that was severely light-limited and relatively low in diatom abundance (Figure 2.5a and Figure B.5). Phylogenetic analysis of environmental sequences indicates that these *FTN* sequences belong to both diatom ferritin groups (Groussman et al 2015), and are derived from multiple *Pseudo-nitzschia* species (Figure B.6 and Figure B.7). A previous iron and B-vitamin enrichment experiment at Ocean Station Papa (P-Low) also detected *FTN* expression in *Pseudo-nitzschia* (Cohen et al 2017a); thus, the abundance of *FTN* transcripts across these sites and experiments indicates that ferritin utilization is a widespread strategy employed by *Pseudo-nitzschia* spp.

Beyond detection, *Pseudo-nitzschia* *FTN* expression appears to be regulated by their initial iron status. At high iron sites, comparatively less variation in *FTN* expression was observed even though iron was either added or removed (Fe vs DFB, Figure 2.5a). Where replication was available (C-High), these differences were not statistically significant. Conversely, the low iron sites showed greater variation in *FTN* expression, with the most extreme case being the oceanic site, P-Low. Here, *Pseudo-nitzschia* *FTN* was undetected in the ambient, severely iron-limited community and then became highly expressed following iron addition (Figure 2.5a). Certainly, if the primary role of ferritin is for long-term iron storage, it would favor *Pseudo-nitzschia* to minimize synthesis of ferritin unless iron is available. *Pseudo-nitzschia* strains within environments where iron is low may be adapted to vary their expression of *FTN* to

accommodate these ephemeral inputs of iron, whereas *Pseudo-nitzschia* in high iron environments appear to constitutively express *FTN* to accommodate the more frequent iron inputs and higher supply.

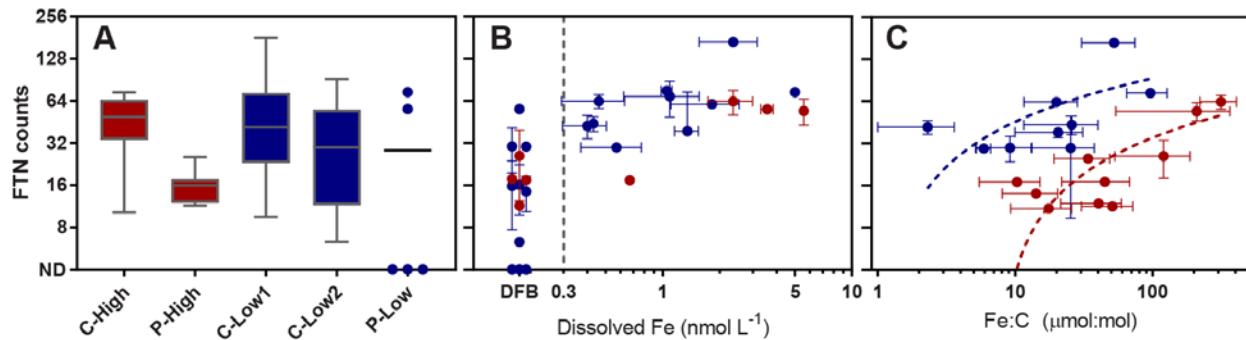


Figure 2.5 - *Pseudo-nitzschia* normalized ferritin (*FTN*) expression for all initial conditions and incubations. ND = no transcripts detected. Points are colored according to high (red) or low (blue) iron sites. (a) Box plots of *FTN* counts at each site. (b) Averaged *FTN* counts as a function of dissolved iron concentrations. In the DFB treatments or where iron was less than 0.05 nmol L^{-1} , *FTN* counts are plotted to the left of the gray vertical dashed line. (c) *FTN* counts as a function of cellular iron content (quota) expressed as Fe:C ($\mu\text{mol:mol}$) and fitted with a semilog line.

In general, *FTN* expression in *Pseudo-nitzschia* declined under low iron concentrations and reached a maximal level of expression at dissolved Fe concentrations over 1 nmol L^{-1} (Figure 2.5b). This expression pattern matches previous laboratory experiments that used iron chelators to control iron availability (Marchetti et al 2009, Marchetti et al 2017) and is unlike expression in *Ostreococcus* where ferritin serves a role in recycling intracellular iron (Botebol et al 2015). Although not tested to higher concentrations here, these data support that *Pseudo-nitzschia* utilize *FTN* for long-term storage, as *FTN* is likely expressed to maintain storage needed for their maximum quotas. Further substantiating this role, there is a significant positive correlation (Pearson) between *FTN* expression and iron quotas (Figure 2.5c) as well as the changes in *FTN* expression and iron quotas (Figure B.8). This concordance in *Pseudo-nitzschia* indicates that *FTN* expression and iron quotas vary synchronously, and that *Pseudo-nitzschia* are likely able to store more iron as *FTN* abundance increases.

Comparisons between the Fe and DFB treatments at each site display these changes in expression in greater detail (Figure 2.6). *FTN* expression was always higher when iron was added and lower when removed except at the second time point of P-High, which was likely influenced by macronutrient depletion (Cohen et al 2018). *FTN* was highly expressed in the low iron sites following iron

addition, consistent with *Pseudo-nitzschia* increasing *FTN* expression in response to increased iron concentrations (Marchetti et al 2009, Marchetti et al 2017). Responses varied when comparing the first to second time points in the Fe treatments, from a significant decrease in transcript abundance at C-Low1 to little change at C-Low2 (Figure 2.6). The strong difference at C-Low1 is largely driven by unusually high *FTN* expression at the first time point, being much higher than any other level of *FTN* expression even at high dissolved iron concentrations (Figure 2.5b). This result may be an artifact of earlier timing of sampling. As noted in a previous laboratory study, maximal expression of *FTN* quickly followed iron resupply to iron-limited *Pseudo-nitzschia multiseries* and then returned to steady-state levels (Marchetti et al 2009). This difference over time could also be affected by macronutrient depletion at the second time point (Figure 2.2; Cohen et al 2018).

As expected, low fold changes were observed where only DFB was added from the first to second time points as the low iron status of the incubation remained constant. In the cases where DFB was added after Fe (FeDFB), *FTN* expression decreased following the addition of DFB at the low iron sites and matched expression within the DFB-only treatments. The exception was at P-Low where expression of *FTN* remained high and only marginally lower than the Fe treatment although DFB had been added for 48 hours, suggesting that these oceanic diatoms display a delayed response to iron removal. Once more, these gene expression patterns are consistent with what has been observed for *Pseudo-nitzschia* in the laboratory (Marchetti et al 2009, Marchetti et al 2017), substantiating that members of this genus use ferritin for long-term storage of iron. *Pseudo-nitzschia* are responsive with their *FTN* expression to fluctuations in iron availability, and a positive relationship between *FTN* expression and iron quotas is apparent.

To examine the prevalence of ferritin-utilization in other dominant diatoms, *FTN* expression was examined in *Chaetoceros* and *Thalassiosira* (Figure 2.6). *FTN* was virtually undetected in *Chaetoceros* with only two assigned contigs from the assembly with low read counts. Phylogenetic analysis shows that these contigs are similar to the reference *FTN* sequence from *Chaetoceros dichaeta* (Figure B.9). As few reference *Chaetoceros* transcriptomes have shown *FTN* sequences, this result was anticipated (Cohen et al 2018). Although many *Thalassiosira* species do not possess *FTN*, some species do; however, *FTN* was only detected in low abundances in *Thalassiosira* (Figure B.5 and B.10) and there were no significant

changes in expression (Figure 2.6). Previous studies have also observed minute changes in *Thalassiosira FTN* expression as a function of iron status apart from one *FTN* homolog in *T. rotula* that had higher expression under iron-limited conditions suggesting a role other than iron storage in this diatom (Cohen et al 2018). Among these predominant diatom genera, *Pseudo-nitzschia* appears to be the distinct utilizer of *FTN* for long-term iron storage in their natural environment, whereas the other diatoms are potentially using other mechanisms to store iron.

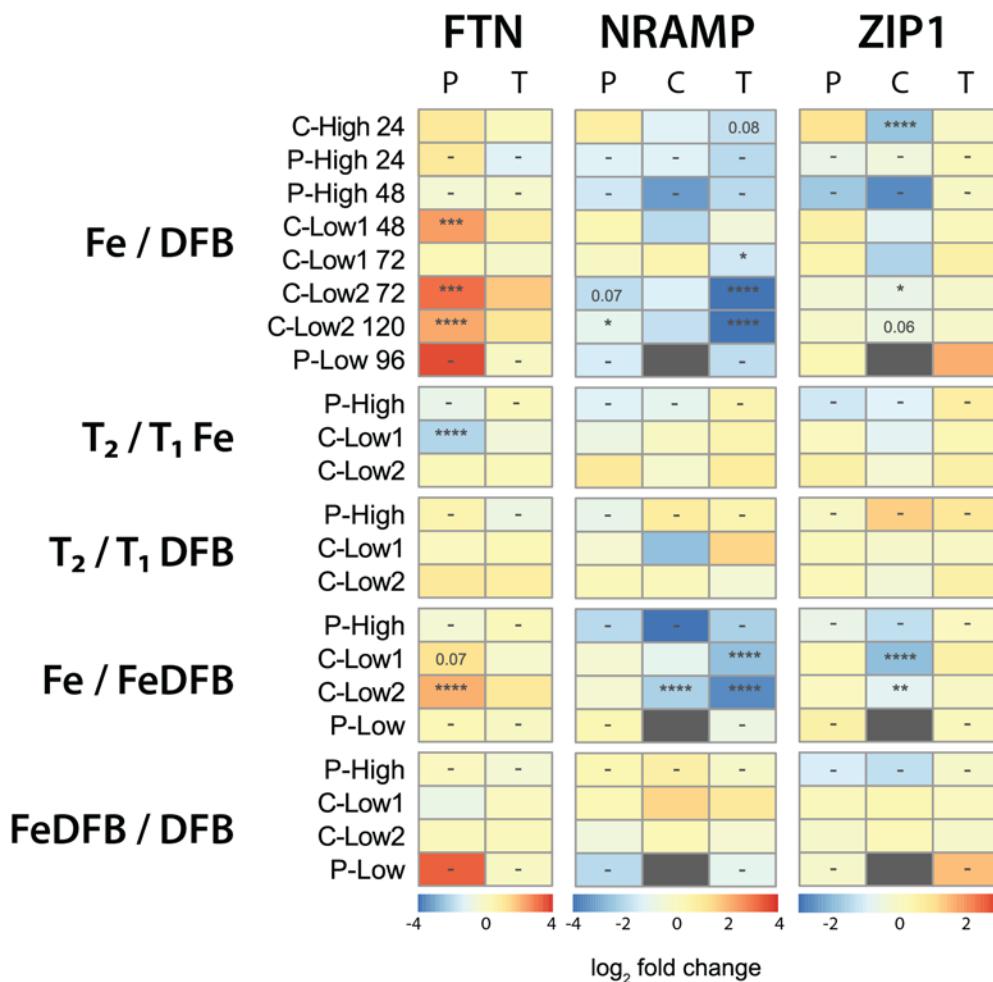


Figure 2.6 - Heatmap of \log_2 fold change values for ferritin (*FTN*), natural resistance-associated macrophage protein (*NRAMP*), and *ZIP1* expression in *Pseudo-nitzschia* (P), *Chaetoceros* (C), and *Thalassiosira* (T). Comparisons are separated into five groups in order from top to bottom: iron addition (Fe) vs iron removal (DFB), T_2 Fe vs T_1 Fe, T_2 DFB vs T_1 DFB, iron addition then removal (FeDFB) vs Fe, and FeDFB vs DFB. Dark gray indicates that the gene was not detected in both samples being compared. Significance in differential expression is shown within each cell where applicable (displayed numerically, $P \leq 0.1$; *, $P \leq 0.05$; **, $P \leq 0.01$; *, $P \leq 0.001$; ****, $P \leq 0.0001$; (-), not applicable).**

Vacuolar-associated Gene Expression

Expression of *NRAMP* was investigated as it may be related to vacuolar storage of iron. This relationship is based on evidence that some *NRAMP* orthologs in other eukaryotes may be involved in vacuolar storage (Lanquar et al 2005, Portnoy et al 2000). *Thalassiosira* *NRAMP* has also been found to be upregulated under low iron quotas and iron concentrations (Kustka et al 2007). As *NRAMP* is a nonspecific divalent metal transporter and Fe uptake in *Thalassiosira* was not inhibited by other divalent metals, *NRAMP* may not be a surface transporter and rather one for transporting iron out of vacuoles (Kustka et al 2007). The presence of iron contained within vacuoles is also supported by spatial elemental analysis of certain non-ferritin containing *Thalassiosira* species (Nuester et al 2012).

Thalassiosira was the genus with highest and most significant differential expression of *NRAMP* (Figure 2.6 and Figure B.5). Expression in the other two genera was usually low, not significantly different, and inconsistent between treatments, indicating that the use of *NRAMP* may be distinct to *Thalassiosira* among the three diatom genera, much like *FTN* in *Pseudo-nitzschia*. Interestingly, *Pseudo-nitzschia* only showed higher expression at P-Low, albeit with low differential expression between treatments (Figure B.5). It is possible that these subarctic oceanic *Pseudo-nitzschia* spp. constitutively express *NRAMP* as a low affinity metal permease on the cell surface. For these diatoms, this function would be advantageous at this site as iron is extremely limiting, Fe(II) oxidation rates are slower in colder waters, and Fe(II) may account for a significant amount of the total dissolved iron (Roy et al 2008). High expression of *NRAMP* in the low iron treatments at this site was also observed in the closely related genus *Fragilaropsis* indicating it may serve a similar role (Figure B.11).

Thalassiosira consistently demonstrated higher *NRAMP* transcript abundance in the DFB treatments with some significant differential expression (Figure 2.6). Expression was consistent when comparing across time in the Fe or DFB treatments. In the iron addition then removal treatment (FeDFB), expression was significantly higher compared to the Fe treatment. Low differences in expression were observed between the FeDFB treatment and the DFB treatment. These results are consistent with the laboratory studies showing higher *NRAMP* expression under low iron concentrations and significant down-regulation following iron resupply in *T. pseudonana* (Kustka et al 2007); thus, they are consistent with what we may expect from the role of *NRAMP* in transporting iron out of a vacuole. In particular, the

high expression in the FeDFB treatment suggests that *Thalassiosira* store iron in vacuoles and then highly express this transporter to shuttle iron out of the vacuole once it is no longer externally available. Given the rapid responsiveness to iron availability observed in the lab (Kustka et al 2007) and the differential expression observed here, it appears likely that *NRAMP* expression is regulated by external iron availability rather than cellular iron quota.

Although this expression supports common use of this vacuolar strategy in *Thalassiosira*, the role of *NRAMP* in diatoms remains rather elusive and may be different between genera as evident from its expression in ferritin-utilizing, oceanic *Pseudo-nitzschia* (Figure 2.6) Phylogenetic analysis of diatom *NRAMP* genes shows that they appear conserved to some degree among diatoms, distinct from *NRAMPs* in other organisms, and to have originated from their red algal ancestor (Figure B.12). The gene, however, is not universal to all diatoms given that an ortholog was not detected in the *Phaeodactylum tricornutum* genome.

Overall, *Chaetoceros* displayed little *FTN* or *NRAMP* expression in these incubations, but iron quotas were similar to that of *Pseudo-nitzschia* within most treatments, particularly at C-Low1 and P-High (Figure 2.4, Figure 2.6, and Figure B.5). From these quotas, it is apparent that *Chaetoceros* is able to store iron at relative quantities similar to that of *Pseudo-nitzschia* within these time frames; however, their iron storage mechanism remains unknown. High expression of *NRAMP* in *Chaetoceros* was found in the FeDFB treatments at C-Low2, but not at C-Low1. The difference appears to be pronounced at P-High, but this result is an artifact of no detection in the Fe treatment (Figure 2.6 and Figure B.5). It is possible that there was some *NRAMP*-utilization by *Chaetoceros* at certain sites, but this response was not always found.

Expression of other iron-related genes that show similar patterns to *FTN* and *NRAMP* were queried. In addition to *NRAMP*, diatoms possess divalent metal transporters belonging to the ZIP family (Allen et al 2008). Like *NRAMP*, these may also be localized to the plasma or vacuolar membrane for passive transport including that of Fe(II) (Eide 2005). Expression of two *ZIP* orthologs (*ZIP1* and *ZIP7*) was detected by all three dominant genera in the metatranscriptomes. *ZIP7* was expressed most by *Pseudo-nitzschia* but not differentially expressed among treatments (Figure B.13). Unlike *FTN* and *NRAMP*, expression of *ZIP1* was expressed at similar levels on average among the three genera;

however, *Chaetoceros* was the only predominant diatom to significantly modify expression of the gene in relation to their Fe status (Figure 2.6). Furthermore, this expression pattern is similar to that of *NRAMP* where increased transcript abundance is observed when DFB is added, particularly when transitioning from high to low iron as in the high iron sites and the FeDFB treatments. As a result, we speculate that *Chaetoceros* may also use a vacuolar storage mechanism but potentially utilizes a different transporter protein, *ZIP1*, for transport out of the vacuole. This gene appears to be similar to that present in some other heterokonts and green algae suggesting different evolutionary origins from *NRAMP* or *FTN* (Figure B.14).

Furthermore, the mechanism of iron import into the vacuole is unknown. Diatoms possess a homolog of the vacuolar iron transporter, *VIT1* or *CCC1*, used for this purpose in *Arabidopsis thaliana* and *Saccharomyces cerevisiae*, respectively (Brembu et al 2011, Kim et al 2006). Expression of these genes was found to be low and inconsistent among treatments in our incubations. Previous experiments show that one of the homologs in *P. tricornutum* was significantly regulated by cadmium; therefore, it is possible that *VIT1* is not related to iron transport in diatoms (Brembu et al 2011).

Biogeochemical and Ecological Implications

The expression of *FTN* in *Pseudo-nitzschia* and *NRAMP* in *Thalassiosira* in these natural communities is consistent with previous laboratory studies substantiating these distinct iron storage strategies—ferritin and vacuolar storage—in ecologically important diatom genera. Out of the three main diatom genera in this study, *Pseudo-nitzschia* was the unique utilizer of *FTN* whereas *NRAMP* was the most highly expressed and regulated in *Thalassiosira*. Transcripts for neither of these genes were abundant in *Chaetoceros* although *Chaetoceros* often had iron quotas similar to those in *Pseudo-nitzschia* leading to speculation that *Chaetoceros* utilizes a different divalent metal transporter, *ZIP1*, for vacuolar iron storage. Interestingly, the presence of these three genes in diatoms also appears to have different evolutionary origins: lateral gene transfer (*FTN*), the red algal ancestor (*NRAMP*), or the green algal ancestor (*ZIP1*). Since ferritin appears to have originated from a lateral gene transfer event from cyanobacteria, but it is present in all four diatom classes (Groussman et al 2015, Marchetti et al 2009), we hypothesize that it was inherited by a diatom ancestor before the first divergence of these lineages,

approximately 56 million years ago (Armbrust 2009). Contributing to diatom diversification, many centric diatoms may have maintained a vacuolar storage mechanism whereas ancestral diatoms to *Pseudo-nitzschia* and *Fragilaropsis* switched to using ferritin for long-term storage and have conserved the gene. As other diatoms maintained a vacuolar storage mechanism, ferritin genes may be absent as a result of gene loss events or present for functions other than long-term storage (Cohen et al 2018).

Ferritin is believed to serve the same function as *Pseudo-nitzschia* in the closely related pennate diatom, *Fragilaropsis* (Marchetti et al 2009). In laboratory studies, ferritin was found to increase in expression under iron-replete conditions in *F. cylindrus* and *F. kerguelensis* (Marchetti et al 2017, Mock et al 2017), while *Fragilaropsis* ferritin expression in the experiments described here was relatively low and sporadic (Figure B.11). Other metatranscriptomes show conflicting results ranging from absent in one from the Southern Ocean (Mock et al 2017), to very low abundances in another at P-Low (Cohen et al 2017a), to abundant but not differentially expressed in another Southern Ocean metatranscriptome (Bertrand et al 2015). These discrepancies suggest that the putative role of ferritin as a long-term iron storage mechanism in *Fragilaropsis* requires further analysis.

Ferritin used for long-term storage provides one explanation for the dominance of *Pseudo-nitzschia* and *Fragilaropsis* in iron-enrichment experiments in the Northeast Pacific and Southern Oceans (de Baar et al 2005, Marchetti et al 2009). As the low iron sites and treatments tended to favor *Pseudo-nitzschia*, we speculate that these differences in storage strategies may have effects on the diatom community composition provided that there is more prolonged iron limitation than what was artificially induced within our incubations. *Fragilaropsis* is known to be abundant in the Southern Ocean, particularly in the open ocean regions where iron delivery is more sporadic and pulsed (Abelmann et al 2006). Closer to the sea ice edge, *Chaetoceros* can be more prevalent. Ferritin may enable facilitated and more uniform distribution of stored iron as the protein may be more easily transferred to daughter cells compared to a vacuole. We speculate that the advantage of iron storage through ferritin versus vacuolar storage applies to regions of ephemeral iron inputs where a high storage capacity can be achieved upon a pulsed input and then passed on to progeny. Over longer time scales without bioavailable iron, ferritin-utilizing diatoms may be able to maintain higher growth rates for a larger and longer-lasting seed population until the next

iron deposition event as observed here at sites C-Low1 and P-Low and in the laboratory (Marchetti et al 2009).

Iron limitation is anticipated to expand as a result of ocean acidification, reduced dust input, and intensified upwelling with increases in upwelled nitrate uncoupled from shelf-derived iron inputs (Capone and Hutchins 2013, Mahowald et al 2005, McQuaid et al 2018, Shi et al 2010). If prolonged iron limitation occurs with greater frequency, a shift towards a system like that observed at P-Low where increases in ferritin-utilizing diatoms such as *Pseudo-nitzschia* in these areas or potentially *Fragilaropsis* in the Southern Ocean are anticipated from gaining longer-lasting seed populations due to their ability to store iron. *Pseudo-nitzschia* in coastal regions are also known for their synthesis of the neurotoxin domoic acid (DA), potentially favoring a system more conducive to its production and leading to an increase in frequency and intensity of harmful algal blooms (Lelong et al 2012). DA was detected in all CUZ incubations with concentrations up to 3.00 ng mL⁻¹ (Figure B.15; Cohen et al 2017b). Furthermore, *Fragilaropsis* can contribute to high biogenic opal burial in the Southern Ocean suggesting an enhancement in the silica pump (Abelmann et al 2006, Assmy et al 2013). Beyond simply the impact that iron limitation has on primary productivity, these long-term community shifts towards ferritin-utilizing diatoms would further influence the marine food web structure and associated biogeochemical cycles.

REFERENCES

- Abelmann A, Gersonde R, Cortese G, Kuhn G, Smetacek V (2006). Extensive phytoplankton blooms in the Atlantic sector of the glacial Southern Ocean. *Paleoceanography* **21**: PA1013.
- Allen AE, Laroche J, Maheswari U, Lommer M, Schauer N, Lopez PJ et al (2008). Whole-cell response of the pennate diatom *Phaeodactylum tricornutum* to iron starvation. *P Natl Acad Sci USA* **105**: 10438-10443.
- Armbrust EV (2009). The life of diatoms in the world's oceans. *Nature* **459**: 185-192.
- Assmy P, Smetacek V, Montresor M, Klaas C, Henjes J, Strass VH et al (2013). Thick-shelled, grazer-protected diatoms decouple ocean carbon and silicon cycles in the iron-limited Antarctic Circumpolar Current. *P Natl Acad Sci USA* **110**: 20633-20638.
- 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.
- Bertrand EM, McCrow JP, Moustafa A, Zheng H, McQuaid JB, Delmont TO et al (2015). Phytoplankton–bacterial interactions mediate micronutrient colimitation at the coastal Antarctic sea ice edge. *P Natl Acad Sci USA* **112**: 9938-9943.
- Birol I, Jackman SD, Nielsen CB, Qian JQ, Varhol R, Stazyk G et al (2009). De novo transcriptome assembly with ABSS. *Bioinformatics* **25**: 2872-2877.
- Bolger AM, Lohse M, Usadel B (2014). Trimmomatic: A flexible trimmer for Illumina Sequence Data. *Bioinformatics* **30**: 2114-2120.
- Botebol H, Lesuisse E, Šuták R, Six C, Lozano J-C, Schatt P et al (2015). Central role for ferritin in the day/night regulation of iron homeostasis in marine phytoplankton. *P Natl Acad Sci USA* **112**: 14652-14657.
- Boyd PW, Jickells T, Law CS, Blain S, Boyle EA, Buesseler KO et al (2007). Mesoscale iron enrichment experiments 1993-2005: synthesis and future directions. *Science* **315**: 612-617.
- Brembu T, Jorstad M, Winge P, Valle KC, Bones AM (2011). Genome-wide profiling of responses to cadmium in the diatom *Phaeodactylum tricornutum*. *Environ Sci Technol* **45**: 7640-7647.
- Bruland KW, Rue EL, Smith GJ (2001). Iron and macronutrients in California coastal upwelling regimes: Implications for diatom blooms. *Limnol Oceanogr* **46**: 1661-1674.
- Capone DG, Hutchins DA (2013). Microbial biogeochemistry of coastal upwelling regimes in a changing ocean. *Nature Geosci* **6**: 711-717.
- Caron DA, Alexander H, Allen AE, Archibald JM, Armbrust EV, Bachy C et al (2017). Probing the evolution, ecology and physiology of marine protists using transcriptomics. *Nat Rev Micro* **15**: 6-20.
- Cohen NR, A. Ellis K, Burns WG, Lampe RH, 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 NR, Ellis KA, Lampe RH, McNair HM, Twining BS, Brzezinski MA et al (2017b). Variations in diatom transcriptional responses to changes in iron availability across ocean provinces. *Front Mar Sci* **4**: 360.

Cohen NR, Mann E, Stemple B, Raushenberg S, Jacquot J, Moreno CM et al (2018). Iron storage capacities and associated ferritin gene expression among marine diatoms. *Limnol Oceanogr*. doi:10.1002/lno.10800

Droop MR (1973). Some thoughts on nutrient limitation in algae. *J Phycol* **9**(3): 264-272.

de Baar HJW, Boyd PW, Coale KH, Landry MR, Tsuda A, Assmy P et al (2005). Synthesis of iron fertilization experiments: From the Iron Age in the Age of Enlightenment. *J Geophys Res-Oceans* **110**: C09S16.

Eide DJ (2005). The Zip Family of Zinc Transporters. In: Iuchi S, Kuldell N (eds). *Zinc Finger Proteins: From Atomic Contact to Cellular Function*. Springer US: Boston, MA. pp 261-264.

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**: 645-656.

Groussman RD, Parker MS, Armbrust EV (2015). Diversity and Evolutionary History of Iron Metabolism Genes in Diatoms. *PLoS ONE* **10**: e0129081.

Harrison PJ (2002). Station Papa Time Series: Insights into Ecosystem Dynamics. *Journal of Oceanography* **58**: 259-264.

Hutchins DA, DiTullio GR, Zhang Y, Bruland KW (1998). An iron limitation mosaic in the California upwelling regime. *Limnol Oceanogr* **43**: 1037-1054.

Kanehisa M, Furumichi M, Tanabe M, Sato Y, Morishima K (2017). KEGG: new perspectives on genomes, pathways, diseases and drugs. *Nucleic Acids Res* **45**: D353-D361.

Keeling PJ, Burki F, Wilcox HM, Allam B, Allen EE, Amaral-Zettler LA 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**: e1001889.

Keren N, Aurora R, Pakrasi HB (2004). Critical Roles of Bacterioferritins in Iron Storage and Proliferation of Cyanobacteria. *Plant Physiol* **135**: 1666-1673.

Kim SA, Punshon T, Lanzilotti A, Li L, Alonso JM, Ecker JR et al (2006). Localization of Iron in *Arabidopsis*; Seed Requires the Vacuolar Membrane Transporter VIT1. *Science* **314**: 1295.

Kustka AB, Allen AE, Morel FMM (2007). Sequence analysis and transcriptional regulation of iron acquisition genes in two marine diatoms. *J Phycol* **43**: 715-729.

Lanquar V, Lelièvre F, Bolte S, Hamès C, Alcon C, Neumann D et al (2005). Mobilization of vacuolar iron by AtNRAMP3 and AtNRAMP4 is essential for seed germination on low iron. *EMBO J* **24**: 4041-4051.

Lelong A, Hégaret H, Soudant P, Bates SS (2012). *Pseudo-nitzschia* (Bacillariophyceae) species, domoic acid and amnesic shellfish poisoning: revisiting previous paradigms. *Phycologia* **51**: 168-216.

Liochev SI, Fridovich I (1999). Superoxide and Iron: Partners in Crime. *IUBMB Life* **48**: 157-161.

Long JC, Sommer F, Allen MD, Lu S-F, Merchant SS (2008). FER1 and FER2 Encoding Two Ferritin Complexes in *Chlamydomonas reinhardtii* Chloroplasts Are Regulated by Iron. *Genetics* **179**: 137-147.

Love MI, Huber W, Anders S (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology* **15**: 550.

- Mahowald N, Baker AR, Bergametti G, Brooks N, Duce RA, Jickells TD *et al* (2005). Global Iron Connections Between Desert Dust, Ocean Biogeochemistry, and Climate. *Science* **308**: 67-71.
- Malviya S, Scalco E, Audic S, Vincent F, Veluchamy A, Poulaire J *et al* (2016). Insights into global diatom distribution and diversity in the world's ocean. *P Natl Acad Sci USA* **113**: E1516-E1525.
- Marchetti A, Parker MS, Moccia LP, Lin EO, Arrieta AL, Ribalet F *et al* (2009). Ferritin is used for iron storage in bloom-forming marine pennate diatoms. *Nature* **457**: 467-470.
- Marchetti A, Schruth DM, Durkin CA, Parker MS, Kodner RB, Berthiaume CT *et al* (2012). Comparative metatranscriptomics identifies molecular bases for the physiological responses of phytoplankton to varying iron availability. *P Natl Acad Sci USA* **109**: E317-E325.
- Marchetti A, Maldonado MT (2016). Iron. In: Borowitzka MA, Beardall J, Raven JA (eds). *The Physiology of Microalgae*. Springer International Publishing: Cham. pp 233-279.
- Marchetti A, Moreno CM, Cohen NR, Oleinikov I, deLong K, Twining BS *et al* (2017). Development of a molecular-based index for assessing iron status in bloom-forming pennate diatoms. *J Phycol* **53**: 820-832.
- Martin JH, Fitzwater SE (1988). Iron deficiency limits phytoplankton growth in the north-east Pacific subarctic. *Nature* **331**: 341-343.
- McQuaid JB, Kustka AB, Oborník M, Horák A, McCrow JP, Karas BJ *et al* (2018). Carbonate-sensitive phytotransferrin controls high-affinity iron uptake in diatoms. *Nature* **555**: 534.
- Mock T, Medlin LK (2012). Genomics and Genetics of Diatoms. *Adv Bot Res* **64**: 245-284.
- Mock T, Ollilar RP, Strauss J, McMullan M, Paajanen P, Schmutz J *et al* (2017). Evolutionary genomics of the cold-adapted diatom *Fragilariaopsis cylindrus*. *Nature* **541**: 536-540.
- Moore JK, Doney SC, Glover DM, Fung IY (2001). Iron cycling and nutrient-limitation patterns in surface waters of the World Ocean. *Deep-Sea Res Pt II* **49**: 463-507.
- Morrissey J, Sutak 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. *Current Biology* **25**: 364-371.
- Nuester J, Vogt S, Twining BS (2012). Localization of iron within centric diatoms of the genus *Thalassiosira*. *J Phycol* **48**: 626-634.
- 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**: 560-566.
- Patro R, Duggal G, Love MI, Irizarry RA, Kingsford C (2017). Salmon provides fast and bias-aware quantification of transcript expression. *Nat Meth* **14**: 417-419.
- Pfaffen S, Bradley JM, Abdulqadir R, Firme MR, Moore GR, Le Brun NE *et al* (2015). A Diatom Ferritin Optimized for Iron Oxidation but not Iron Storage. *J Biol Chem* **290**: 28416-28427.
- Portnoy ME, Liu XF, Culotta VC (2000). *Saccharomyces cerevisiae* expresses three functionally distinct homologues of the nramp family of metal transporters. *Mol cell bio* **20**: 7893-7902.
- Robertson G, Schein J, Chiu R, Corbett R, Field M, Jackman SD *et al* (2010). De novo assembly and analysis of RNA-seq data. *Nat Meth* **7**: 909-912.

Roy EG, Wells ML, King DW (2008). Persistence of iron(II) in surface waters of the western subarctic Pacific. *Limnol Oceanogr* **53**: 89-98.

Shi D, Xu Y, Hopkinson BM, Morel FMM (2010). Effect of Ocean Acidification on Iron Availability to Marine Phytoplankton. *Science* **327**: 676-679.

Twining BS, Baines SB, Fisher NS, Maser J, Vogt S, Jacobsen C et al (2003). Quantifying Trace Elements in Individual Aquatic Protist Cells with a Synchrotron X-ray Fluorescence Microprobe. *Analytical Chemistry* **75**: 3806-3816.

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**: 1827-1850.

Twining BS, Baines SB, Bozard JB, Vogt S, Walker EA, Nelson DM (2011). Metal quotas of plankton in the equatorial Pacific Ocean. *Deep-Sea Res Pt II* **58**: 325-341.

APPENDIX A: CHAPTER 1 SUPPLEMENTARY INFORMATION

Supplementary Materials and Methods

Trace Metal Clean Techniques

Seawater collection was performed using a trace metal clean sampling system consisting of an air-driven PTFE (polytetrafluoroethylene) deck pump (Wilden, Grand Terrace, CA, USA) fitted with PTFE tubing. The pump and tubing were cleaned by pumping 1% trace metal grade HCl overnight followed by an overnight rinse with ultrapure water.

Incubations followed techniques as described by Marchetti et al (2012). Briefly, cleaning the cubitainers included soaking the inside walls in 1.2 mol L⁻¹ hydrochloric acid (reagent grade) for 3 d followed by three rinses with Milli-Q H₂O, soaking in 1.2 mol L⁻¹ hydrochloric acid (trace metal grade) for 1 week followed by three rinses with Milli-Q H₂O, and soaking in 0.1 mol L⁻¹ acetic acid (trace-metal grade). Prior to filling the cubitainers with seawater, the dilute acetic acid was removed and the cubitainers were rinsed thoroughly three times with ambient seawater.

CTD, Satellite, and Meteorological Data

Potential temperature, density, and photosynthetically active radiation were obtained from sensors mounted on a 24-bottle rosette onboard the R/V Melville (Seabird 911+ conductivity-temperature depth sensor). Satellite-derived sea surface temperature data is from the NOAA POES AVHRR satellite courtesy of the NOAA / NESDIS Center for Satellite Applications Research. These data were downloaded from the NOAA CoastWatch Browser (<http://coastwatch.pfeg.noaa.gov/coastwatch/CWBrowser.jsp>). Data are plotted using Interactive Data Language® (Exelis Visual Information Solutions Inc., Boulder, CO, USA). Wind speed and direction was obtained from the shipboard meteorological system (MetAcq) on the R/V Melville.

Microscopy and imaging flow cytometry (FlowCAM)

The viability of cells collected from 96 m was confirmed through imaging flow cytometry using a Flow Cytometer and Microscope (FlowCAM, Fluid Imaging Technologies Inc., Scarborough, ME, USA) which was operated in trigger mode. In this mode, image acquisition by a CCD camera is triggered by

chlorophyll or phycoerythrin fluorescence. Cells with detectable fluorescence were observed, although the initial phytoplankton populations were small (~75 cells mL⁻¹ were found).

Phytoplankton cell abundances and species composition were determined by microscopic examination and by examining images of live cells captured using an imaging flow cytometer. For microscopic counts, 50 mL samples were preserved in 2% Lugol's Iodine and settled for >24 hours in Utermöhl chambers (Utermöhl 1958). Counts were performed at 100x, 200x, and 400x using a Leica DML inverted microscope on a minimum of 400 total cells in at least five fields of view. For data collection using FlowCAM, 50 mL samples were drawn from the well-mixed cubitainers into 50 mL Falcon tubes and stored at 4°C in the dark pending processing (typically within 3 hours of collection). Using FlowCAM's Visual Spreadsheets v3.1 software, a CCD camera captures images of individual cells with a sufficiently large fluorescence signal corresponding to chlorophyll *a* and/or phycoerythrin fluorescence. At least 5 mL of sample was filtered through 300 µm nitex mesh and passed through the system at a flow rate of 0.2-0.3 mL min⁻¹ using a syringe pump equipped with a 5 mL glass syringe. A 100 µm flow cell was used in all sample runs. A digital size filter was applied so that only cells >5 µm were captured in images. The flow cell and tubing were well flushed with Milli-Q water and 70% ethanol between each sample run to avoid cross-contamination.

Fv:Fm

Maximum photochemical yield of Photosystem II ($F_v:F_m$) was measured using fast repetition rate fluorometry and a custom built fluorescence-induction and relaxation system (Gorbunov and Falkowski 2005, Kolber et al 1998). Samples were acclimated to low light for 20 min prior to measuring the minimum (F_0) and maximum (F_m) fluorescence yields. Data were blank corrected using 0.2 µm filtered seawater. The resulting $F_v:F_m$ was derived from the induction profile using a saturating pulse (20,000 µmol photons m² s⁻¹) for a duration of 100-200 µs. The average of 100 iterations was obtained.

Domoic acid

250 mL of seawater was filtered onto a GF/F filter (25 mm) via gentle vacuum pressure (<100 mm Hg) and frozen at -80°C. Filters were extracted with 10 mL of 20% methanol (MeOH) in water in a 15

mL centrifuge tube, sonicated (2 min, 30-40 W) in an ice bath with a Sonicator 3000 equipped with microtip (Misonix, Framingdale, NY, USA), and then centrifuged (10 min, 1399 x g). The supernatant was collected and passed through a 0.22 µm syringe filter into a clean 15 mL centrifuge tube to remove remaining particles. Samples were stored at -20°C until analysis. Concentrations with a detection limit of 0.01 µg L⁻¹ were obtained using an enzyme-linked immunosorbent assay (Abraxis, Warminster, PA, USA) following the manufacturer's protocol including running each sample in duplicate at several dilutions. Final concentrations (pg DA mL extract⁻¹) were calculated using the manufacturer supplied analysis spreadsheet.

Nutrient Analyses

Dissolved nitrate + nitrite (NO₃⁻ + NO₂⁻), phosphate (PO₄³⁻), and silicic acid (H₄SiO₄) concentrations were measured using a Lachat Quick Chem 8000 Flow Injection Analysis system (Parsons et al 1984). Particles were removed by filtering the sample through a GF/F filter using a syringe prior to analysis. Reference materials for nutrients in seawater (Lots BY and CA, KANSO Technos, Osaka, Japan) were run alongside samples for quality control.

Dissolved Iron

Samples for dissolved iron (Fe) analysis were filtered through a 0.4 µm polycarbonate filter (47 mm) held in PTFE filter towers (Savillex Corporation, Eden Praire, MN, USA) into LDPE bottles that had been cleaned as per the GEOTRACES cookbook (<http://www.geotraces.org/science/intercalibration/222-sampling-and-sample-handling-protocols-for-geotraces-cruises>). Samples were acidified at sea with the equivalent of 4 mL 6 N quartz-distilled HCl per liter of seawater (to pH 1.7-1.8). Iron samples were analyzed with an adaptation of Biller and Bruland (2012) as described in Parker et al (2016). Briefly, this method involves preconcentrating the Fe on columns of Nobias chelate PA1 resin and analyzing the eluent on the Thermo-Element high resolution XR ICP-MS. Samples are buffered to pH 6.0 with NH₄Ac immediately before loading sit buffered for about 40 minutes while loading. The samples are eluted from the columns with ~1 mL of 1N quartz-distilled HNO₃, resulting in a concentration factor of roughly 24.

Calibration is performed using a spiked standard curve made in low-metal seawater that is preconcentrated in the same manner as the samples.

Phytoplankton isolation and identification

To complement existing reference sequence databases of phytoplankton isolates, eight species of phytoplankton were isolated from the California Upwelling Zone during the cruise period and their transcriptomes were sequenced. Isolations were performed using an Olympus CKX41 inverted microscope by single cell isolation with a micropipette. Isolates were maintained at 14°C and grown under iron-replete conditions in Aquil culture medium (Price et al 1989). Identification of most species was performed by morphological characterization and 18S rRNA gene sequencing. DNA was extracted with the DNeasy Plant Mini Kit according to the manufacturer's protocols (Qiagen, Hilden, Germany). Amplification of the nuclear 18S rDNA region was conducted with standard PCR protocols using eukaryotic-specific, universal 18S forward and reverse primers. Sequences of primers used are as follows: 18AF 5'- AACCTGGTTGATCCTGCCAGT -3' and 18BR 5'- TGATCCTTCTGCAGGTTCACCTAC -3'. The length of the region amplified is approximately 1600 base pairs. *Pseudo-nitzschia* species were identified through sequencing of the 18S-ITS1-5.8S regions. Amplification of this region was performed with 18ST-euk and 5.8SR-euk primers according to Hubbard et al (2008). PCR products were purified using the QIAquick PCR Purification Kit (Qiagen, Hilden, Germany) and sequenced by Sanger DNA sequencing (Genewiz, Morrisville, NC, USA). Species identity was determined on the basis of sequence homology (BLASTN) against the NCBI non-redundant nucleotide collection with a cutoff identity of 99%.

Isolate transcriptomes

Isolate cultures were grown to late exponential/early stationary phase and filtered onto 3.0 µm polycarbonate filters (25 mm). Filters were stored at -80°C. Total RNA was extracted using the RNAqueous-4PCR Total RNA Isolation Kit (Ambion, Foster City, CA, USA) according to the manufacturer's protocol with an initial bead beating step to disrupt cells. Trace DNA contamination was removed by DNase 1 (Ambion, Foster City, CA, USA) digestion at 37°C for 45 min.

Libraries were created with the Illumina TruSeq Stranded mRNA Library Preparation Kit. Ten samples were barcoded and pooled on a single lane and sequenced on an Illumina MiSeq (300 bp, paired end). Reads were trimmed for quality and to remove adapters using Trimmomatic v0.32 (paired-end mode, adaptive quality trim with 40bp target length and strictness of 0.6, minimum length of 36bp; Bolger et al 2014). The resulting trimmed reads were assembled *de novo* using Trinity v2.0.6 with default parameters for paired reads (Grabherr et al 2011). Proteins were predicted from assembled contigs with GeneMark S-T (Tang et al 2015). A summary of transcriptome sequencing, assembly, and gene prediction is provided in Table S4.

Module-Based Differential Expression and Microarray Study Sources

For KEGG module-based differential expression, quantitative metabolic fingerprinting was used (Alexander et al 2015a). Briefly, read counts annotated for each KEGG Module category were summed and then normalized by the total number of reads for the time point per functional grouping. These data were also visualized with pheatmap v1.0.8. Gene expression ratios for microarray studies and orthologous genes were obtained from The Diatom Portal (Ashworth et al 2016).

Phylogenetic Trees

Reference sequences from *Oryzias latipes* Hd-rR (accession number NW_004089490.1) and *Pseudo-nitzschia multiseries* CLN-47 (protein ID 146448, Joint Genome Institute database) for soluble guanylate cyclase (GUCY1B) were used to query MarineRefII (BLASTP v2.5.0). For nitrate reductase (NR), a reference sequence from *Thalassiosira pseudonana* CCMP 1335 (protein ID 25299, Joint Genome Institute database) was used. Sequences were aligned using MUSCLE (Edgar 2004) in Geneious v5.6.4, and then used to create a maximum likelihood phylogenetic tree with RAxML v8.2.9 (PROTGAMMALG model, 100 bootstraps) (Stamatakis 2014). For NR, contigs with greater than 100 total reads for 0 and 72 hours were placed on the reference tree with pplacer v1.1.alpha18-2-gcb55169 (Matsen et al 2010). Trees were visualized and edited using Archeopteryx v0.9916 (Han and Zmasek 2009).

Supplementary Text

Metatranscriptome Sequencing, Assembly, and Annotation Results

Sequencing of total mRNA from the eukaryotic phytoplankton community yielded over 340 million reads (Table A.1). Metatranscriptome assembly of these reads generated over 3.1 million contigs. Functional assignment of contigs was the most limiting factor with only 37.3% that have an annotation from KEGG (Table A.3). Approximately sixty percent of contigs had a taxonomic annotation and mapped reads (Table A.3). Annotation was validated by conducting phylogenetic analysis of high read abundance contigs for soluble guanylate cyclase (GUCY1B) and nitrate reductase in *Pseudo-nitzschia* and *Chaetoceros* (Figures A.11, A.12). Altogether, this level of annotation is similar to previous metatranscriptomic studies utilizing KEGG and reference transcriptomes from the Marine Microbial Eukaryote Transcriptome Sequencing Project (MMETSP; Alexander et al 2015b, Cohen et al 2017, Gong et al 2017, Keeling et al 2014).

In supplementing the MMETSP transcriptomes with those of isolates from the California Upwelling Zone, the level of taxonomic annotation was improved by an additional 19,604 contigs (0.6%). This result could be expected considering four out of the eight added species already had representation in the MMETSP database (Table A.2). Out of the four isolate species not previously included, three were *Chaetoceros* spp. and the other was *Pseudo-nitzschia americana*; thus, there was little representation added considering such low diversity in the new transcriptomes. Seven percent of contigs, however, changed taxonomic annotation (lower E-value) from one of the MMETSP strains to the isolates we collected which could be due to more accurate assignment of *Chaetoceros* originating contigs or strain-specific differences between species. This improved homology represents an increase in confidence that these transcripts in our environmental transcriptome belonged to the genera that were isolated.

Expression on Different Levels of Annotation

Although KEGG Orthologs (KOs) can be grouped into tighter functional groups, or modules (Kanehisa et al 2012), the amount of annotation on this level is highly limiting. Only 8.8% of contigs met the annotation requirements for analysis on the module level compared to 21.8% with KO annotation (Table A.3). It was often the case that KEGG modules or module classes are not fully represented by

their constituent KOs. This low level of annotation and lack of completeness makes any confident inference from examining module expression such as quantitative metabolic fingerprinting (Alexander et al 2015a) difficult (Figure A.13).

Analysis on the KEGG Ortholog level provides additional limitations as it does not account for all annotated genes. By removing annotated contigs without a KEGG Ortholog assignment, the reduction in annotation is from 37.3% to 21.8% of all contigs (Table A.3), many of which are taxonomically assigned to diatoms. These genes are lacking KEGG Orthology annotation because in most cases, their function is unknown; therefore, they are simply annotated as a hypothetical protein. These are derived from the two complete diatom genomes integrated into KEGG: *T. pseudonana* and *P. tricornutum*.

Of the genes that are annotated and significantly overrepresented in the light, 17 were genes encoding fucoxanthin chlorophyll binding proteins. This result matches our observation of overrepresentation of chlorophyll synthesis genes as well as expression in previous studies when diatoms are growing exponentially (Ashworth et al 2013) or exposed to light after prolonged darkness (Leblanc et al 1999).

Although function of these genes is unknown, expression can be compared to *T. pseudonana* that has been grown exponentially for three days (matching our 72 hour time point) (Ashworth et al 2013). Over 350 unknown genes without KO annotation that were significantly overrepresented at 72 hours in our study were also found to have positive fold changes at the exponential growth phase at 72 hours in *T. pseudonana*. Another study examined *P. tricornutum*'s response to re-illumination after prolonged darkness (Nymark et al 2013). Almost 450 genes without KO annotation were significantly overrepresented at 72 hours had positive fold changes 24 hours after re-illumination. Seventeen of these genes were found to be orthologous between *T. pseudonana* and *P. tricornutum* suggesting they serve the same function across these diatom species. These unknown genes may be important to the growth and light response of diatoms providing basis for future investigations.

Chaetoceros and Pseudo-nitzschia Genes at Depth

Several genes from *Chaetoceros* and *Pseudo-nitzschia* highly expressed in the pre-upwelled condition promote proteasome and ubiquitin activity suggesting that the cells are degrading unneeded or

damaged proteins. RNA turnover was likely also increased with high expression of exosome-related transcripts, while expression of protection of telomeres protein promoted stabilization of DNA (Miyoshi et al 2008) (Figure 1.4b). One highly expressed gene was chondroitin sulfate synthase which is potentially related to transparent exopolymer particle (TEP) production. TEP is found to be generated by *Chaetoceros* within the stationary phase which may have contributed to the aggregation and sinking of these cells as well as the high particulate carbon-to-nitrogen ratios observed in the initial community (Figure 1.d) (Passow 2002a, Passow 2002b).

Soluble guanylate cyclase

The beta subunit of soluble guanylate cyclase was not detected in two *Pseudo-nitzschia* transcriptomes from *P. americana* and *P. granii*, but was found in the *Pseudo-nitzschia multiseries* genome. It may not have been detected in those two species due to lack of sequencing depth for those transcriptomes. The two contigs from our environmental transcriptome with almost all of the read counts for this diatom gene appear to be highly phylogenetically related to the *Pseudo-nitzschia* genes in our reference database further supporting that the taxonomic annotation of this gene is accurate. Interestingly, it is not found in the *Thalassiosira pseudonana* or *Phaeodactylum tricornutum* genomes (Armbrust et al 2004, Bowler et al 2008).

Figure A.1 - Satellite-derived sea surface temperature prior to the incubations. The + indicates the location of seawater collection. (A) Three day average from 02 July 2014 to 04 July 2014. (B) 16 July 2014, one day before the incubations. Data is from the NOAA POES AVHRR satellite courtesy of the NOAA / NESDIS Center for Satellite Applications Research, downloaded from the NOAA CoastWatch Browser, and are plotted using Interactive Data Language v8.6 (Exelis Visual Information Solutions, Inc., Boulder, CO, USA).

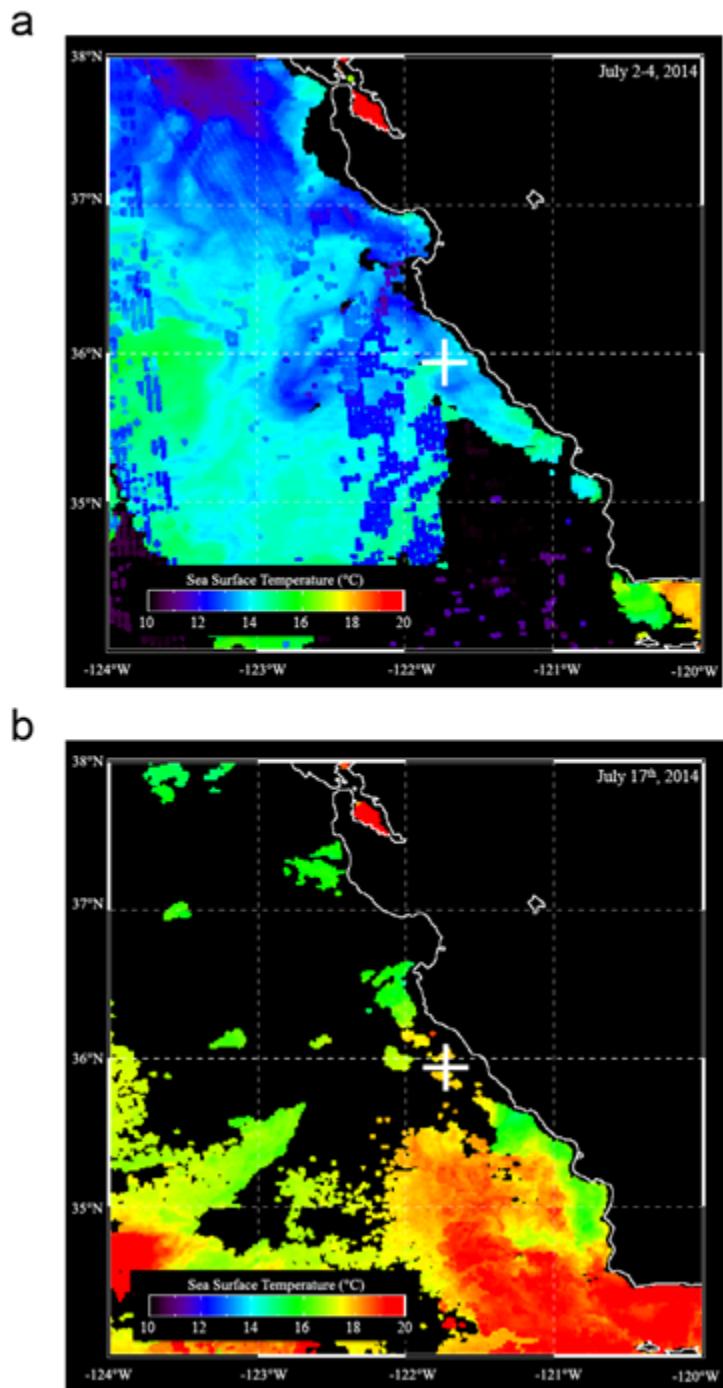


Figure A.2 - Wind speed (m s^{-1}) and direction from the R/V Melville for the 48 hours prior to starting the incubations. Spokes display the frequency of winds blowing from particular directions with color-coded bands showing wind speed ranges.

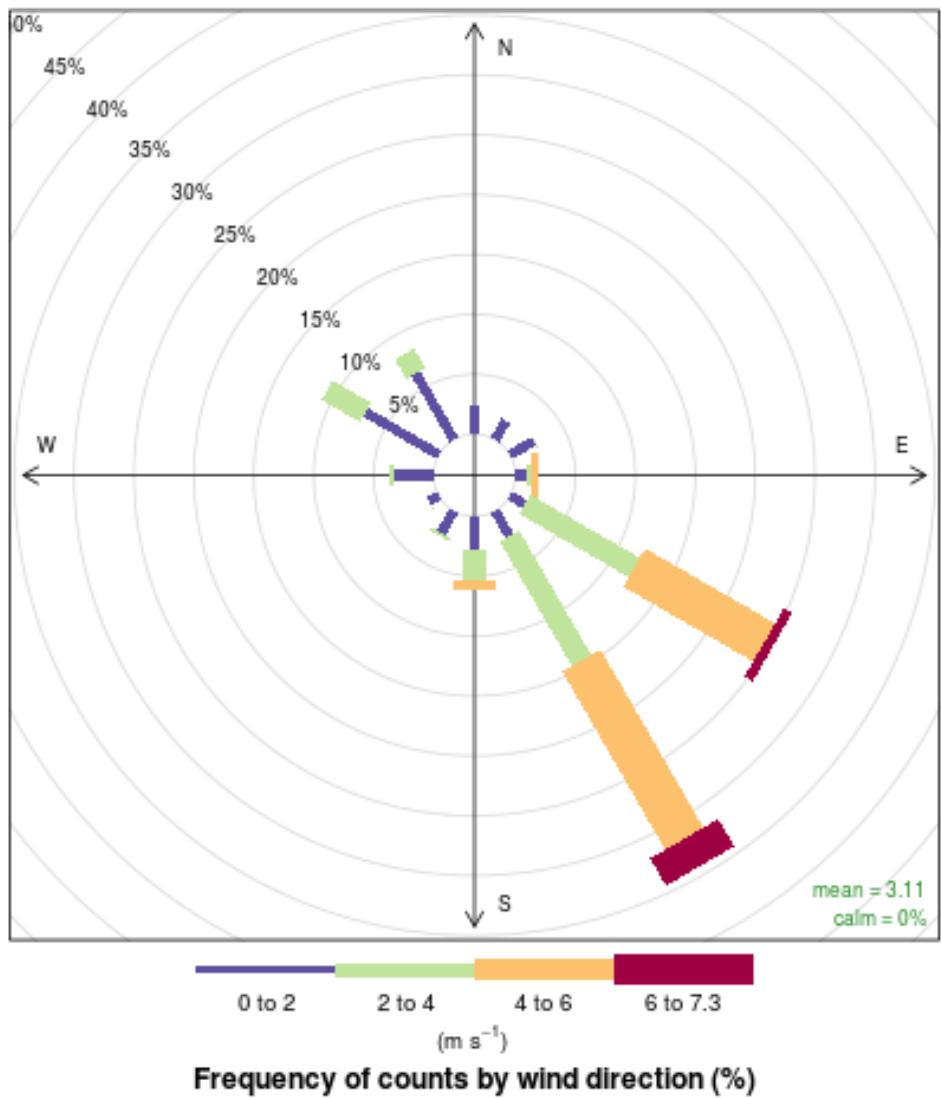


Figure A.3 - CTD (conductivity-temperature-depth) measurements: (A) potential temperature ($^{\circ}\text{C}$), density (σ_t ; kg m^{-3}), (B) fluorescence (raw fluorescence units), and photosynthetically active radiation (PAR; $\mu\text{mol photons m}^{-2} \text{s}^{-1}$) on 16 July, 2014 (PDT), the afternoon prior to the start of incubations.

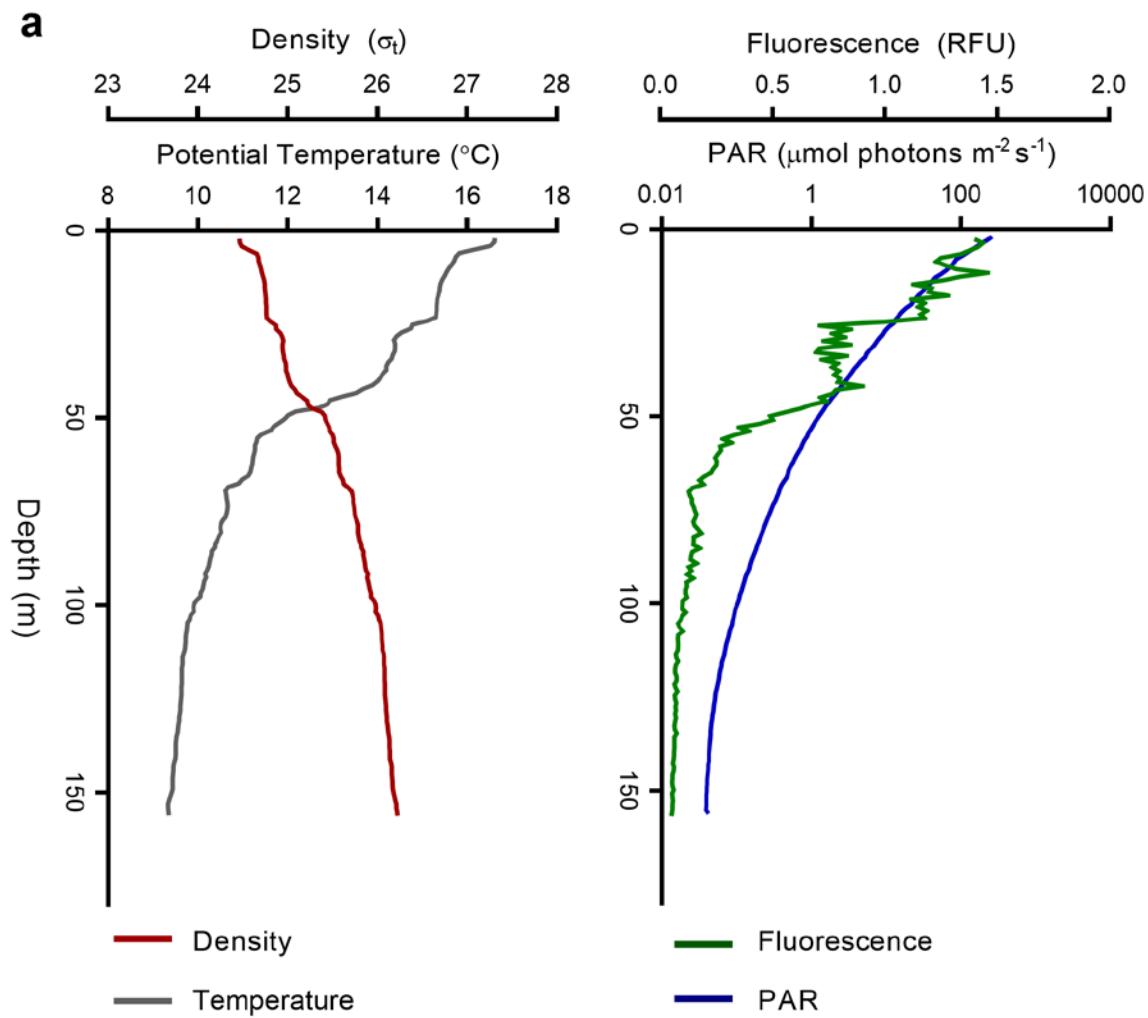


Figure A.4 - Temperature and on-deck photosynthetically active radiation (PAR) during the incubations. Values for every 15 minutes from a HOBO Data Logger (Onset, Cape Code, MA, USA) are plotted.

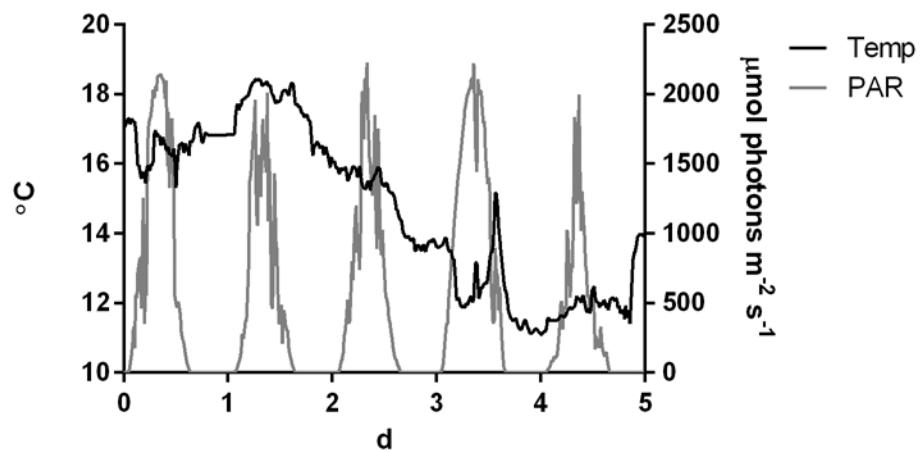


Figure A.5 - Biological coefficient of variation (square root of dispersion) for diatom genes. The biological coefficient of variation is the coefficient of variation with which the true abundance of a gene varies between replicate samples (Chen et al 2014). The tagwise dispersion based on the trended dispersion was used to assess significance in differential expression.

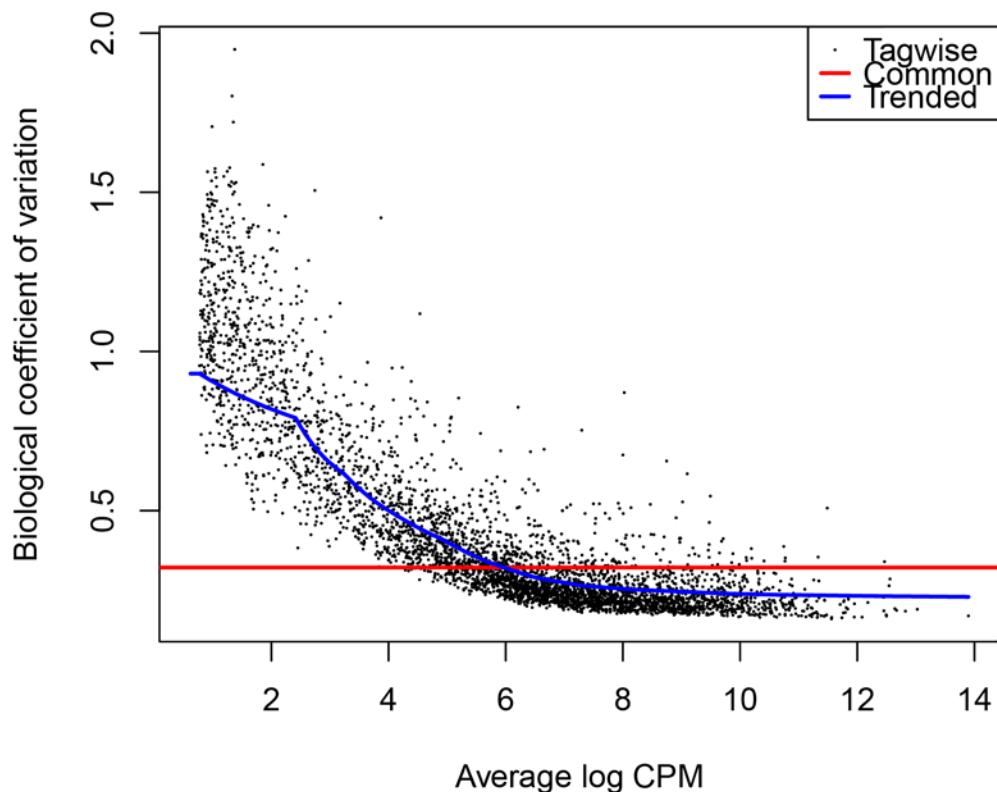


Figure A.6 - Additional measurements from the simulated upwelling incubation experiment: (A) Dissolved macronutrient concentrations: nitrate + nitrite ($\text{NO}_3^- + \text{NO}_2^-$), silicic acid (H_4SiO_4), phosphate (PO_4^{3-}), and iron (Fe). All concentrations are in $\mu\text{mol L}^{-1}$ except Fe which is in nmol L^{-1} . (B) Particulate domoic acid concentrations. (C) Maximum photochemical yield of photosystem II ($F_v:F_m$). (D) Particulate carbon (PC). PC is the only parameter displayed in this figure with size fractionation: $>5 \mu\text{m}$ (red) and $<5 \mu\text{m}$ (blue). Error bars indicate standard deviation of the mean ($n = 3$).

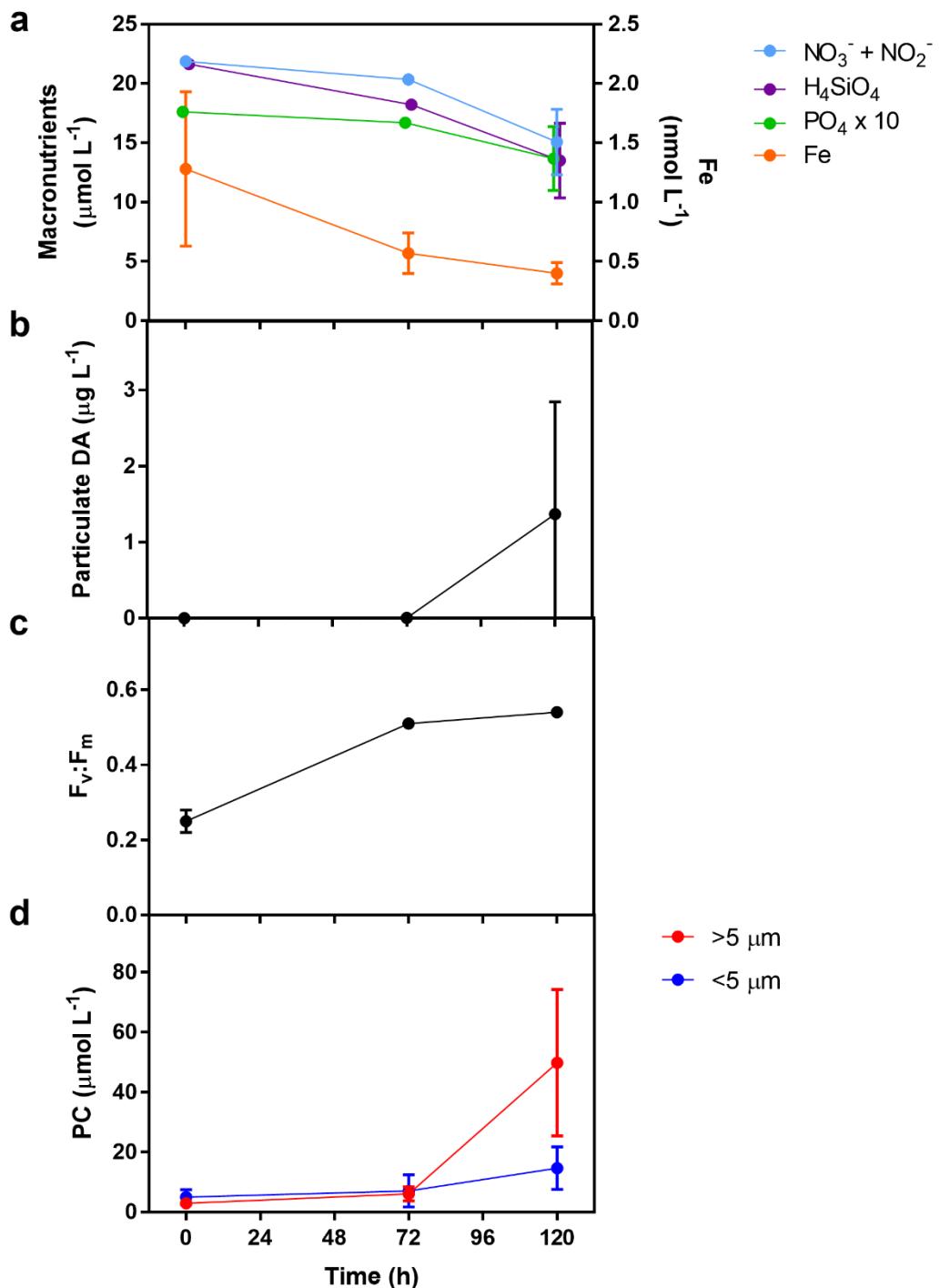


Figure A.7 - Histograms of KEGG Ortholog (KO) gene expression between 72 and 120 hours for the four main phytoplankton groups detected in the study. The four main phytoplankton groups are colored as follows: diatoms (blue), dinoflagellates (red), chlorophytes (green), and haptophytes (orange). KO counts are binned by \log_2 fold change intervals of 1 for 72 and 120 hours. Dashed vertical lines indicate a \log_2 fold change (120 / 72) of -1 or 1. The number and percentage of significantly overrepresented genes at 120 hours (right) and at 72 hours (left) are annotated on each plot.

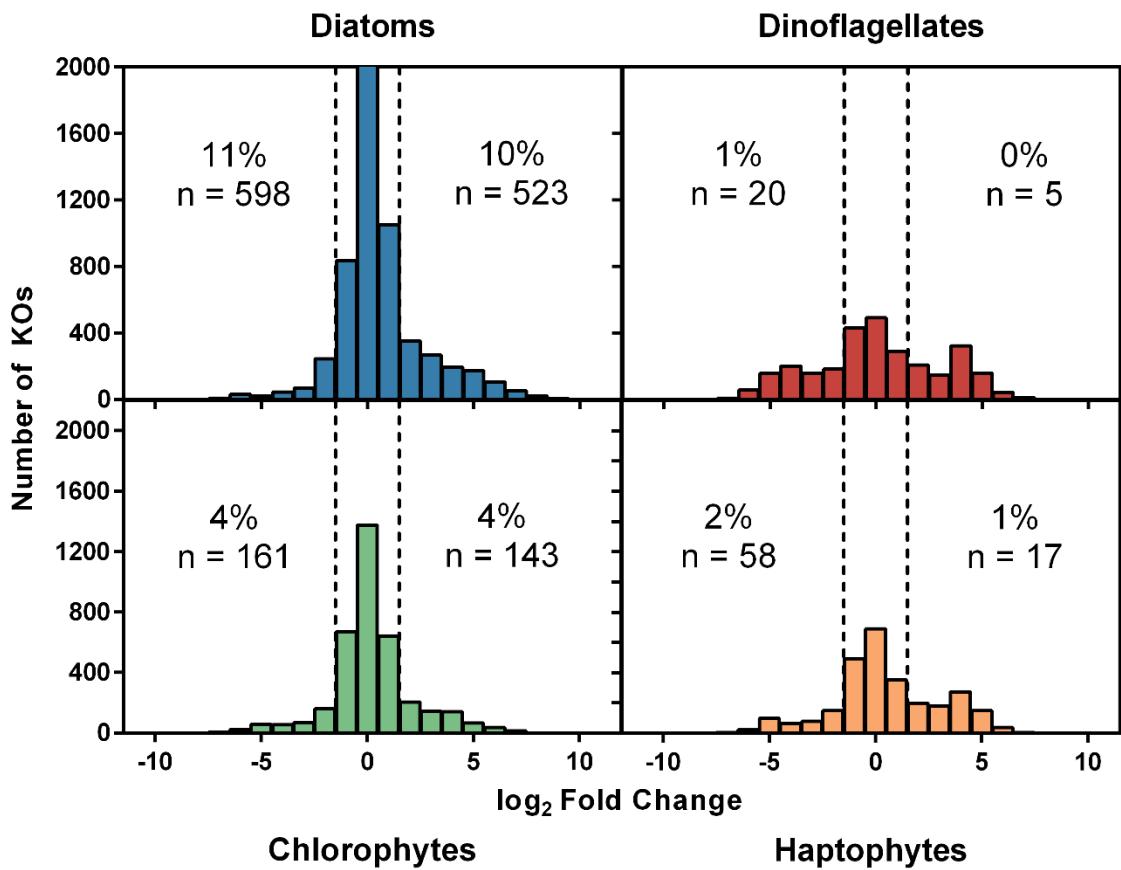


Figure A.8 - Heatmaps for expressed KOs at 72 and 0 hours: (A) the 1,070 expressed KOs where the gene was not detected in one group other than diatoms. Black bars indicate that the gene was not detected. (B) The 1,170 expressed KOs where the gene was detected in diatoms and only one other group.

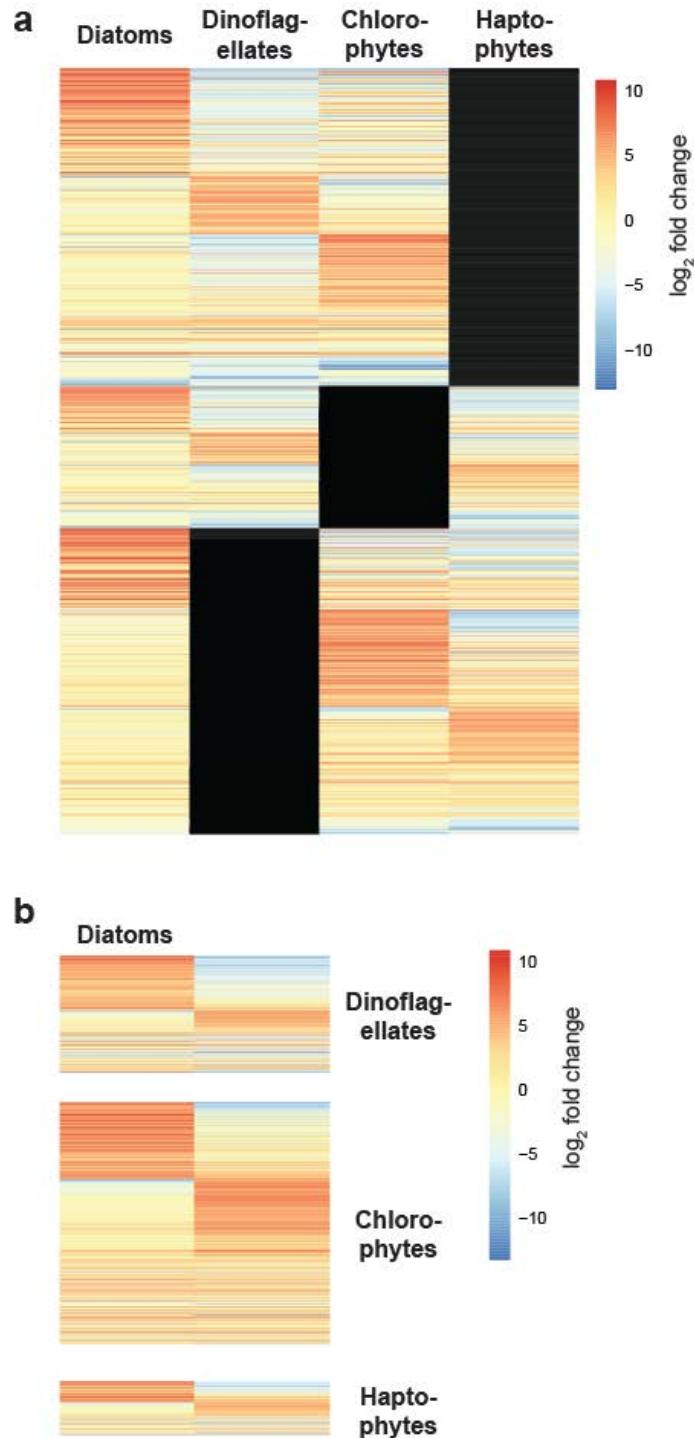


Figure A.9 - Counts and proportions of significantly overrepresented KEGG orthologs within KEGG module class level two groups at 0 and 72 hours: (A) *Pseudo-nitzschia*, (B) *Chaetoceros*, and (C) other diatoms. None refers to no module annotation.

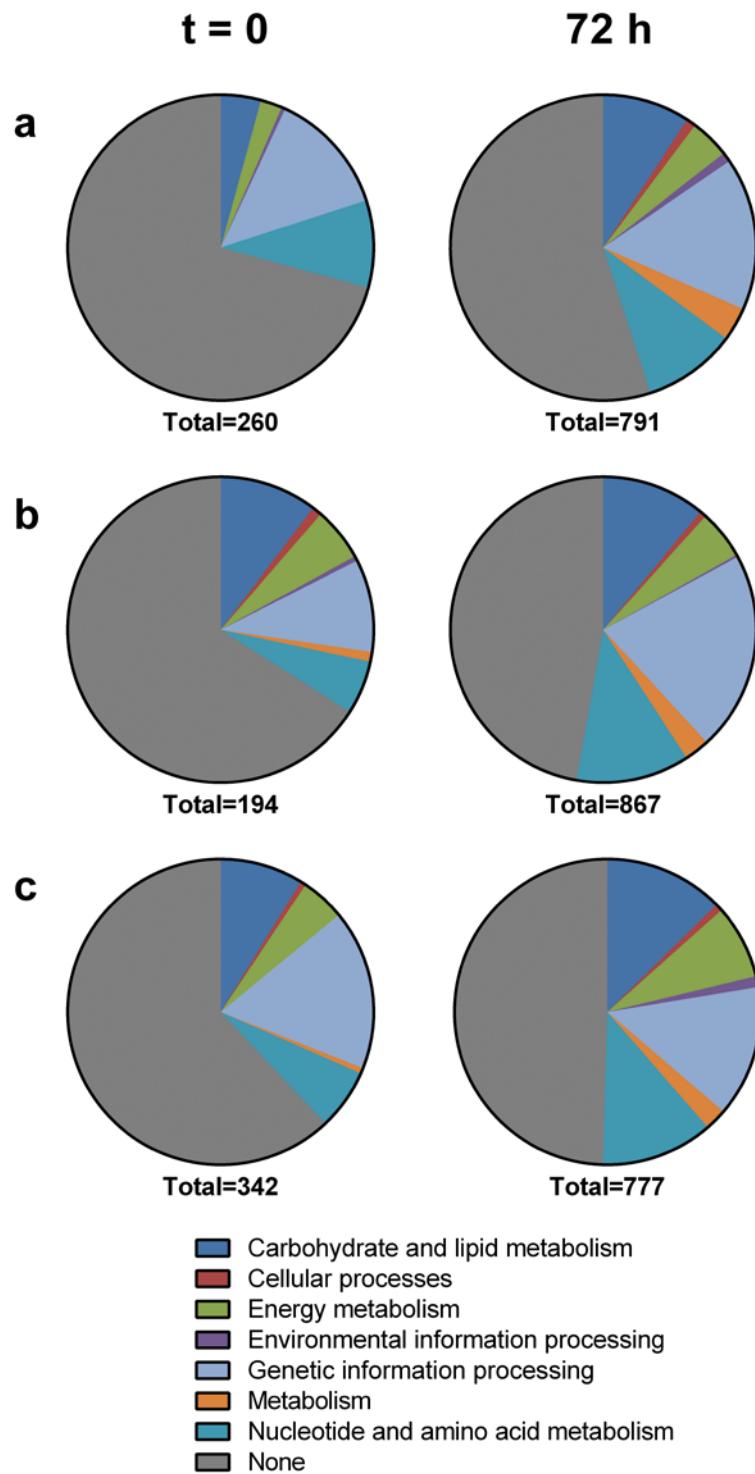


Figure A.10 - Heatmap of nitrogen related genes for *Chaetoceros*, *Pseudo-nitzschia*, other diatoms, dinoflagellates, chlorophytes, and haptophytes at 0 and 72 hours. The color of each box signifies \log_2 fold change while the numbers in each box denotes the average \log_2 counts-per-million of the gene for that group. Darker red (positive fold change) indicates overrepresentation at 72 hours and darker blue (negative fold change) indicates overrepresentation at 0 hours. Rows are sorted by average abundance among the diatoms.

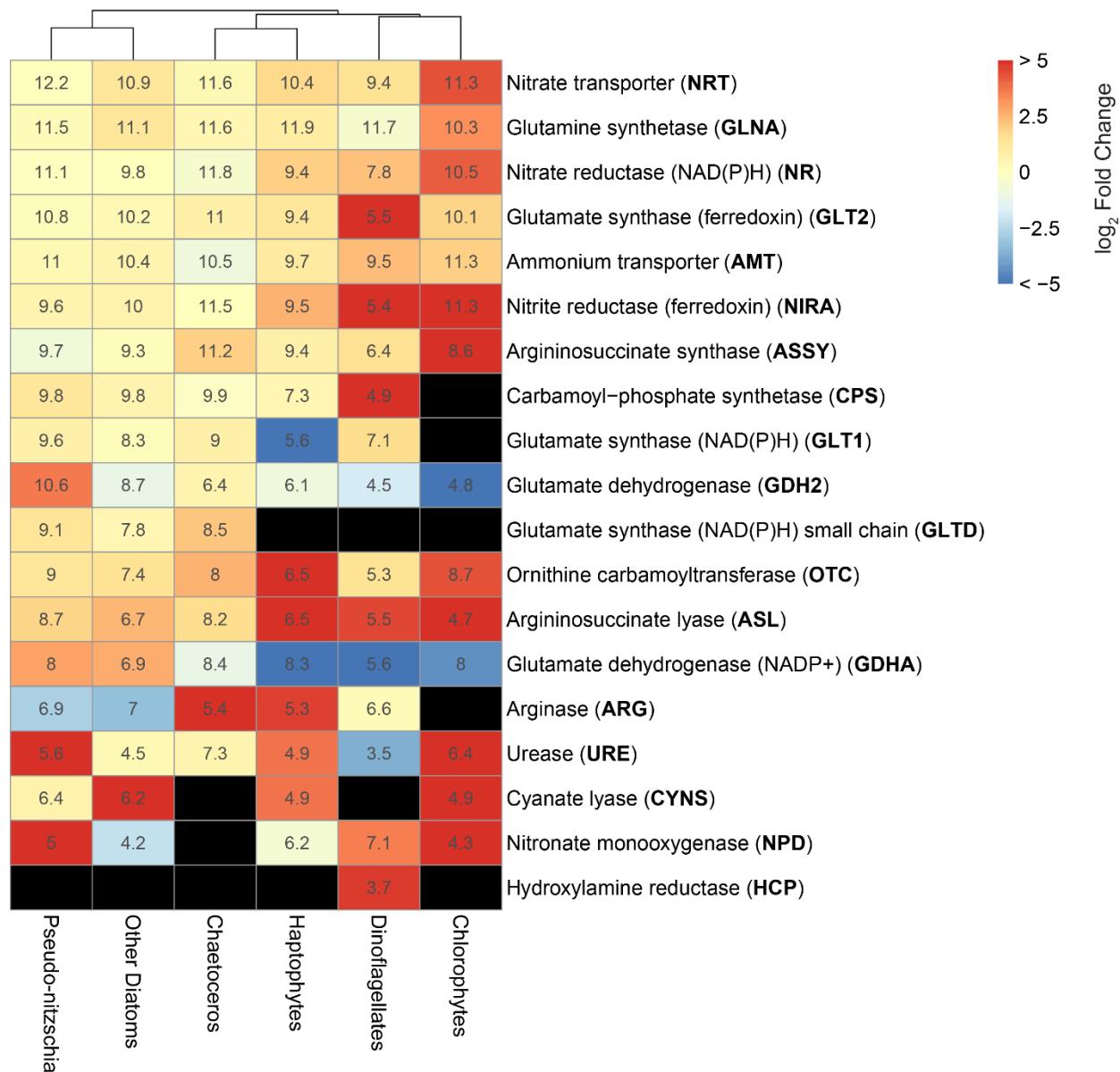


Figure A.11 - Phylogenetic tree of guanylate cyclase soluble subunit beta (GUCY1B). The blue branches, k78.1621874 and k78.943480, denote two full length contigs from our metatranscriptome assembly. Bootstrap values ≥ 50 are indicated at the branch points.

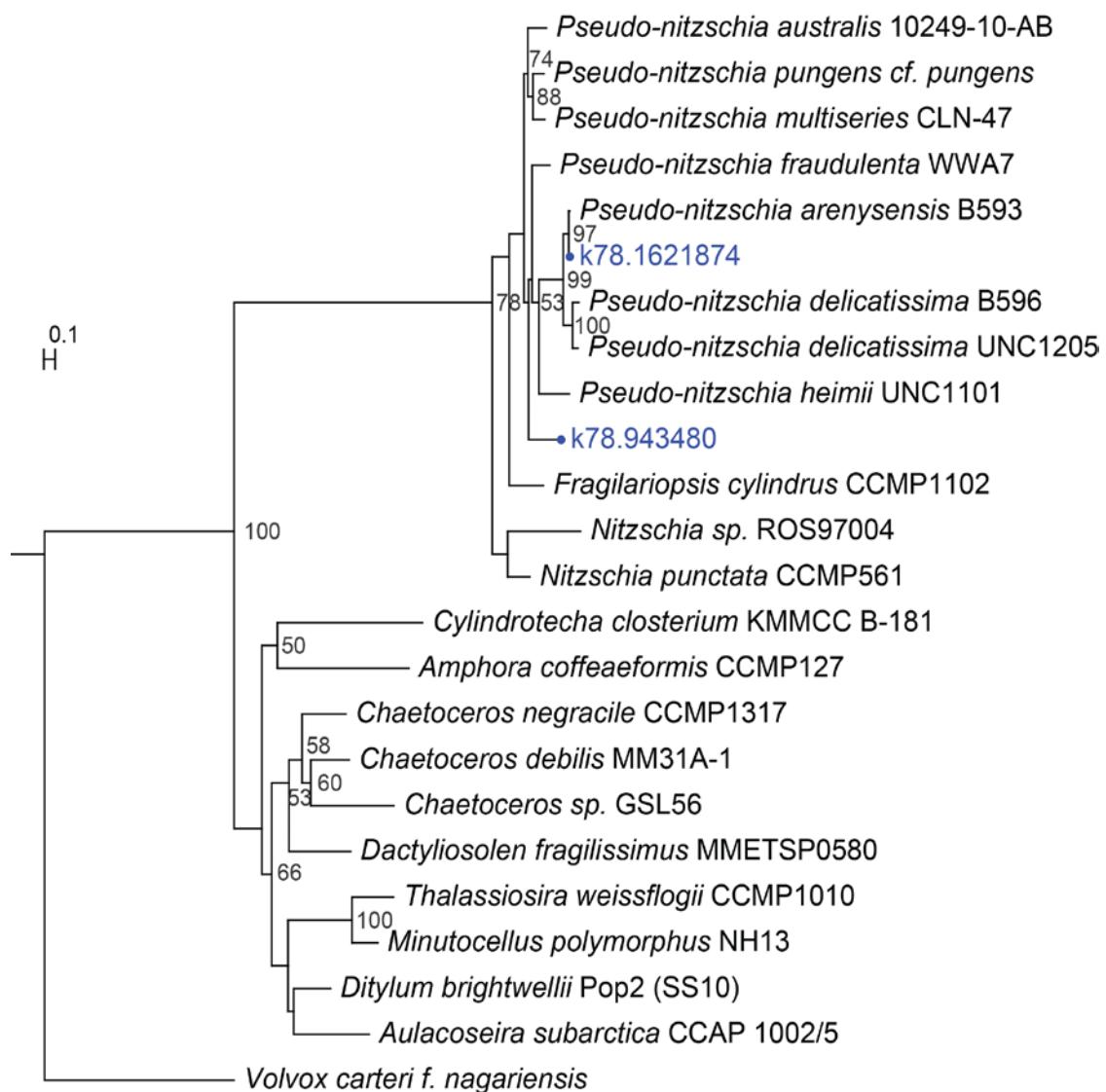


Figure A.12 - Phylogenetic tree of nitrate reductase (NR) in diatoms. Contig IDs are colored according to their taxonomic annotation: *Pseudo-nitzschia* (blue) and *Chaetoceros* (red). Branches exclusively ending in *Pseudo-nitzschia* and *Chaetoceros* on the reference tree are also colored correspondingly.

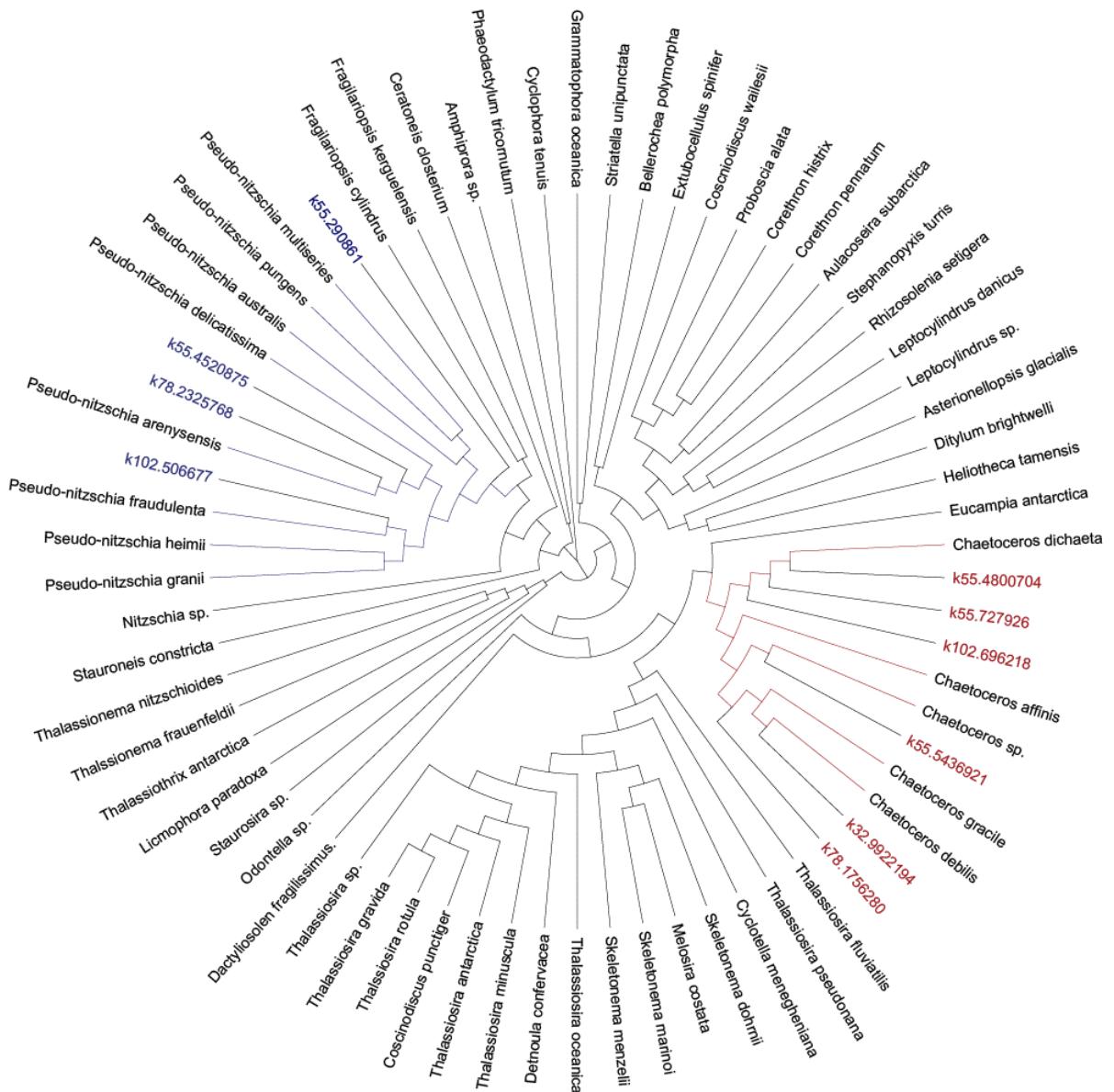


Figure A.13 - Quantitative Metabolic Fingerprint (QMF) depicting relative expression of KEGG modules for each of the four major phytoplankton groups at each time point.

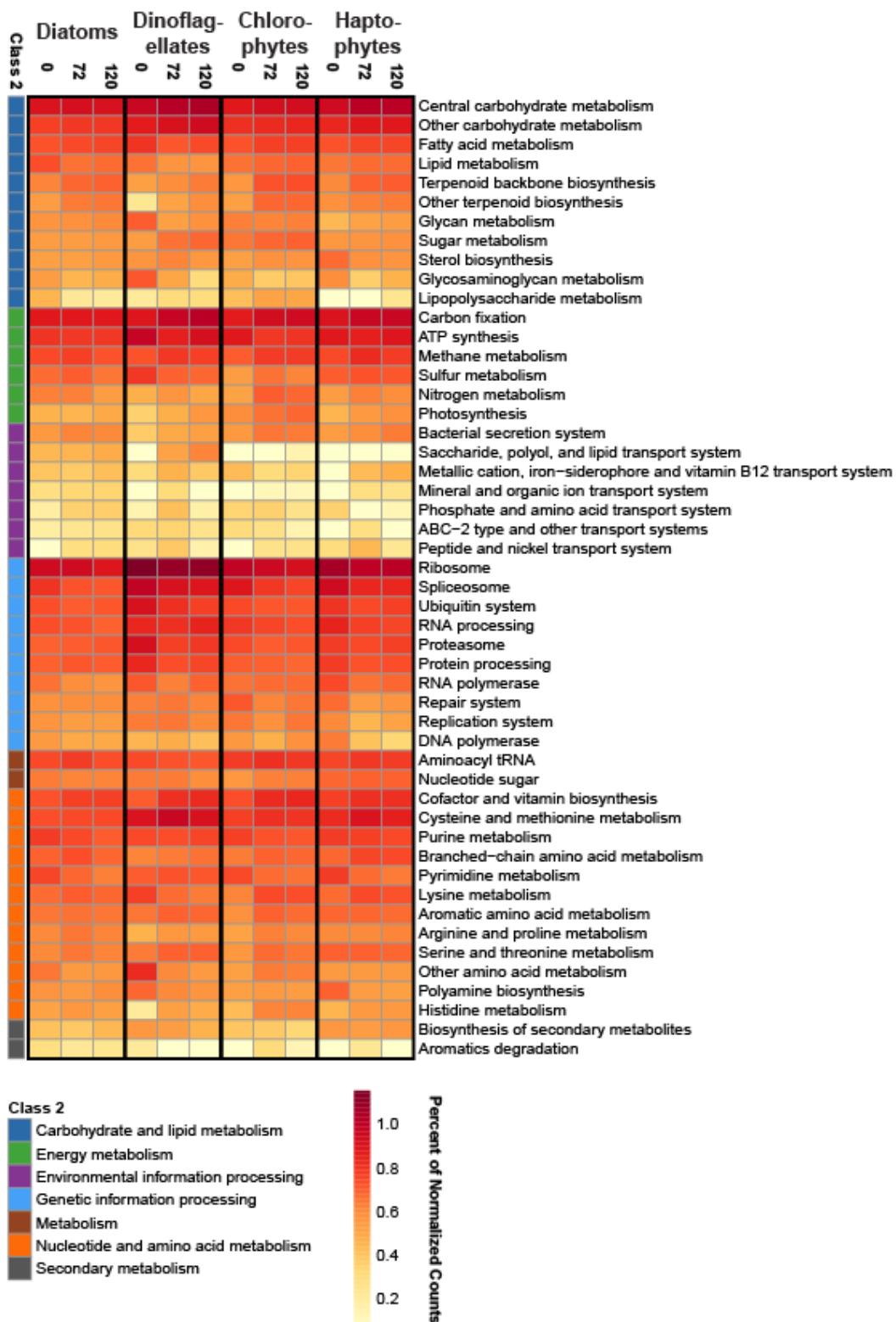


Table A.1 - Summary of environmental RNA sequencing with Illumina HiSeq 2000. Samples with (*) denotes samples used in differential expression analysis. All listed samples were used for assembly and estimating dispersions. PE reads and average length in bases were quantified after trimming adapters and for quality from the original 125 base pair reads. Mapped reads are the number and percentage that have a MAPQ score >= 10 from the total number of paired-end (PE) reads

Sample ID	PE Reads	Avg. Length	Mapped Reads	
T0*	18,385,014	88	4,713,891	25.6%
T72-C*	15,386,368	91	7,647,470	49.7%
T72-Fe	25,308,470	92	12,342,903	48.8%
T72-DFB	22,789,994	87	11,359,979	49.8%
T120-C-A*	26,820,370	89	13,733,582	51.2%
T120-C-B*	19,108,460	91	8,410,083	44.0%
T120-C-C*	16,894,382	90	7,490,886	44.3%
T120-Fe-A	24,207,430	89	11,988,696	49.5%
T120-Fe-B	18,695,494	93	9,252,888	49.5%
T120-Fe-C	20,158,896	92	9,713,759	48.2%
T120-DFB-A	19,130,054	90	9,398,412	49.1%
T120-DFB-B	25,039,482	93	12,054,040	48.1%
T120-DFB-C	27,654,248	87	13,344,671	48.3%
T120-FeDFB-A	18,974,632	91	9,085,579	47.9%
T120-FeDFB-B	17,082,512	89	8,315,221	48.7%
T120-FeDFB-C	27,046,622	85	13,764,934	50.9%
Total:	342,682,428		157,903,103	

Table A.2 - Summary of transcriptome sequencing, assembly, and gene prediction for isolates from the California Upwelling Zone. Paired-end (PE) reads are the total number of reads after sequencing on an Illumina MiSeq. Average read length given in bases is after trimming the adapters and low quality portions from the original 300 base pair reads. The number of contigs and N50 was generated after assembly with Trinity. Predicted proteins and average protein length in amino acids are derived from GeneMark S-T. 18S sequences, raw sequencing files, assemblies, and predicted peptide sequences are deposited in Cyverse (<http://cyverse.org>) under the project name unc_phyto_isolates. Genbank accession numbers for 18S sequences are provided in the table.

Strain ID	Genus species	Genbank Accession Number	PE Reads	Contigs	N50	Predicted Proteins	Avg. Protein Length
UNC1412	<i>Pseudo-nitzschia americana</i>	KX229689	4,475,928	25,917	537	12,753	143
UNC1413	<i>Pseudo-nitzschia fraudulenta</i>	KX229690	3,204,430	33,051	798	21,000	169
UNC1414	<i>Chaetoceros sp.</i>	KX229686	3,479,758	18,534	409	7,571	136
UNC1415	<i>Chaetoceros sp.</i>	KX229687	2,823,336	12,969	422	3,476	136
UNC1416	<i>Chaetoceros decipiens</i>	KX229685	3,440,568	23,461	458	11,657	142
UNC1417	<i>Thalassiosira delicatula</i>	KX229691	7,322,972	50,937	962	32,974	187
UNC1418	<i>Alexandrium tamarensense</i>	KX229684	5,453,530	104,969	530	58,173	149
UNC1419	<i>Emiliania huxleyi</i>	KX229688	5,419,689	47,189	589	22,765	150
Total:			35,620,211			170,369	

Table A.3 - Summary of functional and taxonomic annotation for the environmental RNA assembly. Functional annotation is from the Kyoto Encyclopedia of Genes and Genomes (KEGG), and taxonomic annotation is from MarineRefII (MRII) supplemented with isolate transcriptomes from the California Upwelling Zone. A definition is assigned when a contig has an acceptable BLASTX hit to the database. Ortholog and module annotation is assigned when a contig has a KEGG ortholog or module annotation from one of the top 10 acceptable BLASTX hits. The top group details contigs with annotations, and the second group details contigs with annotations and reads. Percentages are out of the total number of contigs from assembly: 3,151,426.

Annotation	KEGG		MRII + Isolates		Combined	
Definition	1,175,963	37.3%	1,879,961	59.7%	1,155,939	36.7%
Orthology (KO)	686,418	21.8%	-	-	668,275	21.2%
Module (MO)	278,126	8.8%	-	-	274,544	8.7%
Definitions & Reads	1,153,941	36.6%	1,844,730	58.5%	1,124,655	35.7%
KO & Reads	672,698	21.3%	-	-	654,814	20.8%
MO & Reads	271,909	8.6%	-	-	268,393	8.5%

Table A.4 - Cell abundances from microscopic counts ($10^3 \times \text{cells L}^{-1} \pm 1$ standard deviation of the mean of triplicate samples) of four different phytoplankton groups (diatoms, Bacillariophyceae; dinoflagellates, Dinophyceae; green algae, Chlorophyceae; and haptophytes, Haptophyceae) at the initial timpoint and following 72, and 120 hours of incubation. Diatom abundances are shown for the total diatom assemblage (Total), and for members of the genera *Chaetoceros* and *Pseudo-nitzschia*. bd = below detection limit.

Class		T₀	T₇₂	T₁₂₀
Bacillariophyceae	Total	5.8 ± 2.5	250 ± 55	2500 ± 170
	<i>Chaetoceros</i> spp.	0.44 ± 0.63	89 ± 38	980 ± 760
	<i>Pseudo-nitzschia</i> spp.	0.5 ± 0.71	66 ± 55	1200 ± 710
Dinophyceae		6.1 ± 3.8	9.7 ± 4.6	17 ± 24
Chlorophyceae		40 ± 13	250 ± 45	310 ± 85
Haptophyceae		0.90 ± 0.14	bd	bd

REFERENCES

- Alexander H, Jenkins BD, Ryneanson TA, Dyhrman ST (2015a). Metatranscriptome analyses indicate resource partitioning between diatoms in the field. *P Natl Acad Sci USA* **112**: E2182-E2190.
- Alexander H, Rouco M, Haley ST, Wilson ST, Karl DM, Dyhrman ST (2015b). Functional group-specific traits drive phytoplankton dynamics in the oligotrophic ocean. *P Natl Acad Sci USA* **112**: E5972-E5979.
- Armbrust EV, Berge JA, Bowler C, Green BR, Martinez D, Putnam NH et al (2004). The genome of the diatom *Thalassiosira pseudonana*: ecology, evolution, and metabolism. *Science* **306**: 79-86.
- Ashworth J, Coesel S, Lee A, Armbrust EV, Orellana MV, Baliga NS (2013). Genome-wide diel growth state transitions in the diatom *Thalassiosira pseudonana*. *P Natl Acad Sci USA* **110**: 7518-7523.
- Ashworth J, Turkarslan S, Harris M, Orellana MV, Baliga NS (2016). Pan-transcriptomic analysis identifies coordinated and orthologous functional modules in the diatoms *Thalassiosira pseudonana* and *Phaeodactylum tricornutum*. *Mar Genomics* **26**: 21-28.
- 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.
- Bolger AM, Lohse M, Usadel B (2014). Trimmomatic: A flexible trimmer for Illumina Sequence Data. *Bioinformatics* **30**: 2114-2120.
- Bowler C, Allen AE, Badger JH, Grimwood J, Jabbari K, Kuo A et al (2008). The *Phaeodactylum* genome reveals the evolutionary history of diatom genomes. *Nature* **456**: 239-244.
- Chen Y, Lun A, Smyth G (2014). Differential expression analysis of complex RNA-seq experiments using edgeR. In: Datta S, Nettleton DS (eds). *Statistical Analysis of Next Generation Sequence Data*. Springer: New York.
- Cohen NR, Ellis KA, Lampe RH, McNair HM, Twining BS, Brzezinski MA et al (2017). Variations in diatom transcriptional responses to changes in iron availability across ocean provinces. *Front Mar Sci* **4**: 360.
- Edgar RC (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res* **32**: 1792-1797.
- Gong W, Browne J, Hall N, Schruth D, Paerl H, Marchetti A (2017). Molecular insights into a dinoflagellate bloom. *ISME J* **11**: 439-452.
- Gorbunov MY, Falkowski PG (2005). Fluorescence Induction and Relaxation (FIR) technique and instrumentation for monitoring photosynthetic processes and primary production in aquatic ecosystems. In: Van Der Est A (ed). *Photosynthesis: Fundamental Aspects to Global Perspectives: Proceedings of the 13th International Congress on Photosynthesis*. Allen Press, Incorporated. pp 1029-1031.
- Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I et al (2011). Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nat Biotech* **29**: 644-652.
- Han MV, Zmasek CM (2009). phyloXML: XML for evolutionary biology and comparative genomics. *BMC Bioinform* **10**: 1-6.
- Hubbard KA, Rocap G, Armbrust EV (2008). Inter- and intraspecific community structure within the diatom genus *Pseudo-nitzschia* (Bacillariophyceae). *J Phycol* **44**: 637-649.

Kanehisa M, Goto S, Sato Y, Furumichi M, Tanabe M (2012). KEGG for integration and interpretation of large-scale molecular data sets. *Nucleic Acids Res* **40**: D109-D114.

Keeling PJ, Burki F, Wilcox HM, Allam B, Allen EE, Amaral-Zettler LA *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**: e1001889.

Kolber ZS, Prášil O, Falkowski PG (1998). Measurements of variable chlorophyll fluorescence using fast repetition rate techniques: defining methodology and experimental protocols. *BBA-Bioenergetics* **1367**: 88-106.

Leblanc C, Falciatore A, Watanabe M, Bowler C (1999). Semi-quantitative RT-PCR analysis of photoregulated gene expression in marine diatoms. *Plant Mol Biol* **40**: 1031-1044.

Marchetti A, Schruth DM, Durkin CA, Parker MS, Kodner RB, Berthiaume CT *et al* (2012). Comparative metatranscriptomics identifies molecular bases for the physiological responses of phytoplankton to varying iron availability. *P Natl Acad Sci USA* **109**: E317-E325.

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**: 538.

Miyoshi T, Kanoh J, Saito M, Ishikawa F (2008). Fission Yeast Pot1-Tpp1 Protects Telomeres and Regulates Telomere Length. *Science* **320**: 1341-1344.

Nymark M, Valle KC, Hancke K, Winge P, Andresen K, Johnsen G *et al* (2013). Molecular and Photosynthetic Responses to Prolonged Darkness and Subsequent Acclimation to Re-Illumination in the Diatom *Phaeodactylum tricornutum*. *PLoS ONE* **8**: e58722.

Parker CE, Brown MT, Bruland KW (2016). Scandium in the open ocean: A comparison with other group 3 trivalent metals. *Geophys Res Lett* **43**: 2758-2764.

Parsons TR, Maita Y, Lalli CM (1984). *A Manual of Chemical & Biological Methods for Seawater Analysis*. Pergamon: Amsterdam.

Passow U (2002a). Transparent exopolymer particles (TEP) in aquatic environments. *Prog Oceanogr* **55**: 287-333.

Passow U (2002b). Production of transparent exopolymer particles (TEP) by phyto-and bacterioplankton. *Mar Ecol-Prog Ser* **236**: 1-12.

Price NM, Harrison GI, Hering JG, Hudson RJ, Nirel PMV, Palenik B *et al* (1989). Preparation and Chemistry of the Artificial Algal Culture Medium Aquil. *Biol Oceanogr* **6**: 443-461.

Stamatakis A (2014). RAxML Version 8: A tool for Phylogenetic Analysis and Post-Analysis of Large Phylogenies. *Bioinformatics*.

Tang S, Lomsadze A, Borodovsky M (2015). Identification of protein coding regions in RNA transcripts. *Nucleic Acids Res* **43**: e78.

Utermöhl H (1958). Zur Vervollkommnung der quantitativen Phytoplankton-Methodik. *Mitt int Ver theor angew Limnol* **9**: 1-38.

APPENDIX B: CHAPTER 2 SUPPLEMENTARY INFORMATION

Supporting Materials and Methods

Macronutrient Analyses

For the California Upwelling Zone (CUZ) sites, dissolved nitrate + nitrite ($\text{NO}_3^- + \text{NO}_2^-$), phosphate (PO_4^{3-}), and silicic acid (H_4SiO_4) concentrations were measured using a Lachat Quick Chem 8000 Flow Injection Analysis system (Parsons et al 1984) whereas for Line P sites, an Astoria Analyzer was used (Barwell-Clarke and Whitney 1996). 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 (Figure 2.1), interpolated statistical means of nitrate data on a 1° grid for all decades from World Ocean Atlas 2013 (Garcia 2014) was plotted using matplotlib (Hunter 2007) 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 (2012) and Parker et al (2016). 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 an ICP-MS following Milne et al (2010).

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, analysis*.

Phylogenetic Analysis

Internal Transcribed Spacer 1 (ITS1) sequences for known *Pseudo-nitzschia* species (Luisa et al 2004) were compiled from Hubbard, *et al.* (2008, 2014), and Marchetti et al (2008). Reference ITS1 sequences were checked for duplicates with GenomeTools v.1.5.1 (Gremme et al 2013) 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 Chapter 1, Cohen et al (in press), and Moreno et al (2017), and contigs were selected by their annotation as described in Chapter 2. *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 (Edgar 2004) 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 (Stamatakis 2014). Contigs were placed on the tree with pplacer v1.1.alpha18-2-gcb55169 (Matsen et al 2010), and the final trees were visualized with Archaeopteryx v0.9916 (Han and Zmasek 2009).

Supporting 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 (Figure B.2; Marchetti et al 2017). 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* (Chappell et al 2015). 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 (Figure B.2).

Sites C-Low1, C-Low2, and P-Low displayed low iron concentrations and Fe:NO₃ ratios, which were anticipated from their locations (Table 2.1). Although C-Low1 was near a moderate shelf-width area (Figure B.1), relatively low dissolved iron and high micronutrients were still observed (Table 2.1 and

Figure 2.2). This site is offshore the 200 m isobath, and therefore, it is beyond the observed natural boundary for iron-replete high productivity waters (Kudela et al 2006). C-Low2 is located over a narrow shelf region (Figure B.1) explaining the reduced iron. Both of these sites had lower initial dissolved Fe concentrations and Fe:NO₃ ratios well below the theoretical 8 nM:20 μM ratio for near-complete drawdown of nitrate (Bruland et al 2001).

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 (Figure 2.2 and Table B.1). 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 (Figure B.2). 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 (Figure B.2 and Table B.2). 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 (Figure B.1 and B.15). 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 (Figure B.2 and Table B.2).

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 (Whitney et al 2005). The initial conditions are consistent with this dependence as macronutrients were fairly depleted and iron still remained (Figure 2.2 and Table 2.1). After the addition of nitrate to these incubations (Chapter 2 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 possible that other macronutrients (phosphate and silicate) may have become limiting in the iron treatment preventing greater contrast. 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.

Figure B.1 - 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.

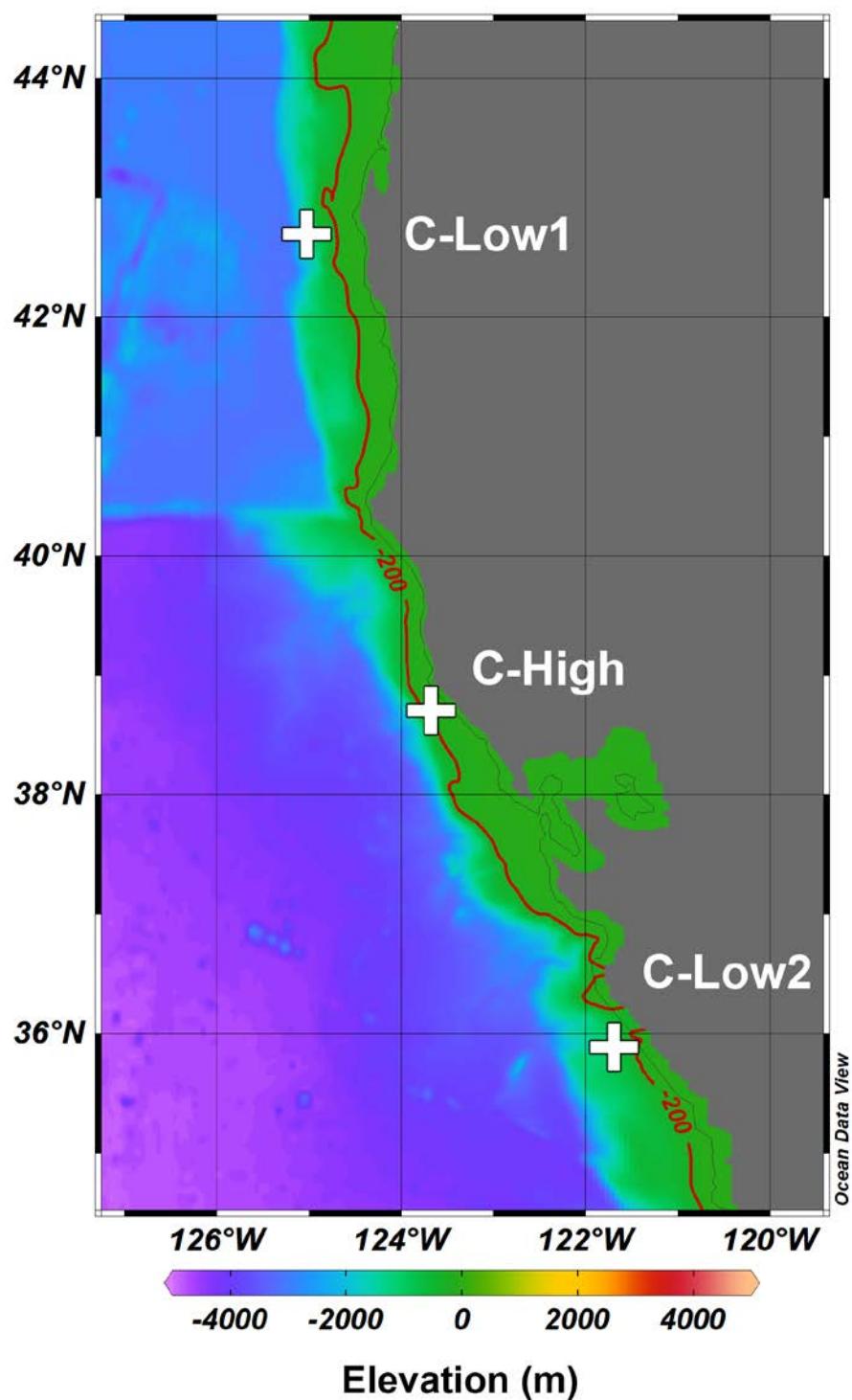


Figure B.2 - Molecular indicators of iron stress or limitation in diatoms as described by Marchetti et al (2017) and Chappell et al (2015). From left to right, the *Pseudo-nitzschia* Iron Limitation Index (*Ps-n* ILI) and flavodoxin (*FLDA1*) and iron starvation induced protein 3 (ISIP3) normalized counts 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.

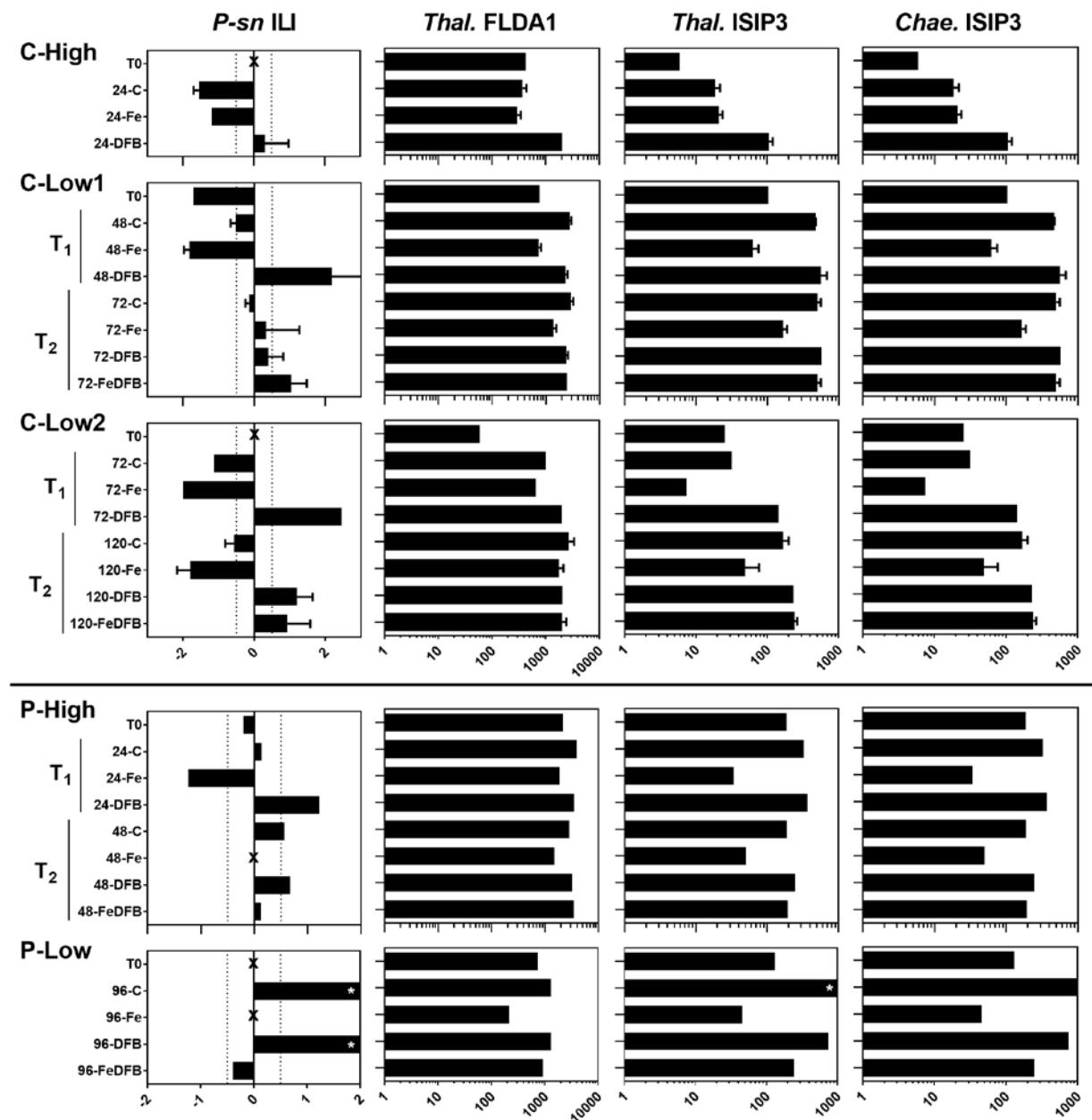


Figure B.3 - Average taxonomic distribution by mapped reads for the whole community (phylum-based) and diatom genera within all diatom assigned reads. Incubations are labelled as follows: control (C), iron addition (Fe), iron removal (DFB), iron addition then removal (FeDFB) and denoted as the first or second time point (T_1 or T_2) where applicable.

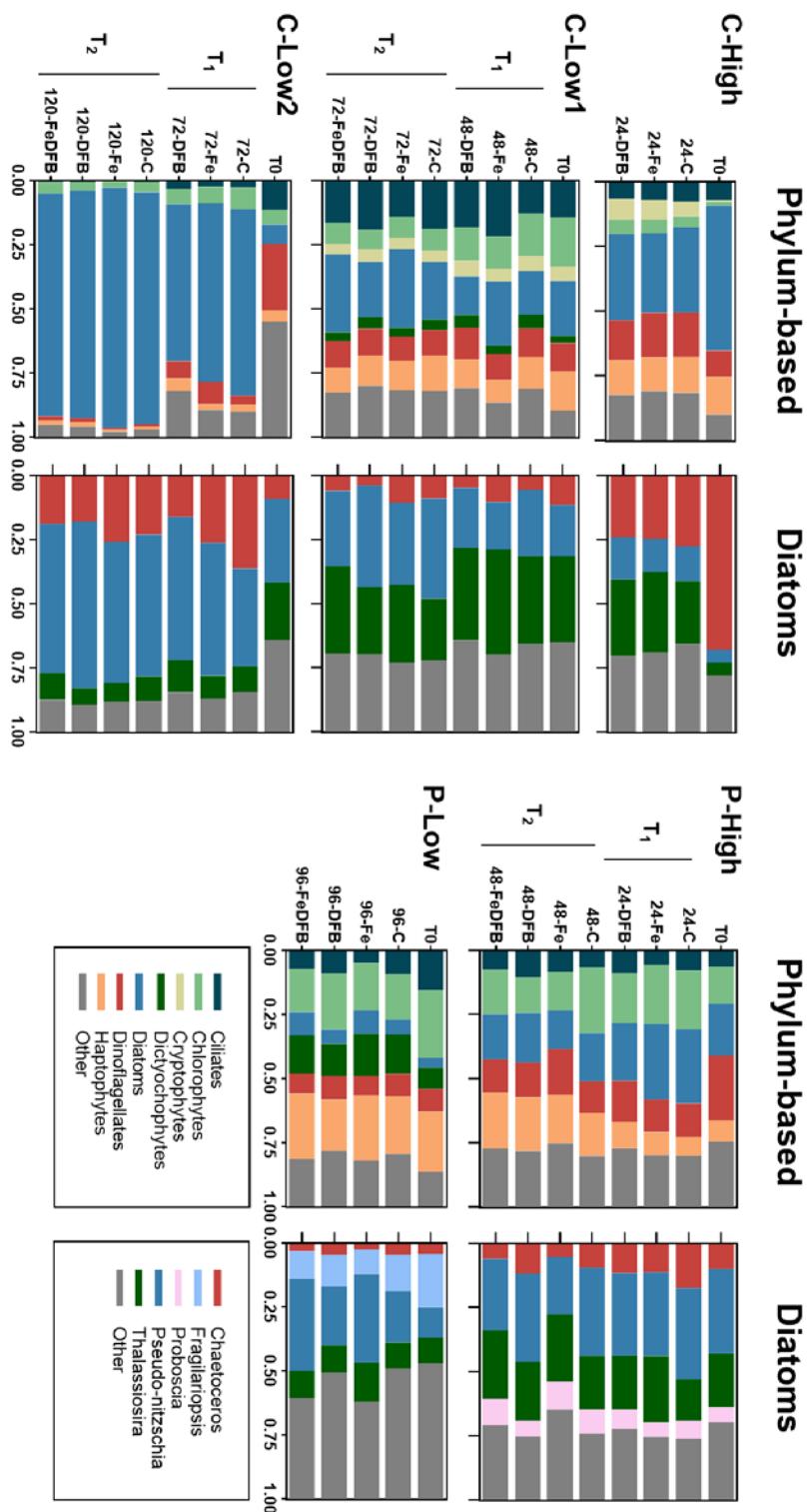


Figure B.4 - Light and synchrotron X-ray fluorescence micrographs of *Chaetoceros* and *Pseudonitzschia* cells from the control and iron (+Fe) incubations at 48 hours at C-Low1.

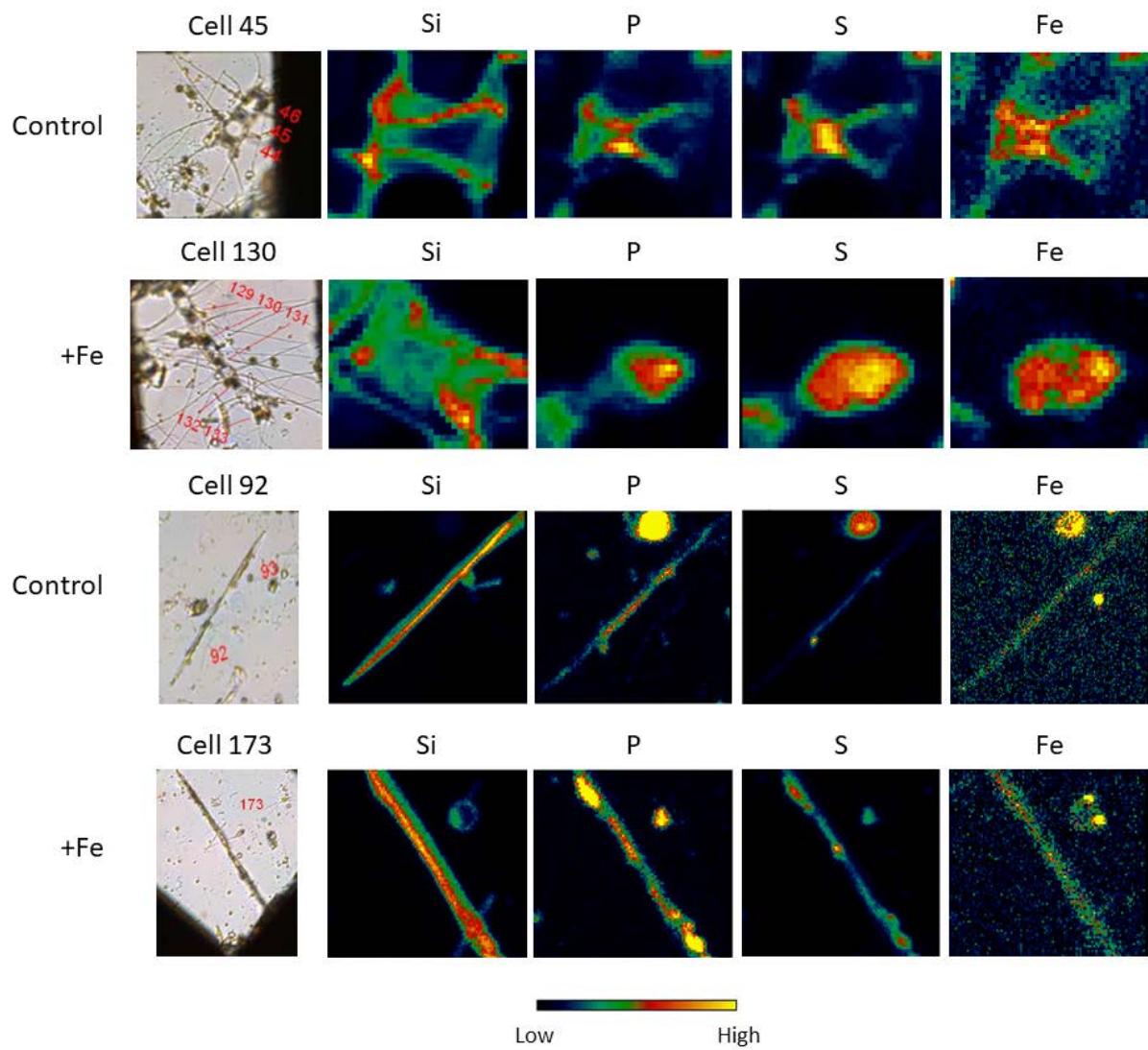


Figure B.5 - Normalized transcript abundances for ferritin (*FTN*), natural resistance-associated macrophage protein (*NRAMP*), and *ZIP1* in *Pseudo-nitzschia* (P), *Chaetoceros* (C), and *Thalassiosira* (T). Dark gray indicates that the gene was not detected.

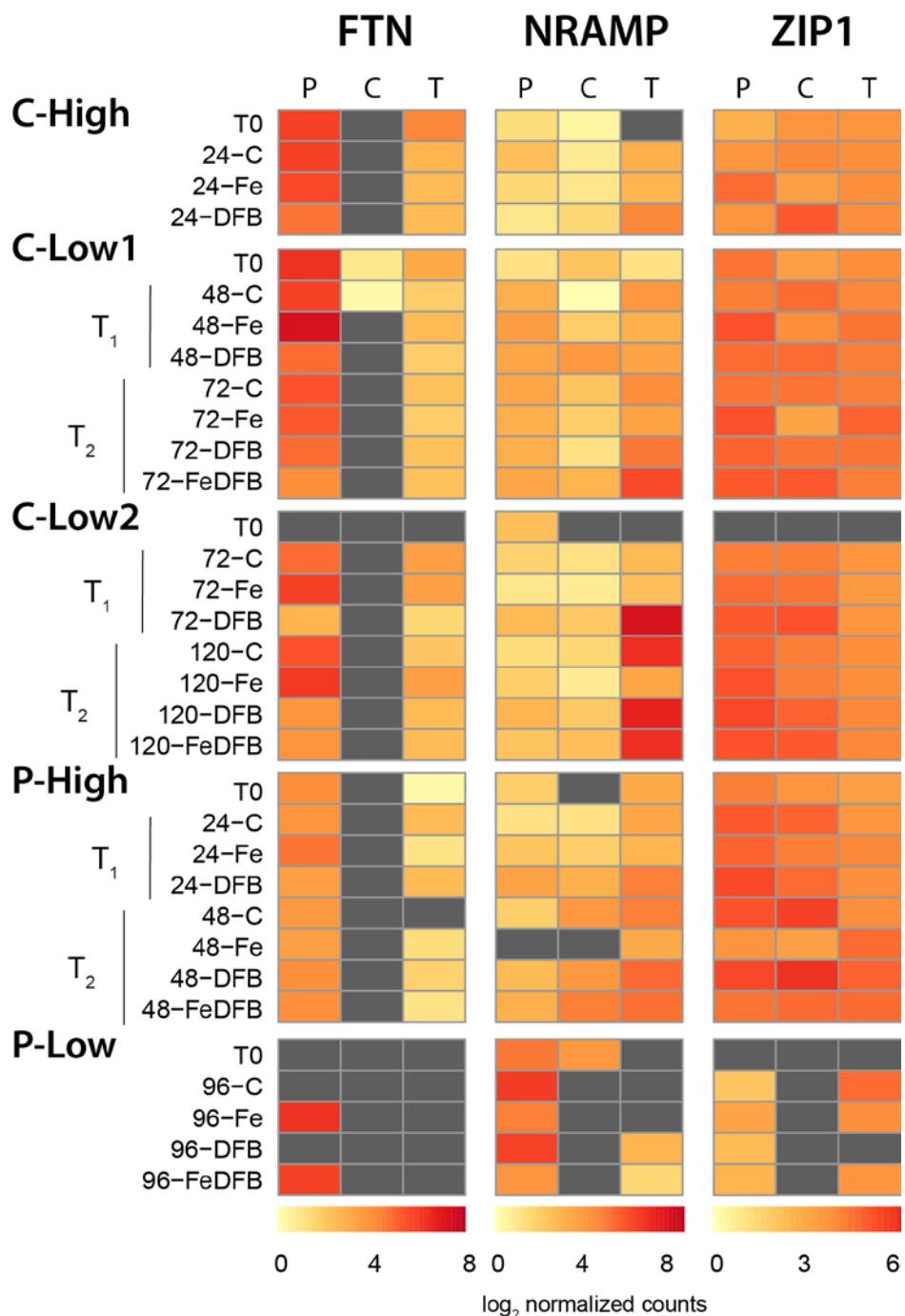


Figure B.6 - Phylogenetic tree of *Pseudo-nitzschia* internal transcribed spacer 1 (*ITS1*) sequences with environmental sequences from the metatranscriptome (blue). Bootstrap values ≥ 50 from the reference sequences are shown. Detection of reads mapping for each site is indicated in blue in the table.

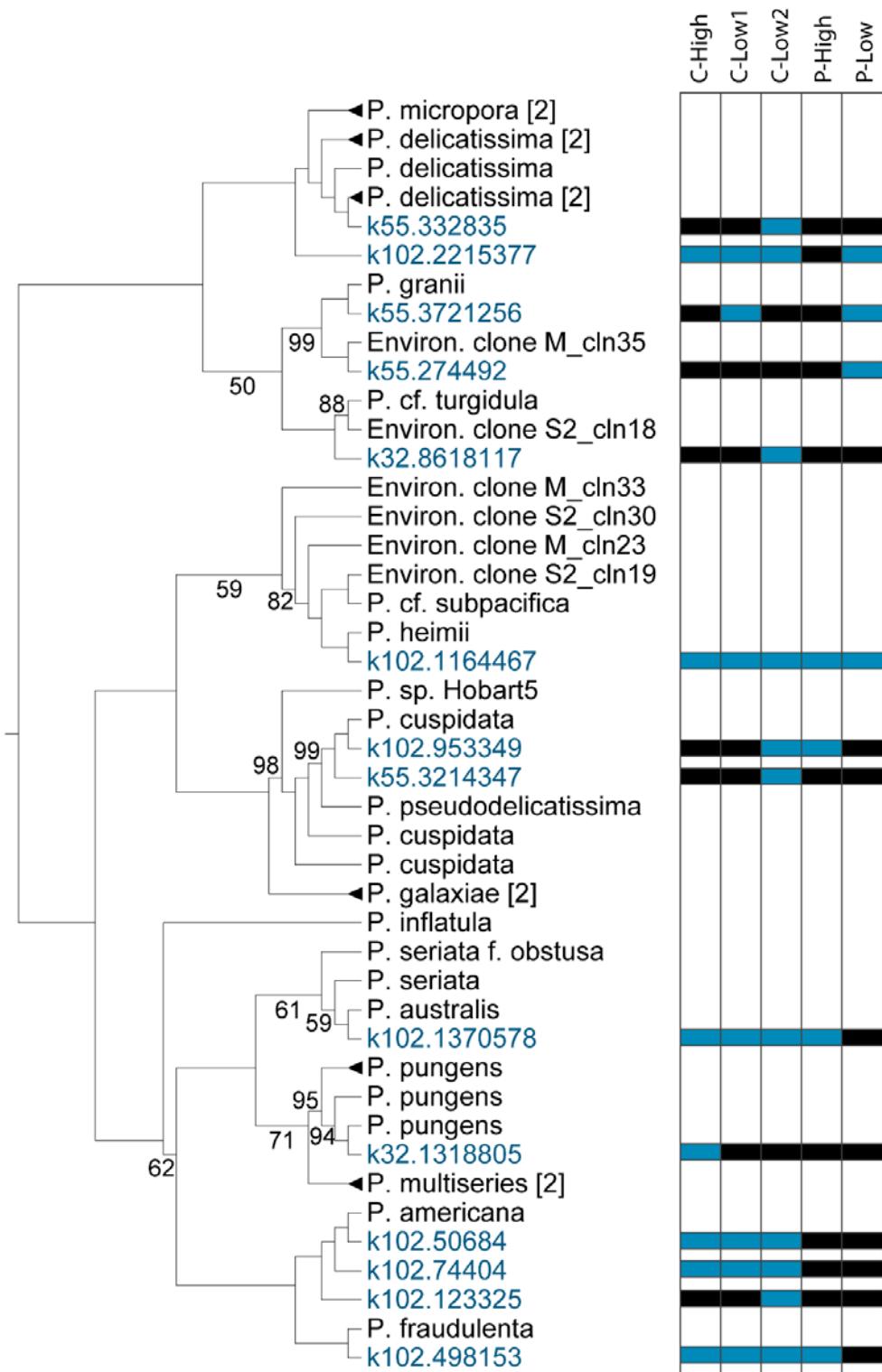


Figure B.7 - Phylogenetic tree of *Pseudo-nitzschia* ferritin sequences with environmental sequences from the metatranscriptome (blue). Bootstrap values ≥ 50 from the reference tree are shown. Detection of reads mapping for each site is indicated in blue in the table.

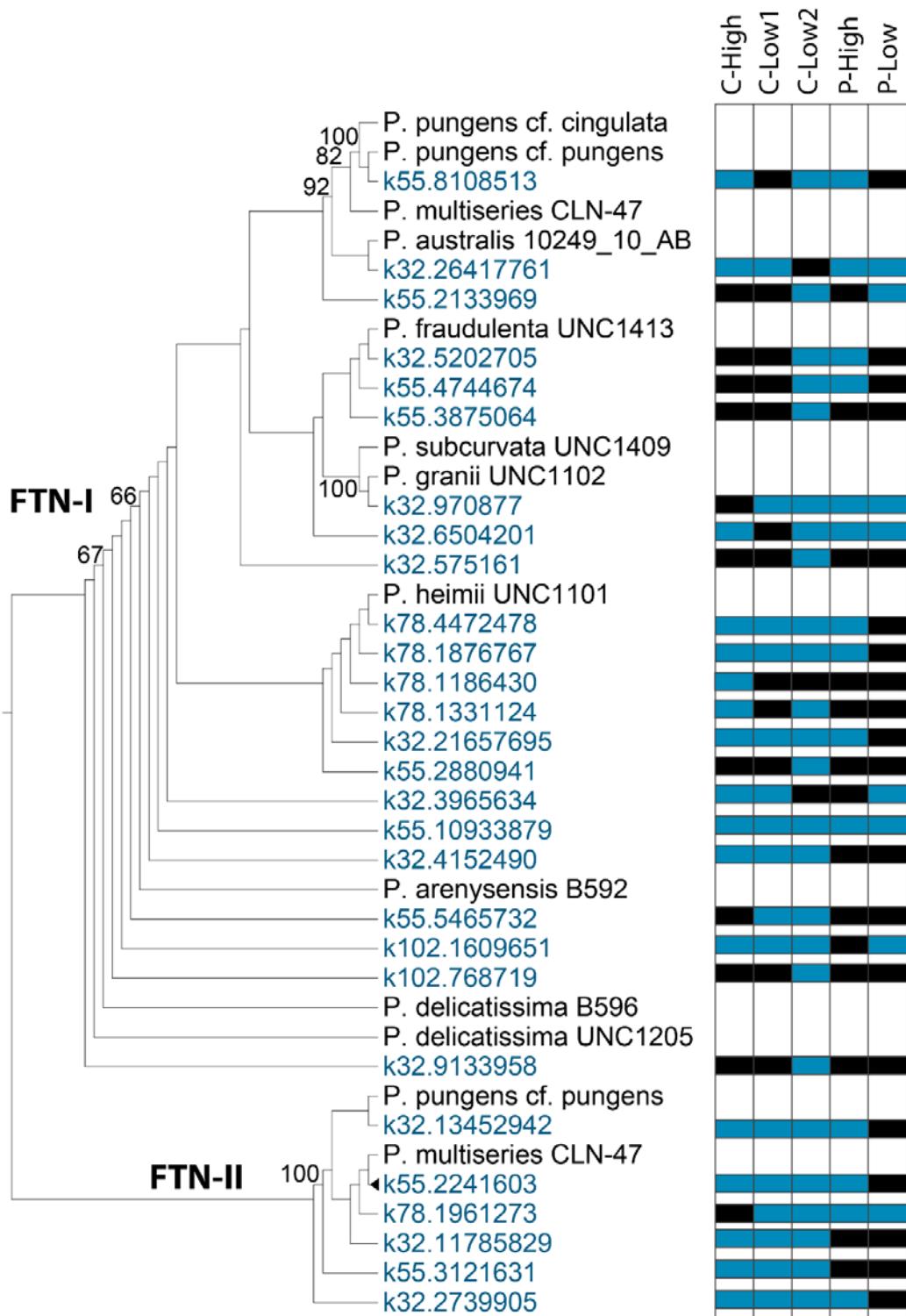


Figure B.8 - Fold changes in *Pseudo-nitzschia FTN* expression (x-axis) versus fold changes in iron quotas ($\mu\text{mol Fe: mol C}$; y-axis) for the Fe or DFB treatments compared to the control treatment.

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 ($p = 0.7551$, $P = 0.0186$).

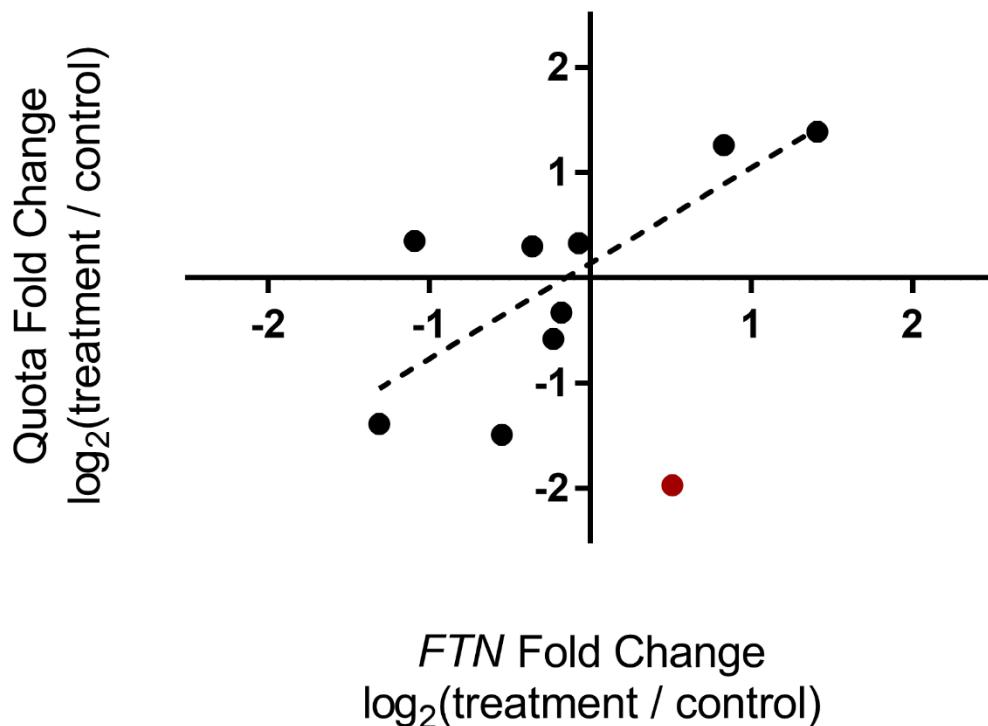


Figure B.9 - Midpoint-rooted phylogenetic tree of diatom ferritins. Branches that do not contain *Chaetoceros* sequences were collapsed with the number of sequences shown. *Chaetoceros*-assigned contigs are highlighted in red. Bootstrap values ≥ 50 are shown.

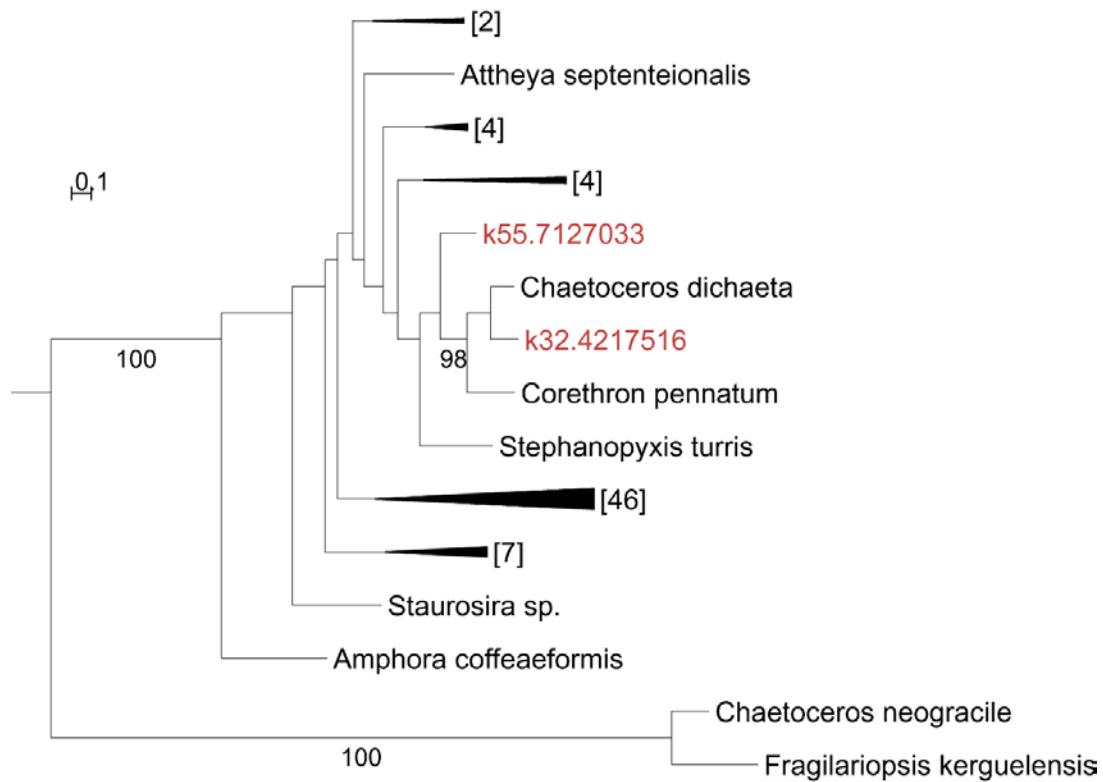


Figure B.10 - Midpoint-rooted phylogenetic tree of diatom ferritins. Branches that do not contain *Thalassiosira* sequences were collapsed with the number of sequence shown. *Thalassiosira*-assigned contigs are highlighted in red. Bootstrap values ≥ 50 are shown.

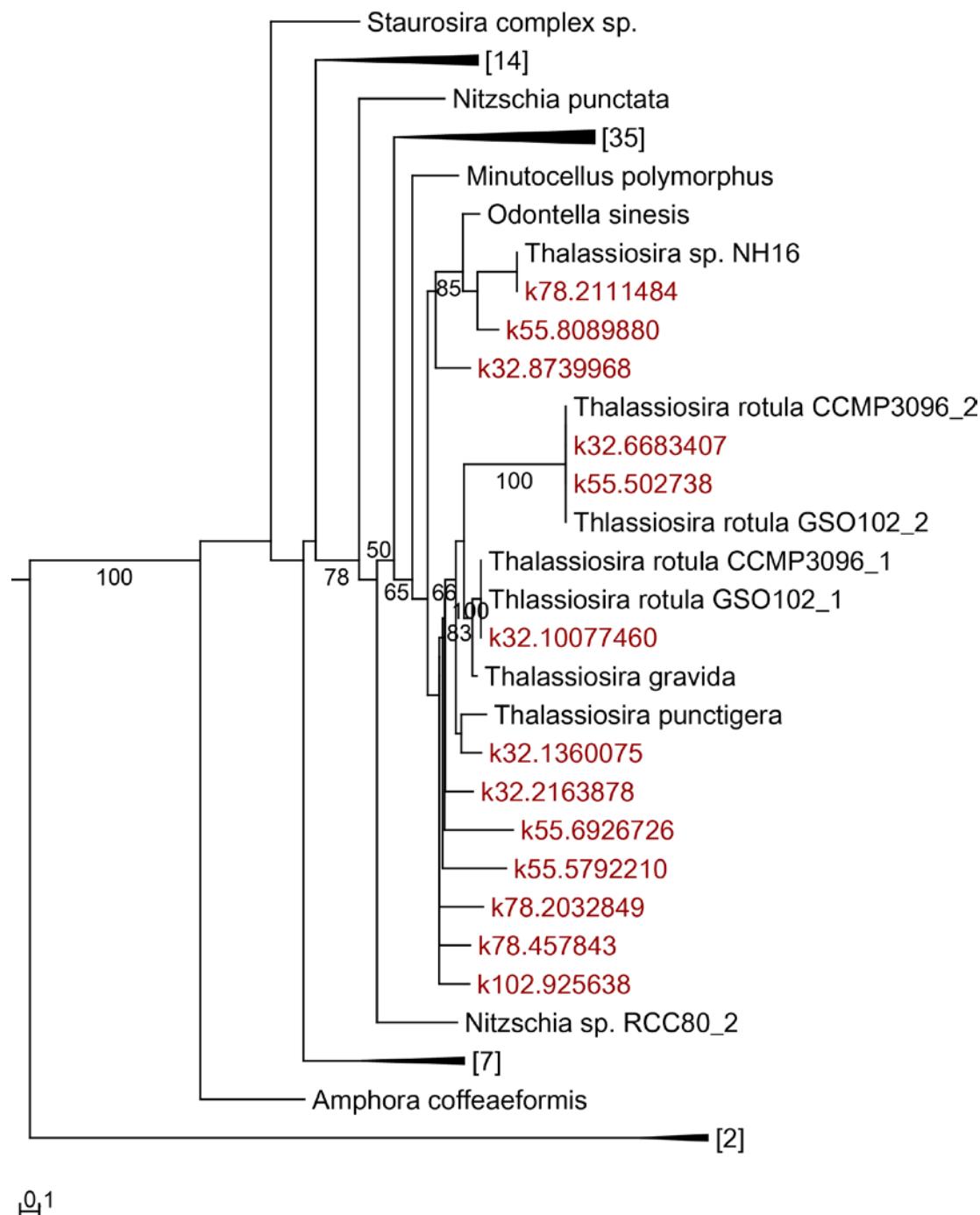


Figure B.11 - Normalized transcript abundances for ferritin (*FTN*) and natural resistance-associated macrophage protein (*NRAMP*) in *Fragilariosis*. Dark gray indicates that the gene was not detected.

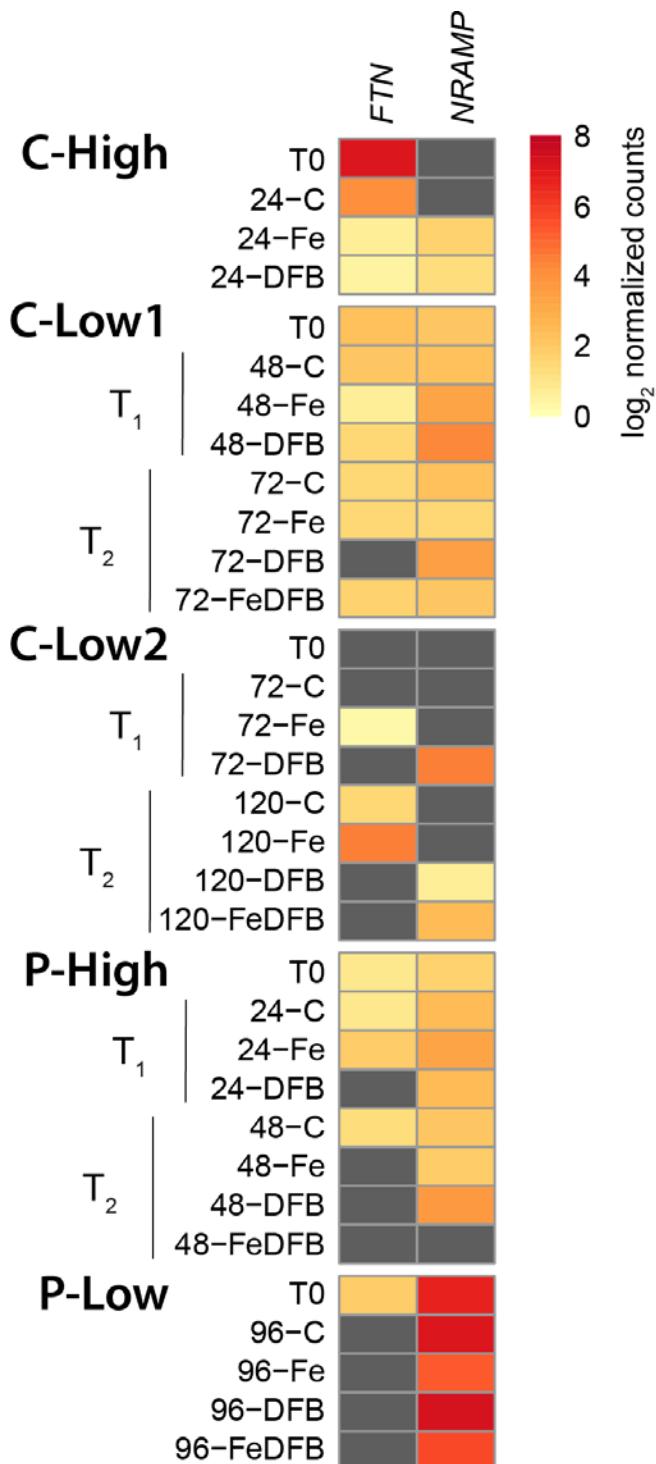


Figure B.12 - Phylogenetic tree of natural resistance-associated macrophage protein (*NRAMP*) homologous genes. Diatom genes (blue) are labeled as Fracy 172829, Psemu 325037, and Thaps 9840 from *Fragilariopsis cylindrus*, *Pseudo-nitzschia multiseries*, and *Thalassiosira pseudonana* with corresponding JGI gene IDs. The red algae, *Cyanidioschyzon merolae* and *Galdieria sulphuraria*, are presented with gene IDs corresponding to entries in Uniprot. Remaining sequences are those presented by Thomine et al (2000) from *Arabidopsis thaliana* (At), *Oryza sativa* (Os), *Saccharomyces cerevisiae* (SMF), *Mus musculus* (Mm), *Deinococcus radiurans* (Dr), *Escherichia coli* (Ec), *Pseudomonas aeruginosa* (Pa), and *Salmonella typhimurium* (St).

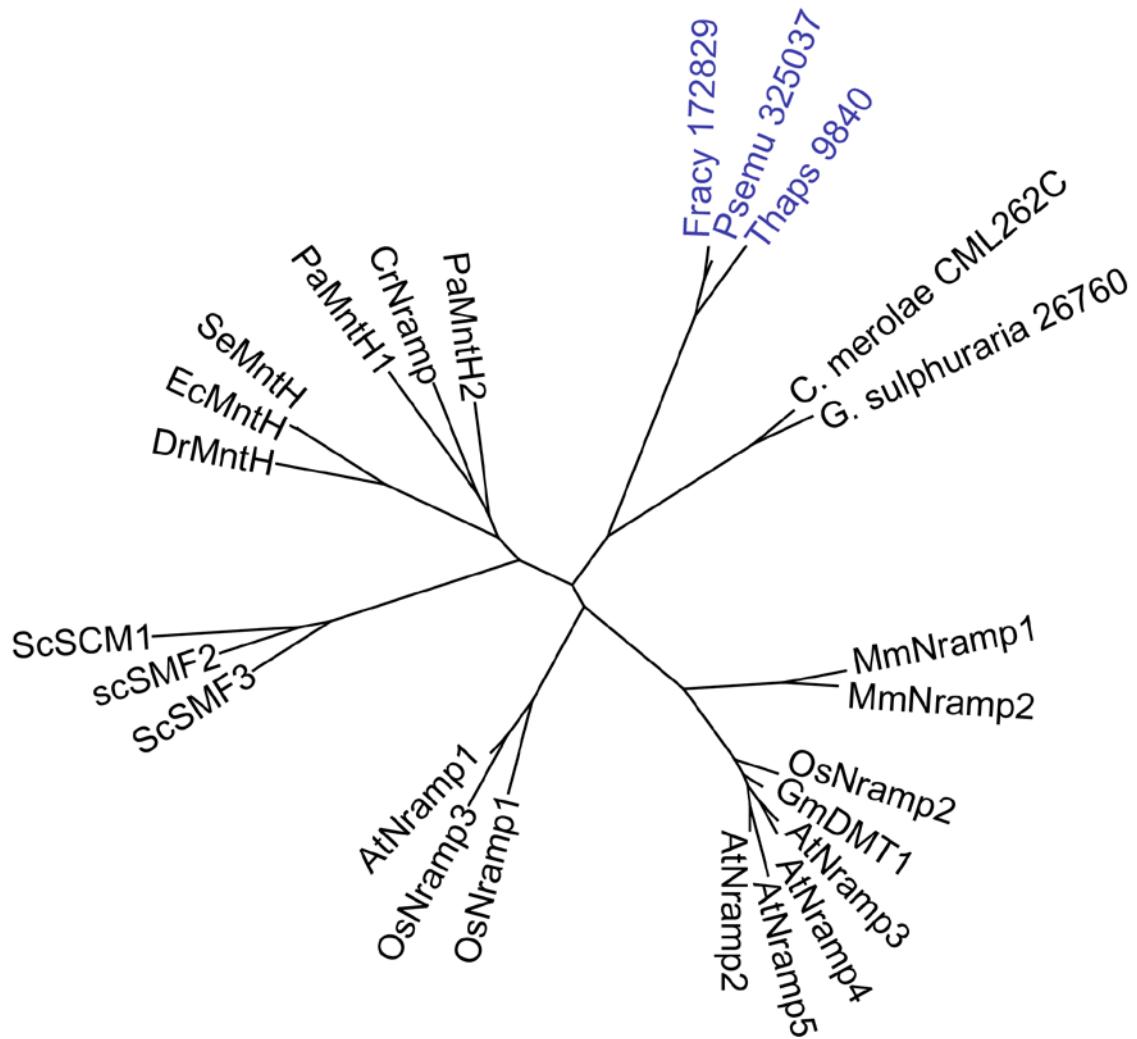


Figure B.13 - Normalized transcript abundances for ZIP7 in *Pseudo-nitzschia* (P), *Chaetoceros* (C), and *Thalassiosira* (T). Dark gray indicates that the gene was not detected.

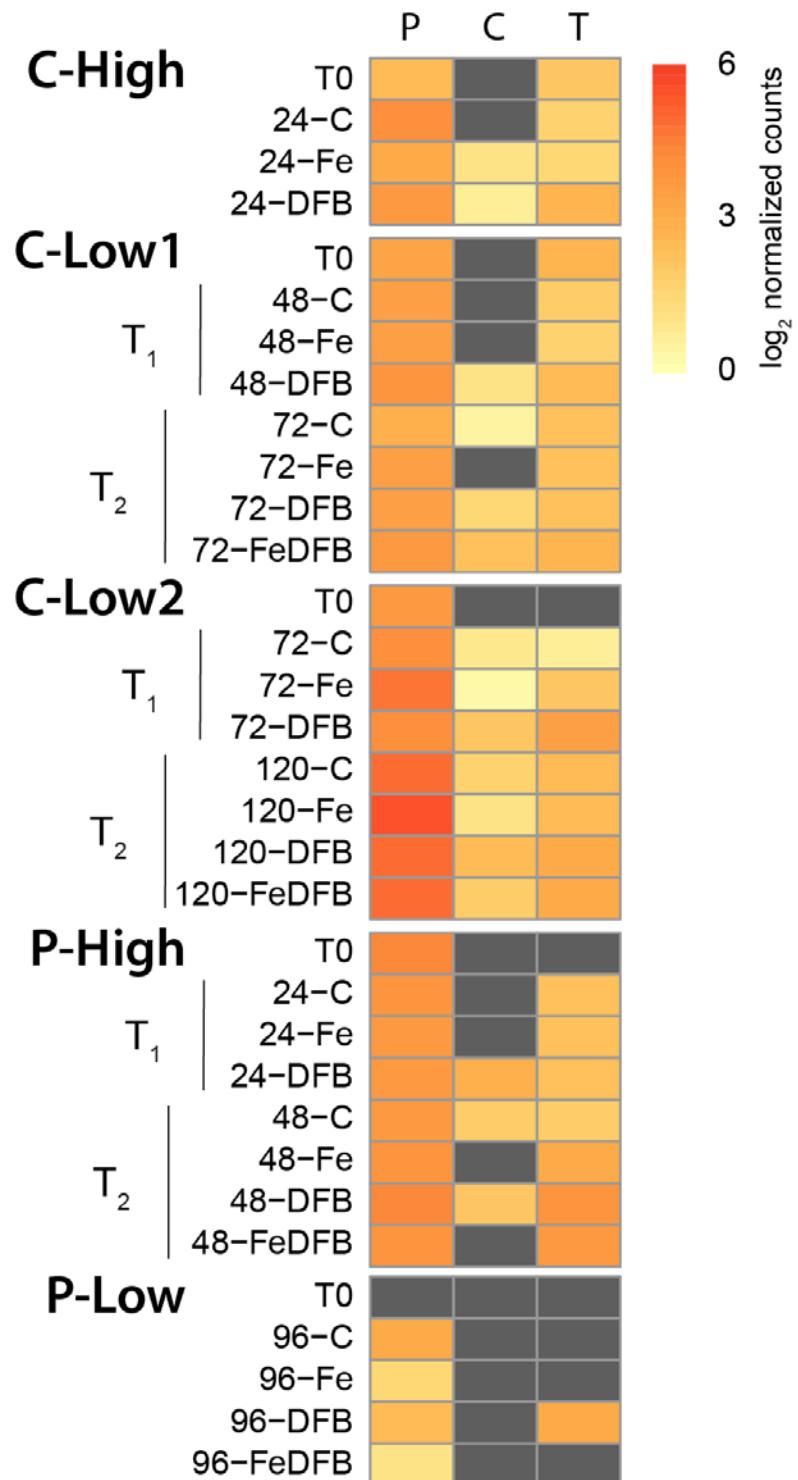


Figure B.14 - Midpoint-rooted phylogenetic tree of ZIP1 homologous genes. Bootstrap values ≥ 50 are shown.

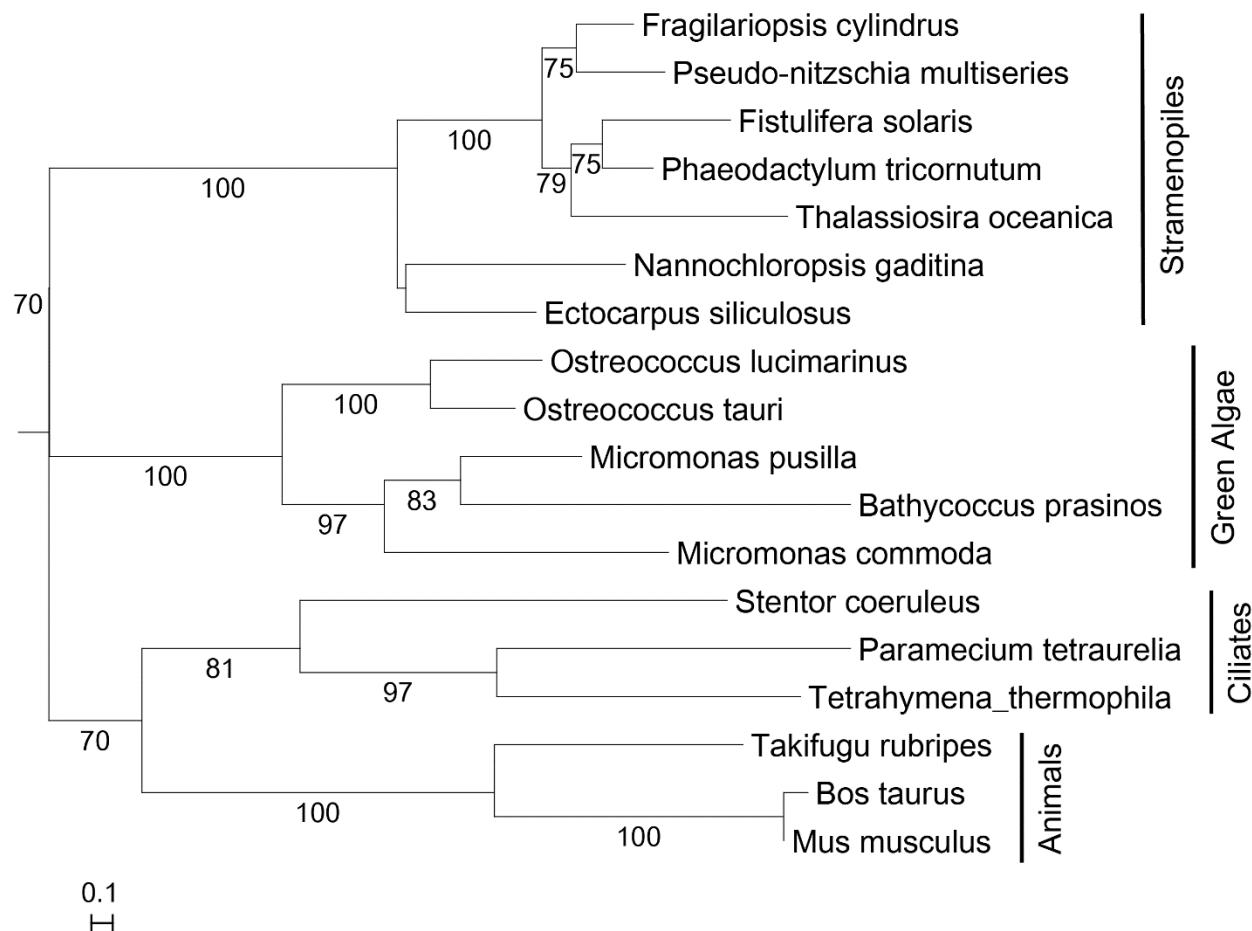


Figure B.15 - Domoic acid (DA) concentrations from C-Low2. Domoic acid was measured for all California Upwelling Zone (CUZ) incubations with the remaining data reported in Cohen et al (2017).

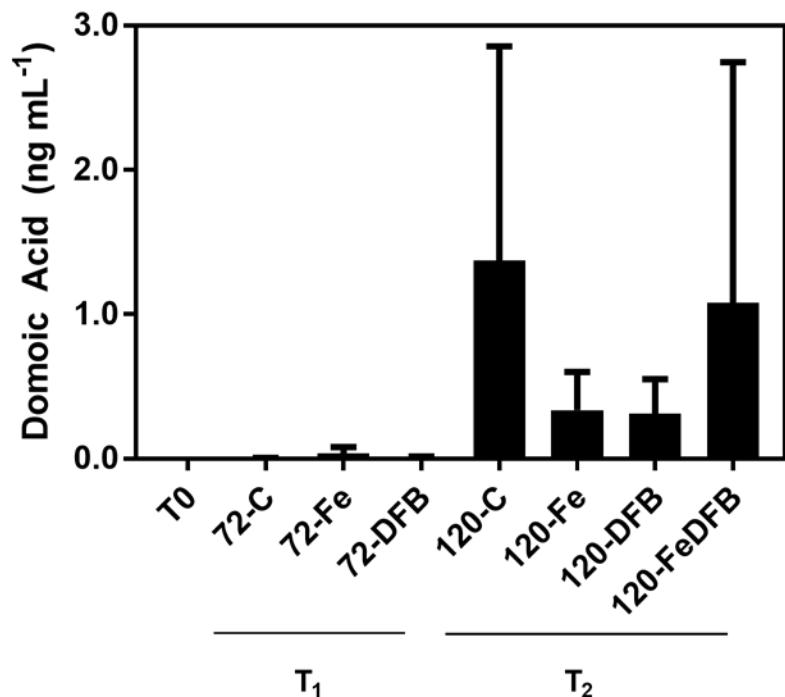


Figure B.16 - Three day average of satellite derived sea surface temperature from 03 June 2014 to 05 June 2014, the day of sampling for C-High (38.7° N, 123.7° W). Data is from the NOAA POES AVHRR satellite courtesy of the NOAA / NESDIS Center for Satellite Applications Research and viewed from the NOAA CoastWatch Browser.

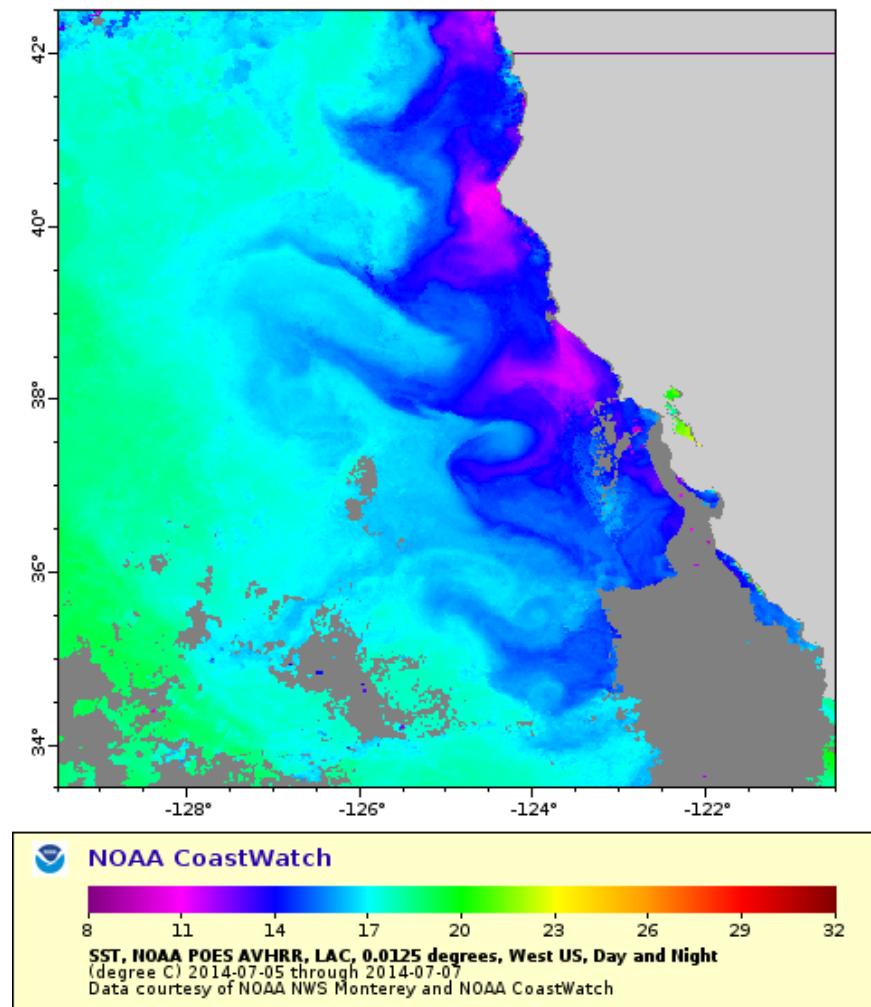


Table B.1 - Time points and initial collection depths for each incubation site.

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 B.2 - 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
NO₃						
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 B.3 - Significance in differential expression of *FLDA1* and *ISIP3* expression for *Thalassiosira* and *Chaetoceros* determined by DESeq2 and shown in Figure B.3. Significance is shown as follows: ns, not significant; *, P ≤ 0.05; **, P ≤ 0.01; ***, P ≤ 0.001; ****, P ≤ 0.0001; or (-), not applicable.

		C-High	C-Low1	C-Low2
<i>Thalassiosira FLDA1</i>				
C vs Fe	T ₁	ns	****	-
	T ₂	-	****	ns
Fe vs DFB	T ₁	****	****	-
	T ₂	-	**	ns
<i>Thalassiosira ISIP3</i>				
C vs Fe	T ₁	ns	****	-
	T ₂	-	****	***
Fe vs DFB	T ₁	****	****	-
	T ₂	-	****	****
<i>Chaetoceros ISIP3</i>				
C vs Fe	T ₁	ns	****	-
	T ₂	-	****	***
Fe vs DFB	T ₁	****	****	-
	T ₂	-	****	****

REFERENCES

- Barwell-Clarke J, Whitney F (1996). IOS nutrient methods and analysis. *Can Tech Rep Hydrol Ocean Sci* **182**: 1-43.
- 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.
- Bruland KW, Rue EL, Smith GJ (2001). Iron and macronutrients in California coastal upwelling regimes: Implications for diatom blooms. *Limnol Oceanogr* **46**: 1661-1674.
- Chappell PD, Whitney LP, Wallace JR, Darer AI, Jean-Charles S, Jenkins BD (2015). Genetic indicators of iron limitation in wild populations of *Thalassiosira oceanica* from the northeast Pacific Ocean. *ISME J* **9**: 592-602.
- Cohen NR, Ellis KA, Lampe RH, McNair HM, Twining BS, Brzezinski MA et al (2017). Variations in diatom transcriptional responses to changes in iron availability across ocean provinces. *Front Mar Sci* **4**: 360.
- Cohen NR, Mann E, Stemple B, Raushenberg S, Jacquot J, Moreno CM et al (in press). Iron storage capacities and associated ferritin gene expression among marine diatoms. *Limnol Oceanogr*.
- Edgar RC (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res* **32**: 1792-1797.
- 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). In: S. Levitus, Mishonov A (eds). *NOAA Atlas NESDIS 76*. p 25.
- 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**: 645-656.
- Han MV, Zmasek CM (2009). phyloXML: XML for evolutionary biology and comparative genomics. *BMC Bioinform* **10**: 1-6.
- Hubbard KA, Rocap G, Armbrust EV (2008). Inter- and intraspecific community structure within the diatom genus *Pseudo-nitzschia* (Bacillariophyceae). *J Phycol* **44**: 637-649.
- 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.
- Hunter JD (2007). Matplotlib: A 2D Graphics Environment. *Comput Sci Eng* **9**: 90-95.
- 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**: 2999-3022.
- 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.
- Marchetti A, Lundholm N, Kotaki Y, Hubbard K, Harrison PJ, Virginia Armbrust E (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**: 650-661.

Marchetti A, Moreno CM, Cohen NR, Oleinikov I, deLong K, Twining BS *et al* (2017). Development of a molecular-based index for assessing iron status in bloom-forming pennate diatoms. *J Phycol* **53**: 820-832.

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**: 538.

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**: 200-207.

Moreno CM, Lin Y, Davies S, Monbureau E, Cassar N, Marchetti A (2017). Examination of gene repertoires and physiological responses to iron and light limitation in Southern Ocean diatoms. *Polar Biol* **4**: 679-696.

Parker CE, Brown MT, Bruland KW (2016). Scandium in the open ocean: A comparison with other group 3 trivalent metals. *Geophys Res Lett* **43**: 2758-2764.

Parsons TR, Maita Y, Lalli CM (1984). *A Manual of Chemical & Biological Methods for Seawater Analysis*. Pergamon: Amsterdam.

Stamatakis A (2014). RAxML Version 8: A tool for Phylogenetic Analysis and Post-Analysis of Large Phylogenies. *Bioinformatics* **30**: 1312-1313..

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**: 4991-4996.

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**: 1055-1067.

APPENDIX C: CHANGES IN GROWTH AND ELEMENTAL COMPOSITION OF A DIATOM AND HAPTOPHYTE THROUGHOUT A SIMULATED UPWELLING CONVEYER BELT CYCLE

Introduction

Phytoplankton communities in coastal upwelling regions contribute to a significant level of global primary production and serve as the base of highly productive food chains (Capone and Hutchins 2013). Within these environments, phytoplankton undergo a cycle in which subsurface populations are transported upwards with upwelled water masses to seed surface blooms and return to depth as upwelling subsides and nutrients are depleted (Wilkerson and Dugdale 2008). Surviving cells at depth may act as seed stocks when upwelling returns. Herein this loop is referred to as the upwelling conveyor belt cycle (UCBC; Figure C.1a).

Diatoms are known to dominate phytoplankton blooms during upwelling (Estrada and Blasco 1985), and respond positively to the shift-up (upwelling) portion of the UCBC. As part of shift-up, it is believed that diatoms respond more quickly to available nitrate compared to other phytoplankton groups (Fawcett and Ward 2011). This response can be explained by frontloading of nitrogen-related genes before the upwelling event occurs (Chapter 1). Coupled to this rapid nitrate uptake, diatoms in natural communities have been observed to transition from relatively high carbon-to-nitrogen (C:N) ratios to those that approach the Redfield-predicted value (Fawcett and Ward 2011, Kudela and Dugdale 2000)(Chapter 1).

As phytoplankton transition throughout the light and nutrient conditions within the UCBC, they likely undergo a number of physiological changes. Importantly, phytoplankton are known to alter their elemental composition based on physiological status (Moore et al 2013) as observed during shift-up. The previously observed high C:N ratios (Fawcett and Ward 2011, Kudela and Dugdale 2000, Chapter 1) imply that the cells faced nitrogen limitation as the upwelled waters aged, and then the cells sank to depth. As these observations are derived from field-based sampling, it is plausible that there was also C-rich detrital material elevating the initial carbon measurement. Moreover, these studies show this rapid change occurring from 0 to 24 hours, but are without observations from within the first day of the shift-up response.

The physiological status of phytoplankton within the UCBC is also likely influenced by the availability of the micronutrient iron. In the California Upwelling Zone for example, iron delivery is primarily

dependent on riverine input and upwelling-driven resuspension of continental shelf sediments (Bruland et al 2001, Hutchins et al 1998). In areas with steep continental shelves, reduced interaction between upwelled waters and the sediment can result in iron limitation which is expected to expand. As upwelling is anticipated to intensify from climate change (Bakun 1990), upwelled nitrate has increased potential to be unmatched by upwelled iron. Furthermore, ongoing changes in ocean chemistry may result in reduced availability of iron to phytoplankton (McQuaid et al 2018, Shi et al 2010)

To examine the physiological responses and resolve changes in elemental composition, lab-based UCBC simulations were conducted with a diatom and haptophyte isolated from the California Upwelling Zone. Samples within 12 hours of returning to light and high nutrients were collected to examine the rapid responsiveness of phytoplankton to shift-up. UCBC simulations were also conducted under low iron conditions to examine the effect of increasing iron stress. We show that the diatom can rapidly respond to the shift-up scenario compared to the haptophyte, and that these responses are impacted by iron stress.

Materials and Methods

Experimental set-up

Two cultures of phytoplankton species from the California Upwelling Zone were used: a diatom, *Chaetoceros decipiens* (UNC1416), and a haptophyte, *Emiliania huxleyi* (UNC1418). Species isolation and identification is described in Appendix A. Cultures were grown in artificial Aquil medium following trace metal clean (TMC) techniques (Marchetti et al 2015, Price et al 1989) at 12°C and 115 µmol photons m⁻² s⁻¹. Macronutrients were added such that silicate would not become limiting for diatom growth (50 µmol L⁻¹ NO₃, 10 µmol L⁻¹ PO₄, 200 µmol L⁻¹ H₄SiO₄). Aquil was chelexed in a TMC room, microwave sterilized, allowed to cool, and then supplemented with filter-sterilized (0.2 µm Acrodisc®) EDTA-trace metals (minus iron) and vitamins (B₁₂, thiamine, and biotin). Premixed Fe-EDTA (1:1) was added separately at a total concentration of 1370 nmol L⁻¹ for the high iron treatments or 3.1 nmol L⁻¹ for the low iron treatments. The resulting iron not complexed to EDTA (Fe') concentrations were 2730 pmol L⁻¹ (high iron) and 6 pmol L⁻¹ (low iron)(Sunda et al 2005). Media were allowed to equilibrate overnight before use.

Cultures were acclimated to medium using the semi-continuous batch culture technique (Brand 1985) at the two iron concentrations in acid-cleaned 28 mL polycarbonate centrifuge tubes with stable growth rates. Cultures were then grown in a 1 L polycarbonate bottle, and when in exponential phase, 5 mL (high iron) or 20 mL (low iron) of each culture was then transferred to triplicate clean 2 L polycarbonate bottles with Teflon tubing. Seawater was stirred continuously at approximately 100 rpm and gently bubbled with clean air. Although sterile techniques were used for all culture work to minimize bacterial contamination, cultures were not considered axenic.

To simulate the upwelling conveyer belt cycle (UCBC), cultures were grown until they reached stationary phase for 3 days to simulate blooming then reaching nutrient limited growth (Figure C.1, Table C.1). At this point bubbling and stirring stopped, and then bottles were moved to a dark (0 $\mu\text{mol photons m}^{-2} \text{s}^{-1}$) section of the incubator to simulate sinking out of the euphotic zone. After 10 days, 500 mL were transferred into 1.5 L of new medium and returned to the light with stirring and bubbling resumed. The 10 day time frame was selected based on satellite observations of timing between upwelling (Chapter 1) and as a point at which significant differences in lag time to return to exponential growth could be observed. Cultures were grown until stationary growth was measured for two days, thus completing the loop. Sampling was performed by first dispensing into acid-washed and Milli-Q rinsed bottles in a laminar flow hood in the TMC room, then aliquoting for this new container. Relative abundance and growth was assessed by regular measurements of blank-corrected raw fluorescence units (RFUs) with a Turner 10-AU fluorometer.

Fv:Fm

Maximum photochemical yield of Photosystem II ($F_v:F_m$) was measured using a Satlantic FIRe (Gorbunov and Falkowski 2005, Kolber et al 1998). Samples were acclimated to low light for 20 min prior to measuring the minimum (F_o) and maximum (F_m) fluorescence yields. Data were blank corrected using microwave-sterilized Aquil medium. The resulting $F_v:F_m$ was derived from the induction profile using a saturating pulse (20,000 $\mu\text{mol photons m}^{-2} \text{s}^{-1}$) for a duration of 100 μs . The average of 50 iterations was obtained.

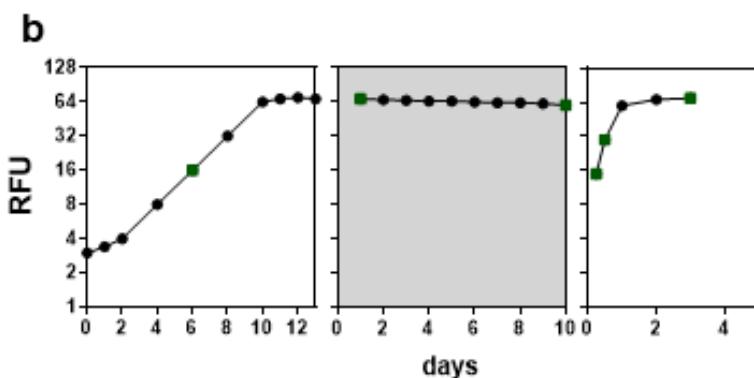
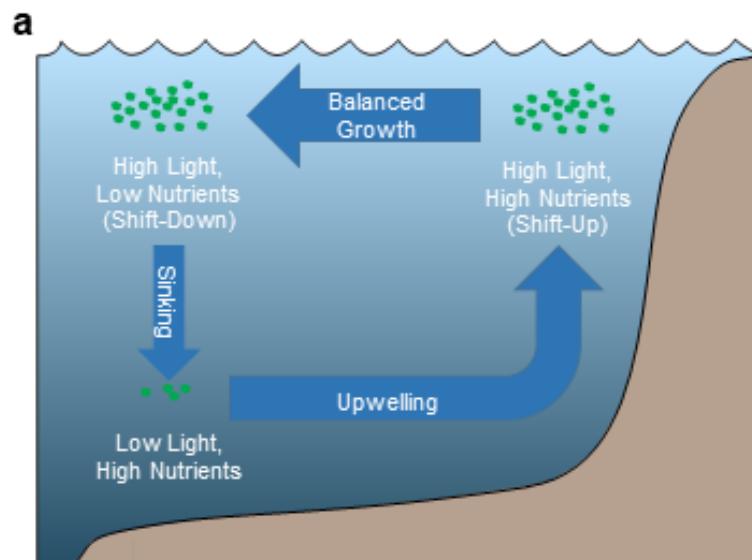


Figure C.1 – Conceptual model of the upwelling conveyor belt cycle. (a) Reproduced from Wilkerson and Dugdale (1987). (b) Modeled growth measurements (raw fluorescence units, RFUs) for a laboratory-based UCBC simulation over the course of 26 days. Green squares indicate sampling time points throughout the experiment (Table C.1).

Time point	Description	Light or Dark
1	Exponential growth 1	Light
2	4 days in stationary phase (1 day in dark)	Dark
3	13 days in stationary phase (10 days in dark)	Dark
4	12 hours after time point 3	Light
5	Exponential growth 2	Light
6	2 days in stationary phase	Light

Table C.1 – Description of sampling time points throughout the UCBC simulations.

Particulate carbon and nitrogen

Particulate carbon (PC) and nitrogen (PN) were obtained by gentle vacuum filtration of 50 mL of water onto a pre-combusted (450° for 5 hours) GF/F filter. Filters were immediately stored in petri dishes at -20°C. Prior to analysis, filters were dried at 65°C for 24 hours then wrapped in tin capsules. Total nitrogen and carbon were quantified with a Costech 1040 CHNOS Elemental Combustion system according to U.S. Environmental Protection Agency Method 440.0 (Zimmermann et al 1997).

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 v7.04.

Results and Discussion

Physiological changes throughout the UCBC under iron-replete conditions

Both *C. decipiens* and *E. huxleyi* showed clear physiological changes to the different conditions in the UCBC when iron was replete (Figure C.2). First, they displayed reductions in photosynthetic efficiency ($F_v:F_m$) between the first exponential growth measurement and stationary phase time points. For *C. decipiens*, these declines were statistically significant ($P < 0.001$) between the first time point and the two time points in the dark. An even greater decrease from 0.376 ± 0.052 to 0.309 ± 0.061 was observed 12 hours after the return to light which is likely a result of non-photochemical quenching. In both species, $F_v:F_m$ values were able to return to higher levels once the cells returned to exponential growth.

Modifications to the C:N ratios are apparent as well. Both species were growing at near-Redfield values during the initial exponential growth phase and showed increases in the C:N ratio during stationary phase. Again in *C. decipiens*, these differences were strongly significant ($P < 0.0001$). These results support that both the stationary phase cells while still in the light and deeper seed communities have altered C:N ratios well above Redfield values, and that the matching field observations (Fawcett and Ward 2011, Kudela and Dugdale 2000, Chapter 1) can be attributed to changes in cells rather than simply C-rich detritus.

Of particular interest is the shift-up response as shown by comparing the samples from 10 days in the dark (T3) with after returning to the light for 12 hours (T4) and reaching exponential growth (T5). *C. decipiens* appeared to return to exponential growth within 24 hours of returning to light and nutrients whereas *E. huxleyi* did not for approximately 48 hours (Figure C.3). Concurrently, the C:N ratio rapidly declined from 20.18 ± 2.13 to 11.89 ± 0.70 within the first 12 hours of returning to light (Figure C.2). By 36 hours after returning to the light as the cells were in exponential growth, their ratio returned to Redfield. Interestingly, *E. huxleyi* maintained a high C:N ratio after the 12 hour return to light, but the ratio dramatically declined to 1.80 ± 0.40 when the cells reached exponential growth (Figure C.2). These results align with the comparatively rapid shift-up response of diatoms to upwelling compared to other phytoplankton groups, and that nitrate uptake rates are likely rapidly increasing to result in this change.

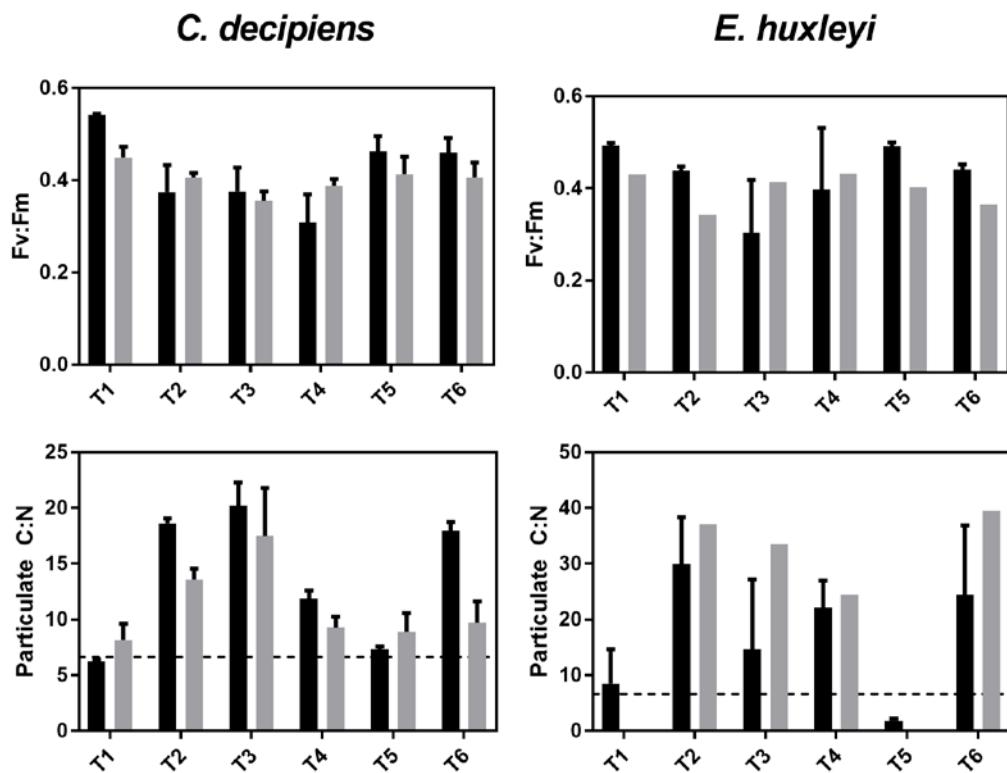


Figure C.2 – $F_v:F_m$ and C:N ratios for *C. decipiens* and *E. huxleyi* throughout the UCBC simulations under iron-replete (black bars) and iron-stressed (gray bars) conditions. In the C:N plots, the Redfield ratio is displayed with a dashed line. For the low iron treatment in *E. huxleyi*, $F_v:F_m$ are only shown for the replicate that completed all 6 time points (Figure C.3), and C and N values were below the detection limit at T1 and T5 for low iron.

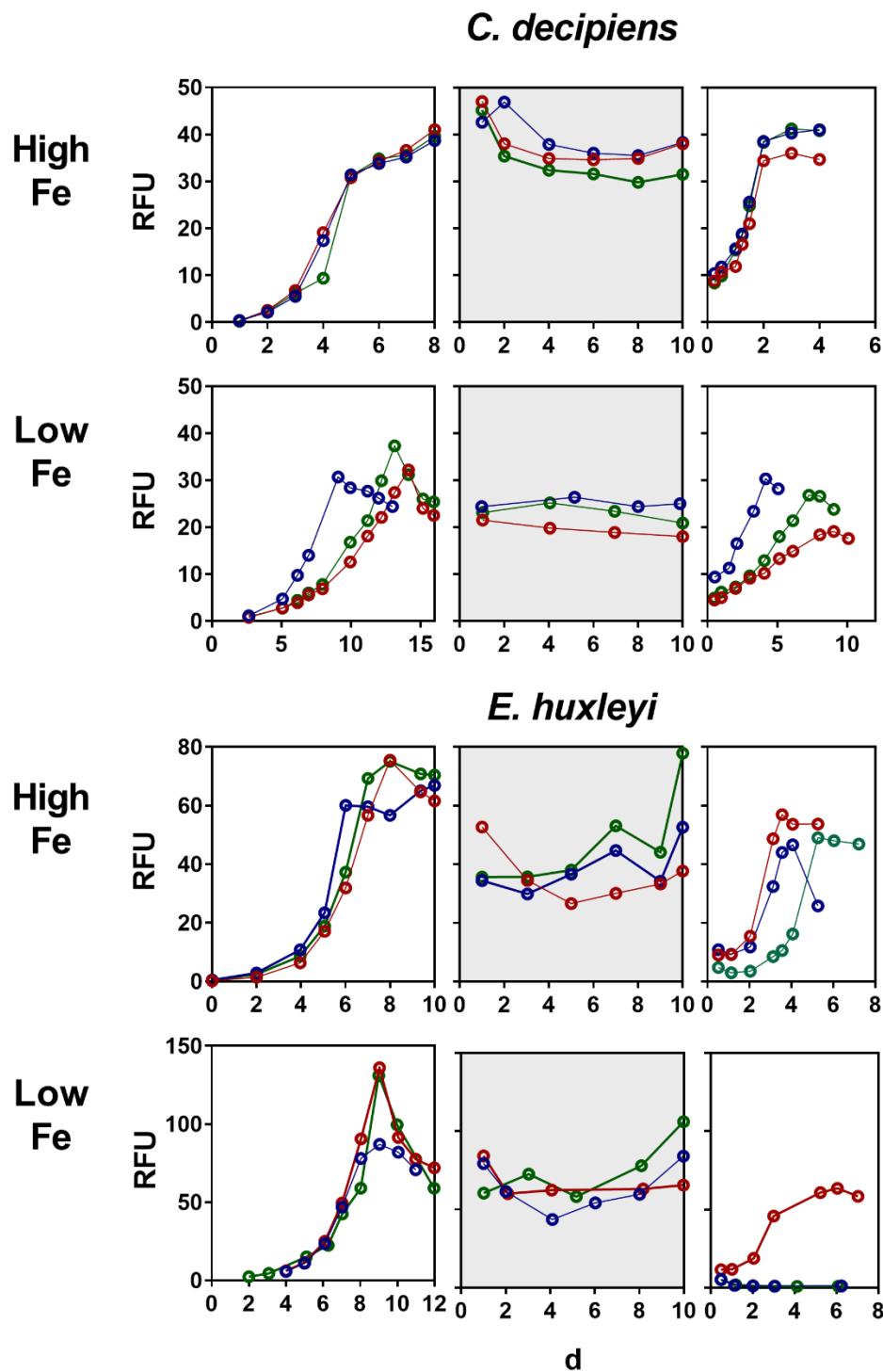


Figure C.3 – Raw fluorescence units (RFUs) of each sample of *C. decipiens* and *E. huxleyi* throughout the UCBC experiments under high and low iron. The first panel in each set is the initial exponential growth and stationary phase in the light, the middle panel is 10 days in stationary/dark phase, and the last is the return to exponential growth and stationary phase in the light.

Effects of iron stress throughout the UCBC

Within the first phase of exponential growth, *C. decipiens* showed significant reductions in growth rate (0.86 d^{-1} to 0.33 d^{-1}). The reduction in growth rates was comparatively smaller for *E. huxleyi* (0.73 d^{-1} to 0.64 d^{-1}), but the iron stressed cells were unable to grow upon returning to light in two out of three replicates (Figure C.3). Importantly, reductions in $F_v:F_m$ were also observed in both showing that the cells exhibited iron stress in the low iron treatments. With $F_v:F_m$ already impacted by iron stress, the variation as a result of light and/or other nutrient stress was comparatively less to when iron was replete.

C:N ratios in *C. decipiens* followed a similar pattern when comparing the iron-limited to iron-stressed conditions (Figure C.2). In early stationary phase (T2 and T6), C:N ratios were significantly lower, but reached similar values after 10 days in the dark. Importantly, C:N ratios still quickly approached Redfield upon returning to the light (T4) indicating that the nitrogen-related shift-up response may not be significantly impacted by low iron conditions. In *E. huxleyi*, C and N values were below the detection limit during exponential growth preventing a comparison for those time points. Similar to the iron-replete conditions, ratios remained high in stationary phase and did not quickly change within the first 12 hours of return to light.

Conclusions

Phytoplankton clearly undergo physiological changes as they transition throughout the different stages of the upwelling conveyer belt cycle. Importantly, these results indicate that previous field observations of high C:N ratios in deeper seed populations (Fawcett and Ward 2011, Kudela and Dugdale 2000, Chapter 1) can be attributed to change in physiological status rather than simply C-rich detrital material affecting results. Furthermore, it is apparent that the diatom shift-up response occurs relatively rapidly with strong differences noted within 12 hours of upwelling. Iron stress did not seem to severely impact the diatom's ability to respond, but did for the haptophyte. Further molecular analysis of these or similar experiments may reveal the underpinnings of these responses in greater detail.

REFERENCES

- Bakun A (1990). Global climate change and intensification of coastal ocean upwelling. *Science* **247**: 198-201.
- Brand LE (1985). Low genetic variability in reproduction rates in populations of *Prorocentrum micans* Ehrenb. (Dinophyceae) over Georges bank. *J Exp Mar Biol Ecol* **88**: 55-65.
- Bruland KW, Rue EL, Smith GJ (2001). Iron and macronutrients in California coastal upwelling regimes: Implications for diatom blooms. *Limnol Oceanogr* **46**: 1661-1674.
- Capone DG, Hutchins DA (2013). Microbial biogeochemistry of coastal upwelling regimes in a changing ocean. *Nature Geosci* **6**: 711-717.
- Estrada M, Blasco D (1985). Phytoplankton assemblages in coastal upwelling areas. In: Bas C, Margalef R, Rubies P (eds). *International Symposium on the Most Important Upwelling Areas off Western Africa*: Barcelona. pp 379-402.
- Fawcett S, Ward B (2011). Phytoplankton succession and nitrogen utilization during the development of an upwelling bloom. *Mar Ecol-Prog Ser* **428**: 13-31.
- Gorbunov MY, Falkowski PG (2005). Fluorescence Induction and Relaxation (FIRe) technique and instrumentation for monitoring photosynthetic processes and primary production in aquatic ecosystems. In: Van Der Est A (ed). *Photosynthesis: Fundamental Aspects to Global Perspectives: Proceedings of the 13th International Congress on Photosynthesis*. Allen Press, Incorporated. pp 1029-1031.
- Hutchins DA, DiTullio GR, Zhang Y, Bruland KW (1998). An iron limitation mosaic in the California upwelling regime. *Limnol Oceanogr* **43**: 1037-1054.
- Kolber ZS, Prášil O, Falkowski PG (1998). Measurements of variable chlorophyll fluorescence using fast repetition rate techniques: defining methodology and experimental protocols. *BBA-Bioenergetics* **1367**: 88-106.
- Kudela RM, Dugdale RC (2000). Nutrient regulation of phytoplankton productivity in Monterey Bay, California. *Deep-Sea Res Pt II* **47**: 1023-1053.
- Marchetti A, Catlett D, Hopkinson BM, Ellis K, Cassar N (2015). Marine diatom proteorhodopsins and their potential role in coping with low iron availability. *ISME J* **9**: 2745.
- McQuaid JB, Kustka AB, Oborník M, Horák A, McCrow JP, Karas BJ et al (2018). Carbonate-sensitive phytotransferrin controls high-affinity iron uptake in diatoms. *Nature* **555**: 534.
- Moore CM, Mills MM, Arrigo KR, Berman-Frank I, Bopp L, Boyd PW et al (2013). Processes and patterns of oceanic nutrient limitation. *Nature Geosci* **6**: 701.
- Price NM, Harrison GI, Hering JG, Hudson RJ, Nirel PMV, Palenik B et al (1989). Preparation and Chemistry of the Artificial Algal Culture Medium Aquil. *Biol Oceanogr* **6**: 443-461.
- Shi D, Xu Y, Hopkinson BM, Morel FMM (2010). Effect of Ocean Acidification on Iron Availability to Marine Phytoplankton. *Science* **327**: 676-679.
- Sunda WG, Price NM, Morel FM (2005). *Algal Culturing Techniques*. Academic Press: Burlington, MA, USA.
- Wilkerson FP, Dugdale RC (1987). The Use of Large Shipboard Barrels and Drifters to Study the Effects of Coastal Upwelling on Phytoplankton Dynamics. *Limnol Oceanogr* **32**: 368-382.

Wilkerson FP, Dugdale RC (2008). *Nitrogen in the Marine Environment (2nd Edition)*. Academic Press: Burlington, MA, USA.

Zimmermann CF, Keefe CW, Bashe J (1997). *Method 440.0: Determination of Carbon and Nitrogen in Sediments and Particulates of Estuarine/Coastal Waters Using Elemental Analysis*. United States Environmental Protection Agency, Office of Research and Development, National Exposure Research Laboratory.